A LOGICAL ANALYSIS OF TOLMAN'S THEORY OF LEARNING

Thesis for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

Joseph Frederick Lambert

1956

This is to certify that the

thesis entitled

A Logical Analysis of Tolman's

Theory of Learning

presented by

Joseph Frederick Lambert

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Philosophy

Henry S. Leonard
Major professor

Date September 14, 1956

O-169

A LOGICAL ANALYSIS OF TOLMAN'S THEORY OF LEARNING

by

Joseph Frederick Lambert

AN ABSTRACT

Submitted to the School of Advanced Graduate Studies of Michigan State University of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Philosophy

1956

Approved: Henry S. Leonard

A LOGICAL ANALYSIS OF TOLMAN'S THEORY OF LEARNING

by

Joseph Frederick Lambert

AN ABSTRACT

Submitted to the School of Advanced Graduate
Studies of Michigan State University of
Agriculture and Applied Science in
partial fulfillment of the
requirements for the
degree of

DOCTOR OF PHILOSOPHY

Department of Philosophy

1956

Approved: Henry S. Leonard

The purposes of this essay are two-fold. First, it will present a logical, or formal construction of Tolman's theory of learning. The construction to be offered in this essay is restricted, by and large, to that version of Tolman's theory as explained and illustrated in his book, _Purposive Behaviorism in Animals and Man_.[1] However, in the formal development of Tolman's theory an attempt will be made to derive certain versions of the so-called latent learning principle. The latent learning issue has its beginnings in the Blodgett experiment of 1929; the experiment is discussed in _Purposive Behaviorism_.[2] It is still a burning issue today. The importance of the latent learning experiments for Tolman's theory of learning cannot be underestimated. For they are generally taken to be the most important of the many experiments which constitute the empirical foundations of Tolman's theory of learning. Investigators have questioned whether (in fact) they can be deduced from Tolman's theory and hence are in doubt as to whether they constitute a test of Tolman's theory of learning.[3] The system in this essay shows that at least some of them are deducible from Tolman's theory.

Secondly, it will present a critical appraisal of

Joseph Frederick Lambert

Tolman's th
the proble
Tolman's th
scope; in
mental iss
methodolog
mental sour

Finall
to be devel
of Tolman's
especially
The presen
tic. This

1. Tolma
Unive
2. Ibid.
3. Feehl
Moder
pps.

Joseph

Tolman's theory of learning. This appraisal is initiated by the problems which arise in the formal construction of Tolman's theory. The appraisal is largely methodological in scope; in general, it does not deal with empirical or experimental issues. Nevertheless, to a certain extent, these methodological points derive their plausibility from experimental sources, for example, the latent learning issue.

Finally, it should be understood that the formal system to be develcped in this essay is not a complete construction of Tolman's theory. Time and space are invulnerable enemies; especially when one is working in uncharted surroundings. The present system though incomplete is more than programmatic. This, I trust, will become clear in the ensuing pages.

---

1. Tolman, E. C., Purposive Behaviorism in Animals and Men, University of California Press, 1932 (Reprinted in 1951).
2. Ibid., pps. 48-50.
3. Meehl, P. and MacCorquodale, K. "Edward C. Tolman", in Modern Learning Theory, Appleton-Century-Crofts, 1954, pps. 127-266.

Joseph Frederick Lambert

A LOGICAL ANALYSIS OF TOLMAN'S THEORY OF LEARNING

By

Joseph Frederick Lambert

A THESIS

Submitted to the School of Advanced Graduate Studies of Michigan State University of Agriculture and Applied Science in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Department of Philosophy

1956

TABLE OF CONTENTS

PREFACE

The Purposes of the Essay

The purposes of this essay are two-fold. First, it will present a logical construction of Tolman's theory of learning. The construction to be offered in this essay is restricted, by and large, to that version of Tolman's theory as explained and illustrated in his book Purposive Behaviorism in Animals and Men.[1] However, in the formal development of Tolman's theory an attempt will be made to derive certain versions of the so-called latent learning principle. The latent learning issue has its beginnings in the Blodgett experiment of 1929; the experiment is discussed in Purposive Behaviorism.[2] The importance of the latent learning experiments for Tolman's theory of learning cannot be underestimated. They are generally taken to be the most important of the many experiments which constitute the empirical foundations of Tolman's theory of learning. Investigators have questioned whether (in fact) they can be deduced from Tolman's theory and hence are in doubt as to whether they constitute a test of Tolman's theory of learning.[3] The system in this essay shows that at least some of these experiments are deducible from Tolman's theory.

---

1. Tolman, E. C., Purposive Behaviorism in Animals and Men, University of California Press, 1932 (Reprinted in 1951). Hereafter, this source will be referred to in the body of this essay as Purposive Behaviorism.
2. Ibid., pp. 48-50.
3. Meehl, P. and MacCorquodale, K., "Edward C. Tolman", in Modern Learning Theory, Appleton-Century-Crofts, 1954, pp. 127-266.

Secondly, this essay will present a critical appraisal of Tolman's theory of learning. This aspect of the essay is a consequence of the problems which arise in the formal construction of Tolman's theory. The appraisal is largely methodological in scope; in general, no attempt is made to extend the scope of this appraisal to empirical or experimental issues. Nevertheless, to a certain extent these methodological points derive their plausibility from experimental sources, for example, from the latent learning issue.

The writer wishes to insist that the formal system to be developed in this essay is not to be taken as being a complete construction of Tolman's theory. Time and space are invulnerable enemies; especially when one is working in uncharted surroundings. The present system though incomplete is more than programmatic. This, I trust, will become clear in the ensuing pages.

A Logical Analysis of Tolman's

Theory of Learning

NOTE: Because of type limitations the comma "," is used in place of the usual single quotation mark.

CHAPTER I

INTRODUCTION

1. The introductory chapter deals with the following points: (1) a general discussion of the technique of formal organization, that is, the technique which will be used in the construction of the system to be presented in this essay, (2) a defense of the formal organization of scientific theories, and (3) a general discussion of the crucial and fundamental elements of Tolman's "system", that is, a general discussion of the function and character of the *intervening variables*. In turn, the first of these points deals with (a) a brief description of a logically organized system, (b) a brief description of the apparatus of *Principia Mathematica*, and (c) the definition of "formal organization"; and the second, with (a) arguments in defense of logical organization to which the present author is opposed and (b) arguments in defense of logical organization to which the present author subscribes.

I

The Logically Organized System

2. A system is a class of statements or propositions. A logically organized system is a system in which every proposition is either a *postulate* or is deducible from a postulate, that is, is a *theorem*. To illustrate, consider the following class of propositions.

(1) Some demanded goal-objects are means-objects.

(2) Where there are organisms which are docile with respect to goal-objects, there are means-objects.

(3) Some goal-objects are means-objects or they are not demanded.

Call this class of propositions 'K'. Let us take (1) as a postulate. Then (2) and (3) follow from (1) and, hence, are theorems. Thus, every proposition in K is either a postulate or a theorem; that is, K is a logically organized system.

3. The above illustration made no use of either logical or mathematical symbols. It is often convenient, however, to translate statements in a natural language into statements in a non-natural or symbolic language. Production of theorems is often facilitated in the sense that symbolically translated statements bring into bright relief those features of propositions which are essential to the deduction of theorems. Since the system to be developed in this essay is not couched in English it will be helpful to see how the logical organization of a system is effected in terms of a symbolic language. For purposes of illustration, K will be the system to be so organized. But first a general outline of the symbolic language to be employed is in order.

Outline of the System of Principia Mathematica

4. The primary symbolic language to be used in this essay

is the system of *Principia Mathematica*.[4] But the system to be constructed in this essay also employs the apparatus and symbolism in Tarski's *Second Axiom System for the Arithmetic of Real Numbers*.[5] However, for present purposes we may put aside discussion of the apparatus of *Real Numbers* until the notion of a functor is discussed in Chapter II. No essential distortion results from this omission in the ensuing account of the logical organization of K by the symbolic apparatus of *Principia*.

5. The system of *Principia* contains various signs which are either formal constants, defined and undefined, or variables. Some of the constants are '.', '~', 'V', '⊃', '≡', '=', and '∃'. Greek and English letters are the variables, for example, 'x', 'y', 'φ', 'ψ', 'p', and 'q'. These constants and variables, by different arrangements, form different logical laws. For example, 'p⊃p' is such a law. Within this set of laws there are certain laws called postulates which express certain properties of the basic constants, for example, basic constants such as '.'and '~'. In addition, Principia contains a set of rules for the deduction of certain laws called theorems from the postulates; 'p⊃p' is a theorem in *Principia*.

---

4. Whitehead, A., and Russell, B., *Principia Mathematica*, Cambridge U. Press, Second Edition, 1950. Hereafter, this system will be called *Principia*.
5. Tarski, A., *Introduction to Logic*, Oxford Press, 1941, pp. 217-218. Hereafter, this system will be referred to as *Real Numbers*.

6. Since it is the logical laws in <u>Principia</u> which guarantee the validity of the steps in a proof, the logical organization of K may be construed as an <u>application</u> of <u>Principia</u> to K. What I mean by "application" is what logicians sometimes mean by "interpretation". Unfortunately the word "interpretation" is often used ambiguously. It has a narrow and broad sense. In its narrow sense "interpretation" simply refers to the various meanings assigned to the formal constants. In its broad sense, it refers both to the various meanings assigned to the formal constants and also to the kinds of things which are allowed to replace the formal variables. For example, when one says that the constant '$\supset$' means "implies", one is giving an interpretation of <u>Principia</u> in the narrow sense of the expression "interpretation". But when one says that, for a <u>given</u> <u>meaning</u> of '$\supset$', the kinds of things which can replace the variables, say 'p' and 'q', are propositions like "Organisms demand states-of-affairs" and "There is a white card", one is proposing an interpretation of <u>Principia</u> in the broad sense of the expression "interpretation". In this essay, the word "application" will be used for the broad sense of "interpretation" and the word "interpretation" is confined to its narrow sense. So defined, it follows that if a logical system has been applied, then it has been interpreted. In this essay, the interpretation of <u>Principia</u> is what Carnap calls the "normal interpretation of a logical calculus". That is, the constants '.', '~', 'V', '$\supset$', '$\equiv$', '=', '$\exists$', mean, respectively, "and", "not", "or", "implies", "if and only if", "identical" and "There is at least

one".

Definition of Formal Organization

7. The first step in logically organizing K in terms of *Principia* consists in translating the propositions of K into symbolic statements which include formal constants in *Principia* "normally" interpreted. For example, consider (1) in K. Let 'xDp' mean 'x demands a state-of-affairs p' and 'pMx' mean 'p is a means-object of x'. Then (1) may be written as

$$(1') \qquad (\exists p)(xDp \,.\, pMx).$$

Again, if 'xTp' means 'x is teachable (docile) with respect to p', then (2) may be written

$$(2') \qquad (\exists p)(xTp) \supset (\exists p)(pMx).$$

Finally, (3) may be written as

$$(3') \qquad (\exists p)(pMx \vee \sim xDp).$$

If (1') is taken to be a postulate, then (2') and (3') will follow as consequences of (1') where the laws which permit these deductions are those found in *Principia*. That is, when 'xDp', 'pMx', and 'xTp' are allowed to replace the variables 'p', 'q', and 'r' in the laws of *Principia*,--when an application of *Principia* is made to K--, (2') and (3') are seen to be theorems following from (1') in accordance with the laws of *Principia*. For example, from (1'), by distribution of the existential quantifier over p, that is, '$(\exists p)$', there follows

$$(\exists p)(xDp) \,.\, (\exists p)(pMx);$$

then, by the *Principia* law "$(p.q) \supset p$",

$(\exists p)(pMx)$

from which, by the Principia law "$p \supset (q \supset p)$",

$(\exists p)(xTp) \supset (\exists p)(pMx)$

follows. This is (2').

Again, from (1'), with the aid of the Principia law "$(p.q) \supset (p \supset q)$", there follows

$(\exists p)(xDp \supset pMx)$;

then, with the aid of the Principia law "$(p \supset q) \equiv (\sim q \supset \sim p)$",

$(\exists p)(\sim(pMx) \supset \sim xDp)$

from which, in accordance with the Principia "$(\sim p \vee q) = _{Df} (p \supset q)$",

$(\exists p)(pMx \vee \sim xDp)$

follows. This is (3').

8. The account in Paragraph 7. thus provides an illustration of the logical organization of K by means of an application of the symbolic language of Principia. K will thus be said to be formally organized. It should be noted that in this illustration, taking (1') as a postulate results in a certain economical reduction of the system K. For it can easily be shown that neither (1') nor (2') follow from (3') and that neither (1') nor (3') follow from (2'). The "core" of the system K is the proposition expressed in (1').

9. Two benefits to be derived from the formal organization of a scientific theory or system have already been hinted at in the preceding paragraphs. It has been suggested that there results a certain facilitation in deduction and also a

certain economy in fundamental principles. The next section of the chapter will be devoted to a discussion of the advantages which are to be derived from formal organzation of a scientific theory. The ensuing section is thus a defense of the endeavor undertaken in this essay. The present author feels compelled to discuss this matter for two reasons. First, in contemporary psychology, there are some misunderstandings and some misgivings about the advantages to be obtained from such an endeavor _for the experimental investigator_. Secondly, there are certain claims which have been made by logicians about formal organization which the present author feels have perhaps led to the above mentioned misunderstandings and misgivings on the part of the experimental investigator. The next section will attempt to resolve these problems.

II

The Rejection of the Argument from Precision

10. In this section some arguments concerning the utility of formal organization will be examined. One of these suggestions is unwarranted; the others are not. Let us first consider the unwarranted argument.

11. Some investigators, notably Carnap[6] and Woodger[7], believe that formal organization is a useful tool to the experimental investigator because it is requisite to the construction of rigorous, comprehensible theories. They argue as follows: Natural languages like English and German are replete with obstructions to precise scientific expression; for example, excessive ambiguity of expression, vagueness of expression, tendency to contradiction, and so on. These obstructions are the result of the"unsystematic and logically imperfect structure(s)" of the natural languages.[8] In other words "the richness of their vocabularies and the arbitrariness of their syntactical rules militate against their suitability for scientific purposes by rendering them difficult of control".[9] Again, they possess moral and religious overtones which make them scientifically suspect.[10] To borrow a Woodgerism, they are "not precise enough" for scientific use.[11] As a result both the lack of rigor and the incomprehensibility that we find in so many scientific theories is not so much due to the theorizer himself as it is to the imprecise language with which he expresses his theory. For, we are assured, "[The scientists] know what

---

6. Carnap, R., Logical Syntax of Language, Harcourt, Brace and Co., 1937.
7. Woodger, J. H., The Technique of Theory Construction, International Encyclopedia of Unified Science, Vol. II, No. 5, 1947. See also Woodger, J.H., Biology and Language, Cambridge, 1954.
8. Op. cit., Logical Syntax of Language, p. 2.
9. Op. cit., The Technique of Theory Construction, p. 2.
10. Ibid., p. 2.
11. Op. cit., Biology and Language, p. 9.

they mean, but the current linguistic apparatus [natural language] makes it very difficult for them to _say_ what they mean".[12] In short, the incomprehensibility of many scientific theories is due to the contradictions, ambiguities and deceptions, that is, to the "dangers", in the natural languages used to express them.[13] Therefore, these investigators recommend, --indeed, they insist, -- that formal systems like _Principia_, which though perhaps not devoid of like dangers at least reduce these dangers to a minimum, must be used in order to organize scientific theories into rigorous, comprehensible scientific instruments.

12. The preceding argument is unwarranted. It fails for this reason: Woodger, Carnap, _et al_, have taken the wrong kind of entities as arguments to the predicates "precise", "ambiguous", and "deceptive". In other words, it is _not_ a language which is precise or imprecise, ambiguous or unambiguous, deceptive or not deceptive. Rather it is _the use which a theorizer makes of a language_ which has these properties. Therefore, the rigor and comprehensibility of a scientific theory _is_ relative to the theorizer's _use_ of a language and not merely to the language. Hence, we conclude that formal organization is not requisite to the construction of rigorous, comprehensible scientific theories.

---

12. _Ibid._, p. 95.
13. _Op. cit._, _Logical Syntax of Language_, pp. 311-312.

13. To illustrate this point one need only consider the fact that certain terms in the so-called "precise" languages are often used imprecisely.[14] Thus, Carnap, when concerned with Russell's ambiguous use of "implication" writes:[15]

> Russell's choice of the designation 'implication' for the sentential junction with the characteristic TFTT has turned out to be a very unfortunate one. The words 'to imply' in the English language mean the same as 'to contain' or 'to involve'. Whether the choice of the name was due to a confusion of implication with the consequence-relation, I do not know; but, in any case, this nomenclature has been the cause of much confusion in the minds of many, and it is even possible that it is to blame for the fact that a number of people, though aware of the difference between implication and the consequence-relation, still think that the symbol of implication ought really to express the consequence-relation, and count it as a failure on the part of this symbol that it does not do so. If we have retained the term'implication' in our system, it is, of course, in a sense entirely divorced from its original meaning; it serves in the syntax merely as the designation of sentential junctions of a particular kind.

In summation, the view that a scientific theory needs to be formally organized in order to be "rigorous and comprehensible" must be rejected. Again, the view that loose and incomprehensible theories are not due so much to the theorizer as to the imprecise instrument he is using is also at fault. For if the belief that it is the theorizer's <u>use</u> of language which is precise or imprecise, and so on, is justified, then the rigor and comprehensibility of his theory depend not so much upon the language he uses to express his theory as it does upon <u>him</u>, that is, upon the use which he makes of that

---

14. See Bennett, A. A., and Baylis, C.A., <u>Formal Logic</u>, pp. 269-271.
15. <u>Op. cit.</u>, <u>Logical Syntax of Language</u>, p. 255.

language.

<u>The Utility of Formal Organization</u>

14. Let us turn now to the arguments in favor of formal organization which the writer considers to be warranted. There are five such suggestions.

15. <u>Variables and symbolization</u>. Here the concern is with the notion of a variable insofar as that notion contributes (1) to the rigor of a given analysis -that is, to the intended logical representation of the structure of a proposition or set of propositions and (2) to the avoidance of difficult circumlocutions often required in order to express a given proposition or set of propositions in an unstandardized natural language. The point to be made here is that a certain rigor and simplicity results from the employment of a symbolic language in order to express the logical relations holding between propositions in a given empirical system. Hence, in this sense, it is <u>expedient</u> to formally organize an empirical theory or system. To illustrate, let us examine the following passage from J. H. Woodger's, <u>Technique of Theory Construction</u>.[16]

> The use of variables is a special instance of symbolization which deserves a little further consideration. It will be noticed that in the English translations of some of the statements of Part II of T we were able to avoid the use of variables without loss of precision.

---

16. Cf. pp. 67-68.

In other cases this was not done, because without taking advantage of this convenient device an accurate English translation would have been possible only at the cost of intolerable circumlocution and prolixity. In further illustration of these facts we may consider the following examples.

Suppose we wish to state that the relation of being earlier than or before in time (in the sense we have denoted by 'T') is transitive, but without using the technical logical predicate 'transitive'. In a word language this can be done as follows:

(a) If any thing is earlier than another thing, and the latter is earlier than a third thing, then the first is also earlier than the third thing.

By the use of individual variables this becomes:

(b) For every $\underline{x}$, $\underline{y}$, and $\underline{z}$, if x is earlier than $\underline{y}$, and $\underline{y}$ is earlier than $\underline{z}$, then $\underline{x}$ is earlier than $\underline{z}$.

Thus the use of variables enables us to eliminate such words as 'another', 'the latter', 'the first', 'a third', etc., without ambiguity, and reduces the length of the statement by about one-half. The use of a single relation-sign 'T' in the place of 'is earlier than' prepares the way for the use of the calculus of relations and reduces the statement to a single line:

(c) For every $\underline{x}$, $\underline{y}$, and $\underline{z}$, if $\underline{x}T\underline{y}$ and $\underline{y}T\underline{z}$, then $\underline{x}T\underline{z}$.

In the notation of the theory T this was expressed by

(d) (All $\underline{x}$) ((All $\underline{y}$) (All $\underline{z}$) (($\underline{x}T\underline{y}$ and $\underline{y}T\underline{z}$) implies $\underline{x}T\underline{z}$))).

In the notation of <u>Principia Mathematica</u> 'All' is omitted from the quantifiers, and the logical constants 'and' and 'implies' are symbolized by '.' and '$\supset$', respectively, so that we reach a complete symbolization:

(e) $(\underline{x}, \underline{y}, \underline{z}) : \underline{x}T\underline{y} \,.\, \underline{y}T\underline{z} \,.\, \supset \,.\, \underline{x}T\underline{z}.$

(Here also some of the parentheses are replaced by dots.) Finally, we reach the highest degree of brevity by formulating the statement by means of 'T' and signs belonging to the calculus of relations:

(f) $T \mid T \subset T.$

> This last formulation owes its brevity to the fact that, by the use of constants belonging to the calculus of relations, we are able to eliminate individual variables.

16. <u>The argument from intuition and calculation</u>. It is a common observation that the results of insights are often wrong. It is precisely because we do not always trust the results of our intuitions that we require certain checks on this sometimes wayward process. Calculation (or deduction) may be regarded as one such check. On the interdependence of intuition and calculation, Woodger writes:[17]

> Although it is by intuition or common sense that we ordinarily think and discover new hypotheses, the use of a logical technique can help intuition in two ways: (i) by providing a check which enables us to determine whether a theorem arrived at by intuition is in fact a consequence of our assumptions or not and (ii) by providing intuition with a guide rope in complicated and unfamiliar regions where, without some such aid, it could not penetrate or would easily go astray. Intuition and calculation are both necessary for science. Neither is infallible. Together they compensate for each other's defects.

Formal languages like <u>Principia</u> are in part created and conventionalized for the purpose of determining when it is possible to calculate that certain propositions follow from certain other propositions. Certain investigators who have been hypnotized by the relatively concise and well organized set of rules for calculation found in formal languages, have concluded that "calculation in a natural language is not possible".[18] This attitude, I believe, stems from the

---

17. <u>Ibid.</u>, pp. 71-72.
18. <u>Ibid.</u>, p. 2.

artificial distinction between "precise" formal languages and "imprecise" natural languages. Rather it is the case that calculation is facilitated in a symbolic language like Principia. Symbolic languages like Principia arise from the desire to abstract away from the plethora of diversified rules which guide the use of signs in natural languages just those rules which apply to calculation. A symbolic language like Principia then is a very simplified segment of natural language. In short, the reason why we do not normally calculate in natural languages is because we find it unnecessary to do so. For if we were to calculate in natural languages we would first have to separate out, that is, abstract away, at least implicitly, just those rules which apply to calculation. But this has already been done, to some extent, by people who have "constructed" symbolic languages like Principia. Here again, we are guided by the principle of expediency.

17. The argument from prediction. Among the problems well calculated to start a violent debate in psychological circles is the problem of prediction. The issue does not turn about the nature of prediction, but rather is concerned with whether a given statement is or is not a prediction in accordance with some theoretical point of view. The problem of deciding whether or not a given proposition is a prediction from a given theory is quite important. For the adequacy of a theory is tested by what it predicts. Hence if there is no means by which one is able to determine with

reasonable certainty that a given theory allows a certain prediction, then experimental test of that theory is rather difficult.

18. As above, the present author does not maintain that decisions as to whether or not a certain state-of-affairs is predictable from certain theoretical grounds are impossible in natural languages. But, again, decisions of this sort are facilitated by formal organization. Again, as in the preceding section, the claim here is that such decisions when made in a natural language are usually circuitous and prolix where the decision is at all complex. In short, decisions of this sort in a natural language are often inexpedient.

19. To illustrate the relationship between calculation (deduction) and prediction (and also the relationship between calculation and the converse of prediction, that is, explanation), consider the following statement by Carnap.[19]

> Of greater practical importance is the deduction of a singular sentence from premisses which include both singular and universal sentences. We are involved in this kind of a deduction if we explain a known fact or if we predict an unknown fact. The form of the deduction is the same for these two cases. ...we find it again in the following example, which contains, besides signs of the logical calculus, some descriptive signs. In an application of the logical calculus, some descriptive signs have to be introduced as primitive; others may then be defined

---

19. Carnap, R., <u>Foundations of Logic and Mathematics</u>, International Encyclopedia of Unified Science, Vol. II, No. 7, 1952, p. 36.

on their basis. (Designata-for-description-signs)-rules must then be laid down in order to establish the interpretation intended by the scientist. Premiss (3) is the law of thermic expansion in qualitative formulation. In later examples we shall apply the same law in quantitative formulation.

	1. <u>c</u> is an iron rod.
Premisses:	2. <u>c</u> is now heated.
	3. For every <u>x</u>, if <u>x</u> is an iron rod and <u>x</u> is heated, <u>x</u> expands.
Conclusion:	4. <u>c</u> now expands.

A deduction of this form can occur in two practically quite different kinds of situations. In the first case we may have found (4) by observation and ask the physicist to explain the fact observed. He gives the <u>explanation</u> by referring to other facts (1) and (2) and a law (3). In the second case we may have found by observation the facts (1) and (2) but not (4). Here the deduction with the help of the law (3) supplies the prediction (4), which may then be tested by further observations.

The example given shows only a very short deduction, still more abbreviated by the omission of the intermediate steps between premisses and conclusion. But a less trivial deduction consisting of many steps of inference has fundamentally the same nature.

20. <u>The argument from "mathematization"</u>. Generally when one speaks of a mathematicized part of a theory, he is thought to be referring to that part of a theory which has been expressed in arithmetical or algebraic formulas, that is, the quantifiable aspects of the theory. In fact, however, the theory as a whole may be mathematicized if the deductive procedure is used. That is, not only those parts of a theory which are expressed in "quantitative" terms may be mathematicized by the technique of formal organization, but also <u>both</u> the quantitative and non-quantitative aspects of a theory can be "mathematically" arranged by the technique of formal organization. What is intended by mathematization of a theory in

this more general sense is aptly expressed by Woodger.[20]

> Modern logic has now been so far developed that it enables us to mathematicize the whole of a scientific theory and not only the numerical or quantitative part. It furnishes us with calculuses which enable us to perform complicated transformations with precision upon statements which contain no signs belonging to traditional mathematics.

21. This consequence of formal organization is quite valuable. Thus, for example, in the case of Tolman's theory, it permits one to mathematicize that theory in which quantitative techniques are either not worked out or are completely lacking. This is important because among many scientists the belief persists that rigorous development of a theory depends solely upon the development of quantificational techniques. On these matters, J. H. Brown writes:[21]

> A systematized science like modern physics uses mathematics in making measurements as psychology has attempted to do, but an equally important application of mathematics in physics is to the construction of theories. In any advanced science most of the measurements performed depend on a close integration of theory, law, and experiment. The older views of the scientific method which supposed that measurements lead to laws through the discovery of correlations between sets of measurements on different entities have been shown to be unsound. In actual scientific practice the theory leads to the law and the law to the possibility of measurement more often than measurement leads to laws and hence to theories (1). The psychologist in his attempt at an empiricism, based on what he supposes to be a sound mechanistic methodology, has neglected

---

20. Op. cit., Technique of Theory Construction, p. 7. Also see Quine, W. V., Mathematical Logic, Norton, 1940, pp. 7-8.
21. Op. cit., Psych. Theories, pp. 234-235.

> the possibilities of applying mathematical procedures to the construction of psychological theory. <u>Psychologists have made wide use of mathematics in measurement, but have scarcely ever used mathematical concepts in theory-building.</u> The purpose of this paper is to call to the attention of the mathematical psychologist certain mathematical procedures which may be used in the construction of psychological theories. Lack of space prevents the mathematical development of these concepts. The various references, however, should enable the reader to pursue the mathematics of this mode of attack further should he so desire.

Then he proceeds to describe the technique most appropriate to theory construction.[22]

> Recent methodological research has also shown us the most fruitful type of theory. The theory should be based on what may be called the hypothetico-deductive method, or the method of constructs. In this method, hypotheses are devised to account for the descriptive data and from these hypotheses, predictions are made which may be tested in experiment. The constructs used in the hypotheses must be capable of operational definition. They must further lead to theoretical postulates which may be tested in critical experiments. There is so much agreement now amongst methodologists on this point that to argue it further would require space which may better be spent on the development of the constructs themselves.

22. In conclusion, it is possible to give a theory some form of rigorous development without first having established any quantificational techniques. In fact, formal organization may suggest to the investigator certain attacks on the problem of quantification itself.

23. <u>New information.</u> Formal organization of a scientific theory may yield new information about the theory, and thus

---

22. <u>Ibid.</u>, pp. 235-236.

may suggest certain experimental situations by which to test the theory. This new information may be of various sorts. For example, a scientist may suspect that a given system is inconsistent. Formal organization of the system allows the scientist to test for consistency; if contradictory propositions are provable in the system, then the system is inconsistent. Again, it is valuable to know just what elements are basic in the system. In the present system there are eight primitive ideas. Notions like appetite or cathexis, aversion, sign-object, signified-object, etc., are reducible by definitions to these more basic ideas. But most important is the suggestion of new experimental "issues" which may result from the formal organization of the system. In the present system certain theorems involving the law of least effort, the concept of disruption, the notion of avoidance, and the like seem quite suggestive of new experimental tacks. Nevertheless, it must be remembered that in this initial task the purpose of the present author is to organize Tolman's theory rather than to find new experimental tacks. In the last chapter certain new experimental tacks are suggested.

24. One point needs to be pointed out in this discussion. It is not being claimed that formal organization is the cure for all theoretical ailments. Indeed, in the present work we shall see that the apparatus itself is, to a certain extent, defective. The above points were rather directed to the defense of formal organization as a legitimate program

in general. Those places in which the particular example of formal organization exhibited in this system is defective will be dealt with in the final chapter.

III

The Function of Intervening Variables

25. This section deals with a general description of Tolman's notion of the intervening variable. Tolman characterizes his "system" in the following way. Purposive behaviorism consists of an "asserted list of intervening variables", that is, of an asserted list of terms denoting inferred processes in the organism which are said to operate between the presentation of a given stimulus situation and the occurrence of a given response situation, and of a certain "asserted set of laws" concerning the nature and operation of these intervening variables.[23] The concept of an intervening variable is, thus, the key to Tolman's conception of a psychological system. It demands greater consideration. More precisely, what is required is a clearer account of the function of an intervening variable as a systematic instrument.

26. What is the systematic function of an intervening variable? Let us begin with this homely observation: there is some reason why Tolman picked "demand" rather than "oink"

---

23. Marx, M., Psychological Theories, MacMillan, 1951, p. 90. See also Tolman, E. C., "The Determiners of Behavior at a Choice Point", <u>Psych</u>. <u>Rev</u>., 1938, pp. 1-41.

to be an intervening variable in his system. That is, the selection of an intervening variable is not an arbitrary matter. Somehow, as Tolman uses the term "demand" in his system, it is seen to be more appropriate as an explanatory device of certain puzzling phenomena than is "oink" or "yak". The following explanation is not intended to be conclusive, but rather only suggestive of the reasons why Tolman picked "demand" rather than "oink" as an intervening variable and, hence, directs attention to a function of intervening variables which constitutes their "appropriateness" as intervening variables.

27. The preceding remarks suggest that an answer to the proposed question might be found by examining the way in which a typical theorist might construct a theory. Necessarily this discussion will be over-simplified. In general, the theorist is confronted with a set of facts which are stated in the form of singular propositions. In erecting a theory the theorist attempts to account for the various discrepancies which are recorded by the various singular propositions. These singular propositions may be observation statements or non-observation statements. For example, observation statements which Tolman might be concerned with are "Rat a eats food", "Rat a runs to the left", "Rat a bites experimenter", and so on. Examples of non-observation statements are "Rat a demands eating food", "Rat a expects to run to the left", "Rat a expects that the experimenter will jump", and so on. The difference between these two

kinds of statements is this: observation statements are directly verifiable by sense observation, while non-observation statements are indirectly verifiable, that is, they depend for their verification on the observation statements.

28. Let us suppose that the theorist is confronted with a class K where K contains the following "facts", that is, observation statements.

> Rat a is put in front of food.
> Rat a eats food.
> Rat b is put in front of food.
> Rat b eats food.
> .
> .
> .
> .
> Rat n is put in front of food.
> Rat n does not eat food.

[What will be said here applies equally well to non-observation statements.] Suppose further that the observations concerning rat n's behavior occur two weeks later than those concerning the behavior of the rats preceding rat n. On the basis of the behavior of rats a, b ..., the theorist generalizes that whenever a rat is put in front of food, he eats it. There are three important things to notice in the theorist's generalizing activity: (1) he has attempted to connect pairs of observation statements, for example "rat a is put in front of food" and "rat a eats food", by making a conditional statement out of them. This is evidenced by the words "... a rat is put in front of food, (then) he eats it"; furthermore, (2) he has affirmed that this conditional holds not merely in the case of the behavior of rat a, but also in the case of the behavior of rat b, indeed, for all rats.

This is evidenced by the word "whenever". In short he has speculated that the following general statement is true:

> (a) For every x, if x is put in front of food, then x eats food.

Finally, (3), it must be noted that (a) is essentially a prediction. That is to say (a), which has the general form '$(x)(\varphi x \supset \psi x)$', allows that if a state-of-affairs "φz" occurs, then one may conclude that "ψz" will also occur. But this generalization is falsified by the behavior of rat n. Thus the problem; how to account for the behavioral disparity of rat n, that is, how to explain the case in which "φn" occurs but "ψn" does not occur? The theorist considers the conditions under which rats a, b, ... <u>and</u> n were operating. He may notice the following features: (1) all of the rats but rat n were very hungry; (2) rat n's metabolism rate was the only one which was normal; and so on. It is thus conjectured that there is some complex process in the rat which will cause his behavior to deviate under these conditions, that is, the theorist hypothesizes a certain intervening element called "demand". The problem now is to construct a new general proposition such that the behaviors of rats a, b, ..., and rat n can be predicted. The new general statement might read:

> (b) For every x, if x is put in front of food, then x eats food <u>if</u> <u>and</u> <u>only</u> <u>if</u> x demands eating food.

(b) requires a few remarks. The expression "if and only if" is italicized because a relation at least as strong as "if and only if" is required if the case

Rat n is put in front of food and Rat n does not eat food

is to be accounted for by the generalization (b). Thus from (b) we can easily deduce

(c) If Rat a is put in front of food and Rat a demands eating food then Rat a eats food

and

(d) If Rat n is put in front of food and Rat n does not demand eating food then Rat n does not eat food.

Hence, (c) explains why Rat a eats food and (d) why Rat n does not eat food -namely, because in (c) he demanded eating food whereas in (d) he did not demand eating food.

29. The preceding account gives a brief description of the function of an intervening variable. It shows that the intervening variable "demand" enabled the theorist to formulate a reliable generalization. An expression like "oink" would hardly have been helpful. It is important to notice that this characteristic of an intervening variable enables the theorist to order the class K in the sense that each fact becomes a constituent in a generalization from those facts. One might say that a set of propositions K allows a generalization from it <u>in accordance with</u> the conditions expressed by the intervening variable "demand"; it is the particular intervening variable which <u>permits</u> the particular formulation of the generalization. Thus, we may say that part of the meaning of any intervening variable is its systematizing or ordering function-which cannot be reduced in any way to observation terms. This property of an

intervening variable is analogous to the '$\supset$' in <u>Principia</u>. The "horseshoe" allows us to arrange propositions only in certain ways; for example, contrast its use with that of "greater than", that is, '$>$', in chapter IV of this essay. Likewise different intervening variables allow us to arrange sets of singular propositions in different ways by permitting different generalizations from these sets of singular propositions. Thus we see that the intervening variables have an ordering or systematic function in addition to factual conditions which are also expressed by them.

30. This concludes the introductory remarks. In the next three chapters the formal organization of Tolman's theory is undertaken. The operations and interrelations between the various intervening variables in Tolman's system will be displayed. Greatest consideration will be given to the two fundamental notions of Tolman's system, namely, "demand" and "expectation".

CHAPTER II

PRIMITIVE IDEAS AND DEFINITIONS OF SYSTEM TI

1. The system to be developed in this essay is divided into two parts. The first part will be called TI; the second TII. TI is concerned with the relationship between the intervening variables independent of their strengths or frequencies. TI tends to stress the 'intervening' characteristic of the intervening variables rather than their 'variability' characteristic. The present chapter and, also chapter III, present TI and are thus more concerned with the methodology of Tolman's theory than with the working consequences of Tolman's theory; the latter objective is met in TII. TI is more general, more abstract than is TII. Its presentation is therefore somewhat simpler than is the presentation of TII.

I

General Apparatus

2. In TI, small English letters from the middle of the alphabet, 'p', 'q', 'r',..., occur as <u>propositional variables</u>. Small English letters from the end of the alphabet, 'x', 'y',..., are used as <u>organism variables</u>. Greek letters, 'φ', 'ψ',..., are <u>predicate variables</u>. Capital English letters are <u>predicate constants</u> (or <u>predicates</u>). The <u>logical constants</u>, both primitive and defined, are those found in <u>Principia</u>, for example, '.', '~', 'v', '$\supset$', '$\equiv$', '=',

'∃', etc. The arithmetical constants, for example, '>', '=', etc. are those found in Real Numbers.[1]

3. The last kind of symbols used in TI are the initial lower case letters of words or synonyms of those words which designate concepts in TI. For example, 'd' is the first letter of the word "demand". These expressions are functor-expressions; more precisely, they are construed as numerical functor-expressions. The notion of a functor is explained in the following passage from Hans Reichenbach's Elements of Symbolic Logic.[2]

> ....we turn now to a second kind of function, which has developed out of descriptions, and which we call descriptional functions.
> Consider a statement in functional notation
> $$f(y_1, x_1) \qquad (1)$$
> for instance, the statement 'y_1 is the father of x_1'. ...we can write the statement (1), in descriptional notation,
> $$y_1 = (\iota y)f(y, x_1), \qquad (3)$$
> [that is, y_1 equals the y such that y stands in the relation f to x_1. '(ιy)' means 'the so and so'.] Introducing the abbreviation
> $$f'(x_1) =_{Df} (\iota y)f(y, x_1) \qquad (4)$$
> we can write (3) in the form
> $$y_1 = f'(x_1) \qquad (5)$$
> Putting variables in the places of the constants, we obtain
> $$y = f'(x). \qquad (6)$$
> We may say that we have solved the functional 'f(y, x)' for the argument 'y'.
> The function 'f'(x)' is called a descriptional function, since its special values, resulting from specialization of 'x', are descriptions. This kind of function, which we indicate by the prime mark, is to be distinguished from propositional functions, whose values are propositions. If we want to construct a proposition by means of a descriptional function, we must not only specialize the argument but also add a symbol like 'y_1 = '. The symbol 'f'(x)' expresses a descriptional function in the form of a contracted term, since the bound variable and the iota-operator are not indicated; the explicit form of a descriptional function is indic-

1. Op. cit., Introduction to Logic, pp. 217-218.
2. Macmillan, 1947, pp. 311-314.

ated by the right-hand side of the definition

$$f'(x) =_{Df} (\iota y) f(y, x) \tag{7}$$

There are also descriptional functions of more than one variable. For instance, the description 'the man who walks between Peter and Paul' can be used for the construction of the descriptional function 'the z walking between x and y'.

Descriptional functions can be extended. Thus the description 'the color which is spectrally between red and yellow' determines the color orange; and it may be written in the form

$$\varphi'(r, s) =_{Df} (\iota f) B(f, r, s) \cdot V(f) \tag{8}$$

where 'r' stands for 'red', 's' for 'yellow', 'B' for 'spectrally between', and 'V' for 'color'. φ'(r, s) represents a descriptional function when the arguments r and s are conceived as variables.

The distinction between proper and improper use of descriptions applies likewise to descriptional functions. In general a descriptional function will be properly used for some arguments, improperly for others. Thus the descriptional function 'the brother of x', applied to persons as arguments, will be properly used when x has one and only one brother; in all other cases it will be improperly used. When a descriptional function is properly used for all values of its argument within a certain range, it will be called a functor with respect to this range. Thus the descriptional function 'the father of x' is a functor with respect to the range given by all human beings.

Among the functors, the unique mathematical functions are of particular importance. They are descriptional functions ..., having numbers as arguments and as descripta. Usually they are functors with respect to real numbers as their range. Thus when we write a mathematical equation in the form

$$y = f'(x) \tag{9}$$

the symbol 'f'' represents a functor with respect to real numbers as the range of its arguments. The range of the descriptum y may be more comprehensive; it may include complex numbers, for instance when the functor 'f'' is given by the square root. Functors whose descripta are numbers may be called numerical functors.]

The use of numbers as descripta of descriptional functions is not restricted to mathematics. There are properties of physical things which are expressed by means of numbers. Thus an individual motion-property f can be characterized by a number indicating the speed. This method has great advantages over the use of names for predicates; it presents the different properties in numerical functor with respect to this range. The range itself then consists in nonnumerical objects,...

As an example of a numerical predicate, let us consider the sentence 'x_1 moves at 50 miles an hour'. It can be written,

$$m_{50}(x_1) =_{Df} (\exists f)f(x_1) \,.\, U(f) \,.\, (f = 50) \qquad (15)$$

[that is, 'x_1 moves at 50 miles an hour' means 'There is a motion property f which belongs to x and f equals 50'] Solving the statement (15) for 'f', i.e., introducing descriptional notation, we obtain

$$(\iota f)f(x_1) \,.\, U(f) = 50 \qquad (16)$$

Introducing the abbreviation (14) we obtain

$$f''(x_1) = 50 \qquad (17)$$

Here the numerical functor '$f''(x)$' signifies 'the speed of x'.

4. Consider the intervening variable "demand". It is a relation between organisms and states-of-affairs. In order to capture the variability characteristic we construe "demand" as a numerical functor. After formula (17) in the above quotation, we write:

(1) $d(x, p) = N$

The functor is the definite description '$d(x, p)$' which may be read as 'the strength of x's demand for p'. The expression '$= N$' written after the functor stipulates that the functor is a <u>numerical functor</u>, that is, it ranges over numbers. In the present system 'N' can only be replaced by positive numbers. Hence, the entire expression in (1) may be read as 'The strength of x's demand for p equals a given positive number, N'. (1) is a proposition and, hence, may be true or may be false. To say that 'x demands p' or that 'x's demand for p exists' or that 'it is true that x demands p' is to say that

(2) $d(x, p) \neq 0$,

that is, 'the strength of x's demand for p is not equal to 0'. In TI and TII, we lay down the convention that 0 is not greater than any number, that is, 0 is the lowest number. Thus where '$>$' means 'greater than', (2) is equivalent to

(3) $d(x, p) > 0$,

that is, 'the strength of x's demand for p is greater than 0'. Again

(4) $\sim d(x, p) = N$,

that is, 'it is false that the strength of x's demand for p equals N' is equivalent to

(5) $d(x, p) \neq N$,

that is, 'the strength of x's demand for p is not equal to N'. Finally, every intervening variable dealt with in TI and TII will be treated in the same fashion as "demand"; that is, they will all be treated as numerical functors.

5. Dealing with propositions as arguments to predicates like 'demand', 'expectation', etc. creates a special problem. For we shall want, in some cases, to write:

(6) $p = q$

that is, "p is identical with q". In <u>Principia</u> '=', is a relation between values of the variables 'x', 'y', 'z' ...,. In order to be able to write a statement like (6) we lay down a convention for interpreting '=' when it holds between propositions, namely; <u>given any pair of propositions p and q, if p is deducible from q and q is deducible from p, that is, if p and q are mutually deducible, then p is identical with q</u>.

To illustrate the use of this convention consider the two propositions

(7) Rat a is satiated

and

(8) Rat a is not hungry.

Now if (7) is true, then (8) follows. Again if (8) is true, then (7) follows. In short (7) and (8) are mutually deducible propositions. In such cases we shall write, in accordance with the convention,

(9) p is identical with q

or, in symbolism, the statement in (6) above. Under this interpretation of identity -- with respect to propositions -- the apparatus of *Principia* is readily applicable.

6. The laws for the manipulation of the logical entities discussed in paragraph 2. are those found in *Principia*. The laws concerning the arithmetical operations are those found in, or which are deducible from, *Real Numbers*. According to the convention mentioned in paragraph 4. that 0 is not greater than any number, supplementing the postulate set in *Real Numbers* in such a way guarantees that the only numbers which may be substituted for N in

$$d(x, p) = N,$$

or in any proposition containing a numerical functor, are positive numbers.

Goal-objects and States-of-affairs

7. It is commonplace that Tolman considers a goal-object to be the kind of thing which organisms demand, expect, and so on. Unfortunately it is not clear what the character of a goal-object is.[3] Tolman wavers between two interpretations of "goal-object". On the one hand, he regards a goal-object, that is, that which is demanded or expected, as _propositional_ in character, that is, as a _state-of-affairs_, or a _situation_. Thus, when he is discussing the properties of molar behavior, he writes:[4]

> The first item in answer to this question is to be found in the fact that behavior, which is behavior in our sense, always seems to have the character of getting-to or getting-from a specific goal-object, or _goal-situation_.[16]
>
> 16 For convenience we shall throughout use the terms goal and end to cover _situations being got away from_, as well as for _situations being arrived at_, i.e., for _termini a quo_ as well as for _termini ad quem_.

Two pages later he writes:[5]

> The animal when presented with alternatives always comes sooner or later to select those only which finally get him to, or from, the given demanded, or to-be-avoided, _goal-object or situation and which get him there by the shorter commerce-with routes_.

---

3. Here goal-object is being used in its broadest sense, for example, means-objects, and final goal-objects, are goal-objects.
4. _Op. cit._, _Purposive Behaviorism_, p. 10. (My italics)
5. _Ibid._, p. 12. (My italics)

And finally his most obvious employment of goal-objects as propositional occurs in the following passage having to do with expectations.[6]

> In these latter cases also, the perception or memory, *as an expectation, is a prior "setting" of the behavior for such subsequent encounter as: "here an opening"; "there a wall"; "here a smellable crevice"*; and the like. Again, the actual encounter which verifies or fails to verify that "this is an opening, a wall, or a crevice", is a temporally separate and later event."

It is also interesting to note that in the Glossary of *Purposive Behaviorism*, he stipulates that by definition, "goal-object", "goal-situation" and "goal" are synonymous. On the other hand, he also thinks of goal-objects as *things* or *qualities*. This is amply illustrated by the following statements. When discussing the phenomenon of disruption he writes:[7]

> We shall suppose, in general, that in behavior there is always immanent the expectation of some *more or less specific type of goal-object*."

And, again, in the same context:[8]

> The appearance of such disruption will define the fact of an *immanent expectation of the previous type of goal-object, i.e., the type of goal-object which, as long as it is present, does not cause disruption*.

And finally, in his most explicit statement of goal-objects as things or qualities, he writes:[9]

---

6. *Ibid.*, p. 84. (My italics)
7. *Ibid.*, p. 76. (My italics)
8. *Ibid.*, p. 74. (My italics)
9. *Ibid.*, p. 74. (My italics)

> It would, that is, undoubtedly be discovered further: (a) <u>that there was a definite range of goal-object qualities, i.e., those of bran mash, or of foods</u> so closely similar to such mash that the rat could not distinguish them from the latter, which would lead to no such disruption as long as they were found at the exit-box.

8. The problem then is this: Are goal-objects things like "food" or qualities like "leafy succulence of lettuce", or are they states-of-affairs like "x eats food" or "x chews leafy lettuce"? In the discussion of demand as a functor in paragraph 4. that which is demanded or expected, was represented by a propositional variable, that is, by 'p'. Goal-objects were there construed as propositional in character. In this essay, this convention is followed throughout. Three arguments are offered in defense of this convention. First, construing goal-objects as states-of-affairs permits one to describe different behavior patterns in which the same thing -- say, food -- is involved; construing goal-objects as things does not. Secondly, certain important behavioral situations, for example, conflict situations, are not accurately represented when goal-objects are construed as things or qualities. Thirdly, construing goal-objects as things or qualities leads Tolman, in certain instances, to an interpretation of the behavior situation which has unfortunate implications. In connection with this last point, in the discussion of the primitive idea of expectation an attempt will be made to show that Tolman himself implicitly construes goal-objects as propositional

in character.

9. Suppose we have a moderately hungry cat who has just caught a fat mouse. Suppose further that our cat spends some time in examining the mouse, for example, smelling it or turning it over, before he eats it. How do we distinguish between the cat eating the food (mouse) and the cat inspecting the food (mouse)? If goal-objects are construed as thing-like the distinction is difficult to make. For in either case what the cat demands -- is food (mouse). Thus either (or both) behavior patterns would have to be written (in the present symbolism) as :

d(x 'food') = M,

that is, 'the strength of x's demand for 'food' equals the number M'. If goal-objects are taken as situations the two behavior patterns noted in the illustration can be discriminated. One of them can be written as:

d(x 'x inspect the food') = M

that is, 'the strength of x's demand that 'x inspect the food' equals the number M'. The other may be written as:

d(x 'x eat the food') = N

that is, 'the strength of x's demand that 'x eat the food' equals the number N'.

10. Consider a case of conflicting demands.[10] Such a situation is not accurately described by saying that one

---

10. These remarks also apply to conflicting expectations.

of the alternatives <u>is not</u> demanded while the other alternative is demanded. Rather <u>both alternatives are demanded</u>. For example, construe goal-objects as things. Suppose an organism to have conflicting demands with respect to a thing -- say, a door. This may be written as:

$$d(x, \text{'door'}) = N. \quad (N > 0)$$

But then we should have to write:

$$\sim d(x\ , \text{'door'}) = N. \ (N > 0)$$

that is, 'x does not demand the 'door'', for the other case. Notice, however, that the conflict has to do with some relation between x and the 'door', for example, opening the door; -- not merely with the door alone. To express this kind of situation we may write:

$$d(x\ , \text{'x opens the door'}) = N \quad (N > 0)$$

that is, 'x demands that 'x open the door'', for the one case and

$$d(x\ , \sim (\text{'x opens the door'})) = N \quad (N > 0)$$

that is, 'x demands that 'x does not open the door'', for the conflicting case. In the last two cases x demands in <u>either case</u>. The demands are thus in conflict.

11. On pages 128-129 of <u>Purposive Behaviorism</u> Tolman gives his interpretation of the behavior of the animals in an experiment by Robinson and Wever on visual distance perception. I shall quote both the experimental results and Tolman's interpretation of them.

We turn now, finally, to an experiment which indicates the ability of the rat to respond to a hierarchy of subordinate and superordinate goals. This is an experiment by Robinson and Wever. It indicated that the rat can choose a given means-object by virtue of whether or not it leads to a given subordinate goal-object (in this case, to be sure, a negative, or avoidance, subordinate goal-object).

'Two paths, Right and Left, led from the entrance to the food, but doors of the vertical sliding type, and of the same material and color as the walls of the maze, were provided....to permit the closing of either path, as desired. Along the top of the paths a row of electric lights gave even illumination. At the choice point, the right and left alleys were obscured from view by two black flannel curtains, making it necessary for the rat to enter the blind in order to see whether the path was open.'

'For about every third trial both paths were left open and the rat made his way unimpeded to the food. But for the remaining trials both doors were closed until the rat had passed into one alley, had turned around and started back; then the door of the unentered alley was quietly opened by means of a cord in the hands of the experimenter and the animal thus permitted to pass along that way.'

The tasks for the animals which we are here interested in were (a) that of mnemonizing that the closed door on either side meant the non-availability of that side as a route to food, and (b) that of perceiving, as soon as possible, after passing under the curtain, the presence or absence of the sub-goal-object (in this case a negative or avoidance goal-object), the closed door. The results indicate that on the first few days the rats ran way up to the door before rejecting a given side and turning back. It appears, further, that they then learned to turn back sooner and sooner until, finally, each animal reached a relatively constant level of performance of turning back at some characteristic distance from the door.

12. In Tolman's writings the expression "negative or avoidance object" is ambiguous. It could mean -- and in fact it does in this context -- that it is false that x

has a demand for the door. But according to the convention established in chapter II of <u>Purposive Behaviorism</u> the expression "avoidance object" means a demand for the opposite goal situation. Tolman is led into this confusion because he is treating <u>things</u> as goal-objects, that is, where 'z' takes as values things like 'doors', 'alleys', etc., '$d(x, z) > 0$' is the appropriate way of symbolizing a positive goal-object. Let us write:

(10) $d(x, z) > 0$.

An avoidance goal-object is then written as

(11) $\sim d(x, z) > 0$.

By making goal-objects things Tolman is forced to write that "...the mnemonizing [remembering] [of] the closed door on either side <u>meant</u> the non-availability of that side as a route to food." The implications of this interpretation of "avoidance object" are unfortunate.

13. It is fair to translate the sentence in the quotation above which reads: "...the closed door on either side meant the non-availability of that side as a means to food." into the following statement in Tolmanian language:

(a) It is false that x expects alley L, that is, 'that side', leads to food,

where 'alley L' is a 'means-object' and 'food' is the 'final goal-object'.

Here an avoidance object has been clearly interpreted as a denial of the organism's demand for a given goal-object (in this case the closed door at the first corner of alley L).

14. Tolman's reasoning here may be reconstructed as follows. Let '$d(x, L) > 0$' mean 'x demands alley L', '$d(x, d_1) > 0$' for 'x demands the closed door', '$d(x, f) > 0$' for 'x demands food', '$e(x, L \supset d_1) > 0$' for 'x expects that alley L leads to food'. Given the avoidance law,

$$(12)\quad [e(x, L \supset d_1) > 0 \;.\; \sim d(x, d_1) > 0] \supset \sim d(x, L) > 0,$$

that is, 'if x expects that alley L leads to the closed door and x doesn't demand the closed door then x doesn't demand alley L', and the approach law

$$(13)\quad [e(x, L \supset f) > 0 \;.\; d(x, f) > 0] \supset d(x, L) > 0,$$

that is, 'if x expects that alley L leads to food and also demands food, then he demands alley L', and the following facts which held in the Robinson-Wever experiment,

$$(14)\quad \sim d(x, d_1) > 0,$$

that is, 'x does not demand the closed door',

$$(15)\quad e(x, L \supset d_1) > 0,$$

that is, 'x expects that alley L leads to the closed door', and

$$(16)\quad d(x, f) > 0$$

that is, 'x demands food', we may deduce

$$(17)\quad \sim d(x, L) > 0,$$

from (12), (14) and (15). From (17) we may deduce

$$(18)\quad \sim e(x, L \supset f) > 0,$$

with the help of (13) and (16). (18) is a symbolic translation of (a).

15. The reasoning here is perfectly proper; but, unfortunately, it does not say what Tolman wants to say. Consider that organisms which are <u>not even in</u> the experimental situation satisfy (18) -- and what is more, (14); an organism in a deep sleep would make both (18) and (14) true. These organisms neither expect that alley L leads to food nor demand the closed door. But surely it is not Tolman's intention to explain the "behavior" of organisms who do <u>not</u> have demands for a given goal-object and who <u>do not</u> expect that given goal-objects lead to other given goal-objects!

16. A further objection is that the above reasoning does not permit an adequate explanation of what actually happened in the above experiment. For there is no way of inferring on the basis of (18) and (14) the behavior of the animals who clearly turned around and got to the food by the alley on the right -- as the experimental results indicate. In other words, it is not true that the following laws hold:

$$(19) \quad \sim d(x, L) > 0 \supset d(x, R) > 0$$

$$(20) \quad \sim e(x, L \supset f) > 0 \supset e(x, R \supset f) > 0$$

where 'R' means 'alley R (on the right)'. For the negations of (6) and (7) are both possible; that is,

$$(21) \quad \sim d(x, L) > 0 \;.\; \sim d(x, r) > 0,$$

and

$$(22) \quad \sim e(x, L \supset f) > 0 \;.\; \sim e(x, R \supset f) > 0$$

are both possible. Situations (21) and (22) would be satisfied by our sleeping organisms.

17. These difficulties are the product of thinking of goal-objects as 'things' rather than 'states-of-affairs'. For it makes no sense to speak of x having a demand for a not-closed door. And hence the only way of describing the avoidance object is

$$\sim d(x, d_1) > 0.$$

To put it in other words treating goal-objects like 'things' forces Tolman to treat avoidance and approach objects as analogous to logical contradictories when they are more accurately treated as analogous to logical contraries.

18. Finally, interpreting goal-objects as propositions would not have permitted these unfortunate consequences. In accordance with the interpretation of goal-objects as situations the laws (12) and (13) would have to be rewritten. Let 'L' mean 'x runs down alley L', 'd_1' mean 'x runs to door$_1$' and 'f' mean 'x eats food'. Then we rewrite (12) and (13) as follows:

$$(23) \quad [e(x, L \supset d_1) > 0 \,.\, d(x, \sim d_1) > 0] \supset d(x, \sim L) > 0$$

$$(24) \quad [e(x, L \supset f) > 0 \,.\, d(x, f) > 0] \supset d(x, L) > 0$$

(23) describes a typical avoidance situation, that is, "If x expects that if he runs down alley L then he can run to the door, and x <u>demands</u> that he not run to door, then x <u>demands</u> that he not run down alley L". (24) may be interpreted analogously; (24) describes a typical approach situation. Again the "facts" (14), (15) and (16) will have to be rewritten under the new interpretation of goal-objects as:

(25) $d(x, \sim d_1) > 0$

(26) $e(x, L \supset d_1) > 0$

(27) $d(x, \sim f) > 0$

Under this new interpretation of goal-objects we cannot arrive at the conclusion in (18). A more appropriate interpretation of the behavior of the Robinson-Wever animals would read:

(28) $e(x, \sim L \supset f) > 0$

or "x expects that if he does not run down alley L then he gets to eat food." This would mean not that the closed door on either side meant the non-availability of that side as a means to food but rather that the closed door on either side meant the availability of the other side, that is, alley R (the Right alley) as a means to eating food. To get (28), laws other than (23) and (24) would be required. What they may be need not deter us here. The point to be made is satisfied by the above discussion; namely, that interpreting goal-objects as propositional in character does not lead to construing an avoidance object as a denial that the organism demands a given goal-object. Further evidence for this contention is shown in the fact that sleeping animals or, more broadly, animals not in the above problem situation would not satisfy the conditions expressed in (25) or in (28). For the conditions in (25) and (28) require that the animals demand and expect given goal-objects. Again, 'x runs down alley R' would satisfy the conditions expressed in (28).

The Primitive Idea of Demand

19. On page 441 of Purposive Behaviorism Tolman describes a demand as

> An innate or acquired urge to get to or from some given instance or type of environmental presence or of physiological quiescence...or disturbance...

This statement suggests that demands are directed toward or are about things; it is a thing which the organism gets to or from. This view of demand is subject to the difficulties pointed out in the immediately preceding section of this chapter; especially, is it subject to the difficulty of not being able adquately to characterize the approach-avoidance object distinction. On page 437 of Purposive Behaviorism, Tolman defines an avoidance object as one "which is to be got from". Hence, he suggests that the organism does demand something in the avoidance situation. But, as has been pointed out in paragraph 12, where the approach situation is taken as an urge to get to a given thing, the avoidance situation must be characterised as the denial of the urge to get to that thing, that is, the denial of the demand for that thing. This, of course, conflicts with the above description of an avoidance object as a "demand-against" or as an urge to get from the object. In accordance with the remedy suggested in the immediately preceding section of this chapter, the above description of demand is reinterpreted by replacing the word 'presence' with 'situation' and introducing the expression 'situation of' between 'physiological' and 'quiescence'.

20. There is another characteristic of demand which is not made sufficiently clear in the above characterization of demand; it is the forward pointing character of demand. When one affirms that an organism demands a state-of-affairs be or come to be true. Our demands or urges are for something in the future; not for something in the present or past. For example, when I demand that my hunger be alleviated, I am now hungry but I am not now alleviated; the state of alleviation is something to be brought about and hence is something postdating my demand. The moment alleviation of hunger occurs, demand for it, in the particular case, ceases. The most natural phraseology here is: x demands that his hunger be alleviated; the demand is for the bringing about of some state-of-affairs. Demands are thus forward pointing. In order to capture the forward pointing character of demand in the above characterization of demand the following emendation is made: Substitute for the expression 'urge to get to or from' the phrase 'urge to bring about', and add the phrase 'or its opposite' immediately after 'environmental situation'.

21. Demands, depending upon the circumstances and the demanding organism, may vary in strength. On page 67 of Purposive Behaviorism, Tolman writes:

> With relatively strong nursing-need the litter was strongly demanded; with a relatively slight nursing-need it was but slightly demanded. And correspond-

> ing to these <u>differences in the demand</u> the maze performance improved and degenerated.

Consequently, we add to the above characterization of demand the phrase 'which varies in strength'.

22. In virtue of the proposed ramifications listed in the preceding three paragraphs, "demand" may be described as

> An innate or acquired urge to bring about some given instance or type of environmental situation (or its opposite) or of a physiological situation of quiescence or disturbance and which varies in strength.

Accordingly the primitive idea of demand may be appropriately expressed in functor notation as

$d(x, p)$,

which means, 'the strength of organism x's demand that p be true'. The conventions for expressing the case where 'x demands p (or not p)' are those discussed in paragraph 4. in this chapter.

23. Finally, demands sometimes may be drives, but the converse is not true. On page 27 of <u>Purposive Behaviorism</u> Tolman defines a first order drive as a

> ...demand for the presence of [a] specific physiological quiescence (appetite) or against the presence of [a] specific physiological disturbance (aversion) which results from initiating organic excitements.

In short, a first order drive is a "demand for (or against)" a final goal-object. But, on the other hand, organisms also demand means or subordinate goal-objects like 'running down the alley', 'turning to the right', etc. These latter demands are <u>not</u> drives.

The Primitive Idea of Expectation

24. The concept of expectation is perhaps the most important concept in Tolman's theory; certainly it is the most novel. It is in terms of this concept that Tolman's theory is usually contrasted with other theories of learning; for example, the s-r reinforcement theory of Hull, the theory of Guthrie, etc.

25. Tolman conceives of an expectation as a setting of the organism for the occurrence of a situation.[11] Less abstractly, to say that an organism expects a given state-of-affairs to be true is to say that he 'believes' that the state-of-affairs will occur. This statement is not to be taken as implying that expectations are always conscious. Tolman's account of expectation, as he himself notes, rests on no assumption of the organism being conscious.[12] An expectation is a belief or judgement, conscious or unconscious, that so and so will occur.[13] Indeed, it might be compared with Russell's notion of "animal inference".[14] The essential difference between demand and expectation is this: demands impel or initiate the organism into action; expectations direct the organism's action.

---

11. The remarks made in paragraphs 24 and 25 apply to all kinds of expectations. The expression of means-end expectation is described in paragraph 27.
12. _Purposive Behaviorism_, p. 204.
13. _Purposive Behaviorism_, p. 29.
14. Russell, B., _Human Knowledge; its Scope and Limits_, London, 1948.

26. An expectation, like a demand, is forward pointing.[15] It is always conceived to be forward-pointing with respect to some immediately distant or future occurrence. It is evoked by present stimuli. On this point there is a seemingly glaring inconsistency to be found between Tolman's account of expectation in chapter IV and his account of expectation in chapter V in Purposive Behaviorism. On page 84 of Purposive Behaviorism he writes:

> Expectations, those of means-objects as well as those of goal-objects, those to be called perceptions as well as those to be called memories, are always forward-pointing. This fact of their forward-pointingness is relatively obvious for expectations of goal-objects. For here the stimuli and the resultant expectations which they release quite evidently precede the later goal-object encounters, which verify or fail to verify such expectations. The expectations in such cases very evidently point forward to the later moments of goal-object encounter and verification. But a similar situation really likewise obtains for expectations of immediate means-objects. In these latter cases also, the perception or memory, as an expectation, is a prior 'setting' of the behavior for such subsequent encounter as 'here an opening'; 'there a wall'; 'here a smellable crevice'; and the like. Again, the actual encounter which verifies or fails to verify that 'this is an opening, a wall, or a crevice,' is a temporally separate and later event.

Here Tolman stresses the point that what is verified is always temporally consequent to the expectation of it. For example, suppose that we have a straight alley maze with food in the goal-box. Suppose further the rather trivial circumstance that if the animal gets to the food,

---

15. Purposive Behaviorism, p. 84.

he can eat the food. Let us assume that animal has been able to build up an appropriate means-end expectation. Now Tolman's description of the test run -- in accordance with the account in the above quotation -- might run as follows: "The animal expects that if he gets to the food, he can eat the food. The animal is allowed to run down alley A so that he gets to the food. 'Getting to the food' is a stimulus situation. As a result there is evoked the expectation that he can eat the food -- an activity which is verifiable after the expectation." Contrast this situation with the one cited in the following quotation on page 81 of <u>Purposive Behaviorism</u>.

> With such an arrangement, the stimuli coming from the goal-object would, during the given trial itself, evoke an expectation as to the character of this goal-object.

Here Tolman seems to be saying: "Under the same conditions in the artificial situation described above, getting to the food evokes the expectation that the animal can get to the food." But notice in this situation what is expected is not something verifiable <u>after</u> the expectation; for the activity, that is, running down the alley has <u>already occurred</u>. Briefly the two situations may be contrasted in this way: the first situation says that the animal expects to do it, <u>then</u> (i.e., later) does it. The latter situation says rather that the animal does it, <u>then</u> (i.e., later) expects to do it -- sort of a James-Lange theory of expectation!

27. The reason this inconsistency is only apparent is

because 'goal-object' is being used ambiguously. On page 84 of _Purposive Behaviorism_, it is a state-of-affairs; on page 81 of _Purposive Behaviorism_ it is a thing. Under the interpretation of goal-objects as things the common "object" in the above situations is 'food'. But, as we have seen earlier, in this chapter, it is extremely difficult, if not impossible, to discriminate between different activities involving the same "object" or thing, e.g., in this case 'food'. Tolman's real intention in the passage on page 81 of _Purposive Behaviorism_ is accurately exhibited by the description in the passage on page 84 of _Purposive Behaviorism_. In terms of that description the stimulating goal-object would be "getting to the food" while the expected goal-object would be "eating the food". And thus the "inconsistency" disappears.

28. Expectations, like demands, are to be construed as relations between organisms and states-of-affairs. Expectations, like demands, are also variables; that is, they vary in strength. One may have a strong expectation or a weak expectation depending upon the situation and the expecting organism. Evidence that Tolman conceives of expectation as a "variable" is presented in the following statement.[16]

> The second feature of our system which may, perhaps, be abhorrent to true Gestalt-ists is that we have included among these determining variables not only

---

16. _Purposive Behaviorism_, p. 420.

> the immanent sign-gestalts and the behavior-adjustments but also: (a) a variety of preceding determinants; viz., capacities and (b) <u>a series of analyzed variables within the sign-gestalts; viz., means-end-readinesses and means-end-expectations, and discriminanda- and manipulanda-readinesses and -expectations</u>.

In terms of the preceding discussion we may crystallize the notion of expectation as follows. When an organism expects a state-of-affairs he has the strong or weak conviction, or has the strong or weak belief, that a given state-of-affairs will be brought about or will occur. Again, like demand, an expectation is appropriately expressed in functor notation as

$$e(x, p)$$

which means 'the strength of x's expectation that p will be true'. Conventions allowing one to say that an expectation exists or does not exist are the same as those for demand.

29. Let us consider the notion of mean-end-readiness in relation to that of means-end-expectation. In this essay, to say that x has a readiness for p as a means to q is to say that x expects that p implies q. In symbolism this may be written:[17]

$$r(x, p, q) =_{Df} e(x, p \supset q).$$

[This definition is informal and will not be listed among the official definitions in the present system.]

The reasons for adopting this convention are twofold. First, interpreting means-end-readiness in this fashion does not

---

17. It should be noticed that a means-end-expectation is simply a substitution instance of 'x expects that p is true', that is, of '$e(x, p) > 0$'. In the definition above '$p \supset q$' has been substituted for 'p'.

perceptibly change the working consequences of Tolman's theory. Indeed, insofar as the working theory is conceived, means-end-expectations do the same job as means-end-readinesses. Lastly, the differences between means-end-readiness and means-end-expectancy, for example, mean-end-expectancies are regarded as "specific" instances of means-end-readinesses,[18] are concerned with a much more general account of Tolman's theory of learning than is being attempted in this essay. As such, what is said in this essay is not genuinely affected by such differences between means-end-readinesses and means-end-expectancies. Tolman's later work seems to support these conclusions. For there is great emphasis on the notion of expectation and little on the notion of readiness. This seems to be evidence that Tolman himself, insofar as he is concerned with the empirical working consequences of his theory, is really concerned with expectations rather than readinesses.

30. The notion of "means-end-demand" may be defined in terms of the concepts of "demand" and "means-end-expectations". A means-end-demand may be "defined" as

$$d(x, p, q) > 0 =_{Df} [d(x, q) > 0 \, . \, e(x, p \supset q) > 0],^{19}$$

---

18. Purposive Behaviorism, Chapter VI.
19. There are certain formal difficulties with this definition which, however, do not adversely affect the development of the present system. Hence, if the reader wishes, he may take '$d(x, p, q)$' as another primitive and replace the above definition with the postulate '$d(x, p, q) > 0 \equiv [d(x, q) > 0 \, . \, e(x, p \supset q) > 0]$'.

that is, 'x demands p as a means to q' means 'x demands q and expects that if p then q'. Evidence for the legitimacy of the definition is found in the following passage on page 29 of Purposive Behaviorism.

> It will be asserted next that the subordination of such secondary demands to the superordinate ones is due to the interconnection of what we shall designate as means-end-readinesses. The rat, who because hungry and satiation-demanding therefore demands food, exhibits, we should assert, an interconnecting means-end-readiness, viz., the readiness for commerce with that type of food as a means to satiation. Similarly, the rat who, because food-demanding, is therefore peculiarly ready for explorable objects, exhibits another more subordinate means-end-readiness, viz., the readiness for commerce with such and such explorable objects as the means to food.

31. The discussion of expectation ends with consideration of a problem having to do with exploratory goal-objects and means-end-demands. Tolman seems to believe that under the exploratory demand, every object is demanded as a means object. For example, consider the following passage on page 33 in Purposive Behaviorism.

> It appears, in short, that it is the hungry, or satiation-demanding rat who is the exploration-demanding rat. And further it also appears that at the height of hunger, the exploratoriness is specifically directed towards food.

Here Tolman is offering the assumption that in all cases where an organism has both a demand for the alleviation of exploration and a demand for some other goal-object, say, hunger alleviation, when the demand for hunger-alleviation gets great enough, the demand for exploration takes on the character of a means goal-object. In effect, this assumption seems to be saying that the demand for the alleviation

of exploration is always instrumental in character. This assumption seems dubious. For the exploratory demand is, in some cases, what Tolman calls an "ultimate demand" and hence not always a means demand. It is not being denied that what is learned under exploratory demand is available for use in order to satisfy a demand for, say, hunger satiation. However, it is being denied that in fact such knowledge is always utilized by the organism when his demand for hunger satiation, or the like, reaches a certain intensity. It is only utilized when the organism establishes a means-end-expectation which enables him to connect the exploration demand with the demand for hunger-alleviation. Speaking plainly, some organisms fail to make this connection because they are probably not as smart as others.

32. The above conclusion is supported by the Buxton study on latent learning. In this study, (1) different groups of rats were allowed to explore a 12 unit T-maze for three, four, six and nine nights respectively; (2) the animals were deprived of food for 48 hours and then fed in the food boxes in the maze for 20-30 seconds; and (3) the animals were put in the starting boxes of the maze and allowed to run the maze with no possibility of retracing.. These latter were the test runs. Fifty percent of the animals attained the error criterion on the first run. The animals' learning was said to be latent because they demonstrated their knowledge which was acquired in the exploratory period when other conditions, that is, hunger, called such learning to

the fore.[20] However, the animals that we are interested in are the ones who did not reach the error criterion on the first run or indeed on any run after the first run. Tolman's initial assumption requires that all of the animals in Buxton's experiment reach the error criterion. This did not occur. Moreover, Thistlethwaite, in his review of latent learning, writes:[21]

> Additional amounts of exploratory training did not produce larger percentages of animals that could meet the error criterion.

Hence, it seems plausible that perhaps some of the animals had a demand for exploration as an ultimate goal in itself. Further, the fact that various animals would have had to take a greater number of trials to reach the criterion, if indeed they ever did reach the criterion, would seem to indicate that a certain connectivity factor is necessary to explain their temporary or permanent inability to establish, in Tolman's parlance, a means-end-expectation between the exploration demand and the hunger-alleviation demand. The point is this: Tolman has over-intellectualized the concept of expectation. That is, though the experimental observer perceives that there is a connection between the exploratory situation and the hunger, thirst situation, this does not permit him to suppose that the learning organism has either a built in means-end-expecta-

---

20. Hilgard, E., Theories of Learning, p. 285.
21. Thistlethwaite, D., Critical Review of Latent Learning and Related Experiments, Psych. Bulletin, v. 48, 1951.

tion for the former situation relative to the latter or that he _automatically_ acquires this readiness at some stage in the test situation. Tolman seems to have forgotten that a means-end-expectation is a cognition relative to the cognizer. Learning does not proceed in an "as-iffy" fashion.

The Primitive Idea of Sensory Reception

33. The notion of "sensory reception" is a replacement for Tolman's "commerce-with". The primitive ideas of TI are all construed as relations of one sort or another which stand between organisms and _propositions_ or _situations_; the relation of "commerce-with" obtains between organisms and _things_. This is the principle difference between the meaning "sensory reception" and the meaning "commerce-with". On page 440 of _Purposive Behaviorism_, he writes:

> _Commerce-with._ Any behavior-act in going off involves an intimate interchange with (support from, enjoyment of, intercourse with) environmental features (discriminanda, manipulanda, and means-end-relations). For such interchanges or enjoyments with behavior-supports (q.v.) we have coined the term commerce-with.

t should be noted that "environmental features", "disiminanda", "manipulanda" are thing-like in character. ice, the definition of "commerce-with" betrays the very ing Tolman emphasizes in Chapter V of _Purposive Behavior-n_; "commerce-with" is a relation between _organisms and_ _ings_. If one substitutes for the expression "environmental features", the expression "environmental situations" and omits "discriminanda" and "manipulanda" in the above quotation, one thereby obtains a pretty fair picture of the meaning of sensory reception.

34. In the above explanation of the meaning of "sensory reception", appeal is made to expressions such as "intimate interchange with", "enjoyment of" and "intercourse with". As in the case of 'expectation', these concepts do not imply consciousness or awareness. These expressions suggest a very important characteristic in the meaning of sensory reception, namely, that of immediate verification or confirmation. Indeed, what is here being called "sensory reception" seems quite similar to Carnap's notion of <u>direct confirmation</u>;[22] confirmation by immediate sense-inspection. For example, the state-of-affairs 'There is a white card before x' is the kind of thing sensorily received by x; it is what one means when one says "x sees the white card before him". Furthermore, this "intimate interchange" between x and the state-of-affairs 'There is a white card before x' amounts to a direct confirmation by x <u>that</u> there is a white card before x. Confirmation, in Carnap's sense, is synonymous with "takes to be true".

35. In TI, we treat sensory reception as an ordinary relation (that is, not as a functor) holding between organisms and states-of-affairs. Again, sensory reception is treated in the past tense, that is, as sensorily received.

---

22. Carnap, R., Testability and Meaning, Yale University, 1950. (See especially pages 455 and 11)

This relation may be symbolized as follows:

$pR^{e}x$,

which means 'p is sensorily received by x'.

36. The kinds of things which may be substituted for 'p' may be either discriminanda situations, for example, 'There is a white card before x', or manipulanda situations, for example, 'x runs down the alley'. Tolman's distinction between discriminanda and manipulanda is abandoned in the concept of sensory reception. As a matter of fact, Tolman's distinction between manipulanda and discriminanda does not seem to be of very great value either experimentally or methodologically. For example, Tolman, on page 86 of Purposive Behaviorism, writes:

> Actually, of course, in any given case, a rat will be set for, i.e., expect, both discriminanda and manipulanda. Thus in Elliott's experiments the rat obviously was actually set for, not only the immediate maze-parts to "look" and "smell" and "feel" so and so, i.e., discriminanda, but also for them to "be" so and so for the purposes of running, turning, swimming, i.e., manipulanda. And, conversely, in Carr and Watson's and in Macfarlane's experiments, the rat was, of course, actually set for not only the maze-alleys to "be" so and so for running, swimming, and the like, but also for them to "look" and "feel" and "smell" so and so. The one sort of "set" (expectation) can never actually function without the other.

Now if it is true that an expectation of the discriminanda cannot occur without an expectation of the manipulanda -- and vice versa -- the purpose in discriminating between them becomes quite unclear from either a methodological or experimental viewpoint. For, in either case, it is clear that the consequences of the one will be the same

as the consequences of the other.

37. Treating sensory reception in the past tense has another advantage. It emphasizes the backward-pointing character of sensory reception. In this matter, sensory reception is different from demand and expectation and is like the next primitive idea, "confirmation". For example, when 'There is a white card before x' is sensorily received by x, the sensory reception, that is, the direct confirmation, postdates the occurrence of the state-of-affairs.

38. There are two principle reasons for requiring a concept like sensory reception in Tolman's system. First, it has value in explicitly limiting the range of situations which may be characterized as stimulus situations. This latter notion is somewhat vague since the determination of the spatio-temporal limits and the degree of intensity necessary for reception-excitation have not been generally established to the satisfaction of all psychologists or physiologists. As such, the elimination of the vagueness associated with the notion of sensory reception is an empirical problem and thus falls outside the purview of this essay. Let me proceed to an illustration.

39. It is well known that the conditions which Tolman puts upon the responses of an organism include those of demanding and expecting. That is, before an animal responds to a given situation, -- is stimulated by that situation, -- he must not only demand it but expect it. This

[illegible]

is a restatement of Tolman's position that not all situations in environmental proximity to the animal are stimulating situations. In more established terms, unless the animal "attends to" the situation, it will not evoke a response. But now certainly Tolman does not intend that any old demanded and expected situation will induce a response in the organism to that situation. For example an animal might expect that if he's put in the starting box of maze A and furthermore runs down alley B of that maze, then he gets to food. Furthermore, he may be hungry, and may demand running down the alley. But as long as the animal is in some way prohibited from the sensory reception of being in the starting box he will not respond -- that is, he will not run down alley B! In short, only sensorily received situations are stimulating situations.

40. The second reason why sensory reception is required in TI is this: it assumes, in part, the responsibilities delegated to the notion of "commerce-with" in Tolman's theory. Since each primitive idea in TI takes as values of its variable organisms or states-of-affairs (or situations), the notion of sensory reception as a replacement for "commerce-with" is required for the sake of the coherence of TI.

The Primitive Idea of Confirmation

41. The difference between confirmation and sensory reception is a matter of degree. Indeed, "confirmation" might

be explained as an <u>indirect</u> interchange with, an <u>indirect</u> enjoyment of, or an <u>indirect</u> intercourse with environmental situations. As sensory reception corresponds to Carnap's direct confirmation, the present concept of confirmation corresponds to Carnap's notion of <u>indirect confirmation</u>.[23] To put it another way, to say that a state-of-affairs has been confirmed is to say that its confirmation is based on confirmation by sense inspection (sensory reception) of certain other states-of-affairs. For example, consider the case where x has sensorily received that there is an open alley before him, that is, x "sees" the open alley before him, and has sensorily received that the door is open at the end of the alley. On the basis of this information we might say that x has confirmed (by induction) that the alley which turns to the left at the open door leads to food. Confirmation, in our sense, is an indirect taking to be true; an inductive taking to be true.

42. Confirmation, like sensory reception, is (1) treated in the past tense and is (2) taken to be an ordinary relation between organisms and situations. Hence we may write:

xCp

which means 'x has confirmed that p is true (or exists)'. Again, like sensory reception, confirmation is backward

---

23. Op. cit., <u>Testability and Meaning</u>, pp. 11 and 455.

pointing. That is, when x has confirmed that the alley which goes left is free of obstacles, the confirmation of this situation postdates the occurrence of that situation.

43. The next consideration has to do with the forward pointing character of expectation chains. That is, before a given stimulus can "release" various expectations in a chain, the conditions for the releasing of these expectations must be, at least, indirectly confirmed. Let us examine a concrete example. Suppose that we have trained an animal in a straight alley maze. Suppose further that the animal has built up a means-end-expectation (means-end-readiness) during this training that if when he is in the starting box he runs down the alley, he gets to food. In the test series our animal is put in the starting box. This is a stimulus situation. The means-end-expectation that if he runs down the alley then he gets to food is thus "released". This means-end-expectation is "released" by the fact that the animal has confirmed or verified the first term in the expectation chain, that is, he has confirmed that he is in the starting box. The result is that the animal's confirmation of being in the starting box takes on the character of a stimulus which in turn "releases" the expectation of getting to food. The following passage on page 85 in _Purposive Behaviorism_ seems to support this view.

> This statement of the duality of stimulation and verification is but another way of saying that any behavior-act requires not only stimuli to release

> it, but also, later, more substantial environmental actualities such as fulcra, media, and planes, to support it (i.e., verify it, make it possible). Stimuli by themselves are not enough; supports also are needed. Behavior cannot go off in vacuo. It requires a complementary "supporting" or "holding-up". The organism, as a result of stimuli, expects that such and such "behavior-supports" are going to be in the environment. A rat cannot "run down an alley" without an actual floor space ahead to catapult into. And in a discrimination-box, he cannot "choose" the white side from the black without actual choice. Behavior-acts and their immanent expectations are released by stimuli; but they demand and are sustained by later coming behavior-supports.
>
> In parenthesis it may be remarked that this fact that supports, and not merely stimuli, are needed for the actual going-off of any act and are expected by such an act, is a feature about behavior which orthodox psychologies, both stimulus-response psychologies and mentalisms, seem hitherto to have overlooked.

And further on pages 86 and 87, Tolman writes

> The final definitions of expected discriminanda can, it must now be noted, be determined in the last analysis only by a whole series of experiments. Suppose, for example, we discover that a given behavior seems to be "expecting" the presence of a certain specific color. This behavior it is found, will, under the given conditions, continue to go off only so long as this specific color proves actually to be there.

44. One final point; in the above illustration we use the more general "has confirmed" in order to release the expectation for two reasons. First, Tolman himself takes a very broad view of what "releases" or "fires" a chain of expectations as is evident in the above quoted passages. Secondly, it is a law in the present system that sensory reception implies confirmation. In symbols, this law is written:

$$ph^{e}x \supset xCp.$$

Hence, in Carnap's terms, if we allow that an indirectly confirmed state-of-affairs may set off a chain of expectations then we must allow that a directly confirmed state-of-affairs may set off an expectation chain. Indeed, the illustration in paragraph 43 depicts direct confirmation as releasing the described expectation. Nevertheless, an expectation chain can be set off by indirect confirmation, for example, by the kind of induction described in paragraph 41.

The Primitive Idea of Response Tendency

45. The term "response" is used in at least two distinct ways by psychologists. It often means each individual occurrence of a certain type of activity which is initiated by a given state-of-affairs. For example, when I stick my finger with a pin, I respond by snapping back my hand to the left; or I may snap my hand back to the right, back and above, etc. The point is this: the type of activity constituting the response to the stimulus which is described by the proposition 'x sticks his finger with a pin' is that of "snapping back the hand". Let us designate this meaning of response as 'R_1'. The second meaning of response is more accurately described as the frequency of occurrence of R_1. For example, consider the case where the psychologist wants an animal to learn always to jump to the white card in the jumping stand. Suppose the criterion of learning is 14 correct choices in 15 consecutive jumps. In such a case, the psychologist often means

by "response" the frequency of occurrence of a certain type of activity, for example, here, 'jumping to the white card', which is initiated by a given state-of-affairs. In other words, to say the organism has performed is to say the organism has responded, in this latter sense, to a given state-of-affairs. Let us designate this second meaning of the term "response" as 'R_2'. In TI, "response" means response in the sense of R_2; more specifically, R_2 captures the meaning of "response tendency".

46. "Response" in the sense of 'R_2' clearly is an arithmetical variable. For example, during a block of five trials it might be predicted that the organism would make three correct responses (in the sense of R_1) out of 5 possibilities. The organism would then be said to have the tendency to respond in a given manner to a given state-of-affairs 3 times out of 5; or at the ratio of 3/5. It is clear that the notion of response tendency is easily expressible in functor notation. How do we write this idea?

47. Notice first that the concept of response tendency stresses the particular type or kind of resultant activity to a given state-of-affairs. It suggests further that the only conditions under which one may be allowed to say that the proposition x responds by so and so to such and such a state-of-affairs is when the frequency of that activity reaches a certain predetermined level. Again, this resultant activity is causally consequent on an initiating state-of-affairs. With these points in mind the notion of res-

ponse tendency may be written as essentially a functor having three arguments, that is, as

$r(x, \varphi, p)$

which means 'the tendency of x to respond by the type of activity φ to the state-of-affairs p'; 'φ' is a predicate variable, 'x' is an organism variable, and 'p' is a propositional variable. Hence, in accordance with the remarks at the beginning of this paragraph, if we let the number 1 be the predetermined learning or performance criterion, then the proposition that x responds by φ to p is the same as saying that

$r(x, \varphi, p) = 1.$

Indeed, the following law would seem to hold:

(29) $r(x, \varphi, p) = 1 \supset \varphi x,$

that is, 'if x's tendency to respond by φ to p equals 1 then φx is true'. Interpreted, this law might read in a given case, 'if x's tendency to respond by "jumping to the white card" to the state-of-affairs "there is a white card before x" is equal to 1, then "x does jump to the white card". When the conditions in law (29) are met we might say that 'φ' is a realized predicate. When the number is less than 1 but greater than 0 (the performance criterion number), that is, when

(30) $r(x, \varphi, p) = N \,.\, 0 < N < 1$

is true, 'φ' is a realizable predicate. Finally, to say that x has a tendency to respond by φ to p we may write:

$r(x, \varphi, p) > 0.$

[illegible]

of [illegible]

pre[illegible]

is [illegible]

how[illegible]

we [illegible]

the [illegible]

alt[illegible]

[illegible]

for[illegible]

in [illegible]

the [illegible]

[illegible]

the [illegible]

[illegible]

co[illegible]

in [illegible]

la[illegible]

[illegible]

[illegible]

an[illegible]

[illegible]

the [illegible]

of [illegible]

for[illegible]

[illegible]

48. There are certain problems connected with this view of response tendency. Notice that, in a given situation, predicting a certain strength of response tendency say,

$$r(x, \varphi, p) = .70,$$

is like assigning a certain probability to '$r(x, \varphi, p)$'. How to make this assignment is the problem. Generally, we would make this assignment in one of two ways: (1) on the basis of the <u>known</u> total number of responses in a given situation or (2) on the basis of the number of possible responses or trials in arbitrarily selected blocks of trials for each of those blocks --say, blocks of 5 trials each. In the case of (1), we seldom <u>know</u> in a given situation the total number of responses which it is possible for the organism to make and hence the basis of the assignment of the probability number to '$r(x, \varphi, p)$' is questionable. We might fall back on method (2). But obvious difficulties occur here also as soon as we begin comparing the increase in tendency to respond in, say, blocks <u>a</u> through <u>b</u> with later blocks <u>d</u> through <u>e</u> or the increase in tendency to respond in the total blocks of trials in one study with the total blocks of trials in another study where the two totals are different. The solution to these problems, however, lies outside the scope of the present study. It is the hope of the present author that despite these difficulties the notion of response tendency may prove to be useful to the investigator at least as a suggestive device.

49. The final point concerns the circularity of stimulus

and response. Much ado is made in systematic courses in psychology about the so-called circularity of stimulus and response. The discussion usually proceeds in this fashion: "In order to study behavior we must start with certain understood concepts. Among these understood concepts ('stimulus' and 'response'), the one which is not taken as undefined or understood is defined in terms of the other. Here a special problem arises. Suppose, for example, we take 'response' as understood. Then we define a stimulus, quite generally, as 'that which evokes a response'. But now suppose we ask what it means to be a response? The answer to this question amounts to this: a response is 'that which is evoked by a stimulus'. Hence we have the problem of the circularity of stimulus and response." I should like to point out that the problem of circularity of stimulus and response, insofar as the logic of system building is concerned, is a pseudo-problem. In the above example the question concerning the meaning of response in invalid. First of all, there may be an ambiguity in the question -- probably associated with the speaker's use of the word "meaning". Does the question ask for a <u>definition</u> of response or does it ask for a <u>characterization</u> of response -- for example, such as one gives in the case of the primitive idea of <u>alternation</u> in the system of <u>Principia</u>? If the former is meant, the question is clearly improper. For the question would be asking: what is the definition of the undefined term "response"? On the other hand, if the second case is intended there would be no question of circularity because the intent here is not to define response, but rather to give some clue as to how the term "response" is to be interpreted. The circularity then is merely apparent because the intention is not to define "response" in terms of "stimulus", but rather to characterize or explain, independently of its occurrence in the system, its general use. Indeed, in the present system, on the basis of the above account of response tendency, it is very easy to define "stimulus" in the following terms:

$$pSx =_{Df} (\exists\varphi)(r(x, \varphi, p) > 0),$$

that is, 'p is a stimulus of x' means 'There is a φ such that x has a tendency to respond by φ to p'.

Two Concepts of Docility

50. In Tolman's theory the concept of docility is difficult to grasp. The difficulty is largely the result of using "docility" in two distinct ways. Strictly speaking, there is no single primitive idea of docility in Tolman's theory; rather there are at least two distinct primitive ideas of docility. The first meaning of "docility" is "teachability". The second meaning of "docility" is "taught". Surely these are distinct concepts -- though they do have some features in common, for example, both are variables. Let us take up these ideas one at a time.

51. When one explains "docility" as meaning "teachability" he is treating docility as a disposition predicate. A dis-

position predicate designates a disposition or capacity of the organism. This does not mean that the capacity is always being put to work, that is, is active. We characterize people as having drives even when they are not in any way impelled to act in accordance with that drive. In short, a disposition is similar to what John Locke refers to as the "power" to do something or act on something.

52. "Teachability" means essentially "the capability of being taught". Thus it is possible to define "teachability" in terms of "taught". However, the meaning and formulation of "capability" is an extremely difficult and subtle thing to accomplish. Both of these terms, then, will be regarded as primitive in TI.

53. "Teachability" is a variable. For example, we may characterize one animal as more teachable than another. Again, "teachability" is a predicate holding between organisms and states-of-affairs which are means to other states-of-affairs. Teachability is thus appropriately expressed in functor notation as

$$t(x, p, q)$$

which means 'the degree to which x is teachable with respect to p as a means to q'. Other conventions governing the use of functors as described in paragraph 4. of this chapter are appropriate here.

54. When one treats "docility" as "taughtness" he is treating docility as a non-dispositional predicate. As

the dis

and 2 a

sense

55.

etc.

and st

which

is a

56.

suggested above, "taughtness" is really the activation of the disposition. Thus when the child adds the figures 2 and 2 and gets 4 rather than 5, we say he is docile in the sense of being taught. He has been taught that the correct addition is that 2 and 2 are 4. In general, when an organism is docile in the sense of being taught, one is saying that he has been led to see the "correct" (best, most efficient, etc.) means of solving a problem.

55. "Taughtness" like "teachability" is a variable. Witness, for example, the following cases; "So and so has been better taught", "So and so has been really taught!", etc. It is, too, a relation holding between an organism and at least two states-of-affairs. Hence it may be described in functor notation as

$$t^A(x, p, q)$$

which means 'the degree to which x has been taught that p is a means to q'. Other conventions governing the employment of functors as described in paragraph 4. of this chapter are appropriate here.

56. The employment of "docility" as "teachability" by Tolman is evidenced in his explanation of the phenomenon of disruption.[24] Tolman explains that when a "better" goal-object is substituted for a "poorer" one during the course of an experiment, the behavior of the animal will indicate

---

24. Purposive Behaviorism, p. 74.

"surpris

error o

provided

respect

axial w

the bett

power

as "best

has not

ible to

taught.

is that

one the

that is,

more th

other wo

able wi

57.

as evid

the or

the co

that x

a state

given

notes

as a re

wal is a

"surprise", that is, the behavior will show a trial-and-error character. However, after the disruption effect, provided that the animal is (and has been) docile with respect to the better goal-object, the behavior of the animal will tend to become systematic and the choices of the better goal-object will exceed the choices of the poorer goal-object. Here Tolman is employing "docile" as "teachable"; for the animal at the time of disruption has not been taught anything. To be sure, it is not possible to determine teachability unless the animal can be taught. But the point above that Tolman is emphasizing is that the disposition of teachableness is required throughout the experiment in order that the animal can be taught; that is, will choose the better or more demanded goal-object more times than the poorer goal-object after disruption. In other words the animal will not recover unless he is teachable with respect to the better goal-object.

57. In contrast, Tolman's characterization of "docile" as evidenced by the fact that the organism always picks the more efficient of the routes to a final goal construes the concept of docility as "taught". In other words to say that x has been taught that a state-of-affairs p leads to a state-of-affairs q amounts to the claim that for any given set of routes, if p is the most efficient of those routes to q, x will choose p more often than any other route as a means to q. Of course, this would imply that the animal is also "teachable". But the above is an explanation

of an <u>active</u> <u>teachability</u> that is, of a "taughtness", if you will.

58. Two final points: (1) It is evident from the preceding discussion that "taughtness" implies teachability. The converse is not true. It follows that any law holding for teachability holds, also, for taughtness: (2) The concept of taught is required in order to make more clear the concept of fixation employed by Tolman. For though we may say that when someone is fixated on something 'p' he has not been taught with respect to p, we may not say that when he is fixated he is not teachable. For if this were true we could never break down a fixation. This point will be considered in the discussion of the theorems of TI in the next chapter.

A List of The Primitive Ideas

59. The primitive ideas of TI are listed in the order of their occurrence in this chapter.

P1 '$d(x, p)$' means 'the strength of x's demand that p is true'.

P2 '$e(x, p)$' means 'the strength of x's expectation that p is true'.

P3 '$p\hbar^{e}x$' means 'p is sensorily received by x'.

P4 'xCp' means 'x has confirmed that p is true'.

P5 '$r(x, \varphi, p)$' means 'the strength of x's tendency to respond by φ to p'.

P6 '$t(x, p, q)$' means 'the degree to which x is teachable relative to p as a means to q'.

P7 '$t^A(x, p, q)$' means 'the degree to which x has been taught that p is a means to q'.

The Definitions of TI

60. Because some of the postulates of TI are expressed in terms of defined ideas rather than in terms of primitives, it is appropriate to list the definitions first. There are five definitions.

TI. 1.0 $d(x, p, q) > 0 =_{Df} [d(x, q) > 0 \,.\, e(x, p \supset q) > 0]$

This is the definition of means-end-demand. It reads: 'x demands p as a means to q' means 'x demands q and expects that if p then q.

61. The following passage on page 29 of Purposive Behaviorism supports the accuracy of this interpretation of means-end demand.

> It will be asserted next that the subordination of such secondary demands to the superordinate ones is due to the interconnection of what we shall designate as means-end-readinesses. The rat, who because hungry and satiation-demanding therefore demands, food, exhibits, we should assert, an interconnecting means-end-readiness, viz., the readiness for commerce with that type of food as a means to satiation. Similarly, the rat who, because food-demanding, is therefore peculiarly ready for explorable objects, exhibits another more subordinate means-end-readiness, viz., the readiness for commerce with such and such explorable objects as the means to food.

TI. 1.1 $pMx =_{Df} (\exists q)[d(x, p, q) > 0]$

This is the definition of means-object. It reads: 'p is a means-object of x' means 'There is a q such that x demands p as a means to q'.

TI. 1.2 $pUx =_{Df} [d(x, p) > 0 \;.\; \sim(\exists q)(d(x, p, q) > 0]$

This is the definition of an ultimate goal-object. It reads: 'p is an ultimate goal-object of x' means 'p is demanded but p is not demanded as a means to any other state-of-affairs, q'.

62. In Purposive Behaviorism, Tolman's remarks on the notion of ultimate goal-object are a bit confusing. Consider, for example, the following passage on page 36 of Purposive Behaviorism.

> The ultimate goal-objects for the rat, or other animal, are certain finally-to-be-sought, to to-be-avoided, physiological states of quiescence and disturbance due to initiating physiological states or conditions. Subordinate to the demand for or against these there are various types of environmental presence, e.g., food, electric grill, etc., demanded for or against as the last steps in reaching such quiescences or in avoiding such disturbances. And subordinate to these latter there may be still other types of still more subordinate environmental presences also to be sought and avoided.

There are really three kinds of goal-object described in this passage. The first sentence describes what Tolman calls an ultimate goal-object. The second sentence describes what might be called a final goal-object. And the last sentence describes a means-object. The confusion arises in this way. Tolman often uses final and ultimate goal-objects as synonyms. Thus an ambiguity is created. In this essay ultimate (or final) goal-objects are used in the sense in which we talk about hunger-satiation, thirst-satiation, and the like as ultimate goal-objects.

Hence, what Tolman takes as final goal-objects, drinking water, etc., shall be construed as means-objects. Other considerations reinforce this decision. For example, it makes little sense to speak of conflicting ultimate (physiological) goals. If an organism is hungry he does not have conflicting demands about the demand for hunger satiation. But he may be in conflict with respect to the means for satisfying hunger. Accordingly, if ultimate goal-objects were situations like "eating food" etc., we should have to reject the above view about conflict with respect to ultimate goal-objects. For it is clear that we <u>can</u> be in conflict with respect to "eating food". Furthermore, it seems that the employment of "ultimate goal-object" described above is consonant with the way in which Tolman actually uses it in his interpretations of experimental results.

63. There is another point in this definition of ultimate goal-object. The above definition should not be taken as categorizing in some absolute manner those goal-objects which are ultimate from those which are merely means. Rather it should be taken as suggesting that in <u>any</u> given situation what is designated as an ultimate or final goal-object is not a means to any other goal-object. This means that an ultimate goal-object in one situation, for example, alleviation of need for urination, might be taken as a means object in another situation; for example, on a trip, as a means to enjoying the ride in the automob-

ile. Indeed it may prove to be the case that one and the same goal-object can be treated as both a means and/or an ultimate goal object in the same situation. If this latter case were found to be true, TI. 1.2 would have to be abandoned.

TI. 1.3 $F(x, p, q) =_{Df} [d(x, p, q) > 0 \,.\, \sim(\exists r)[xC(r = \sim p) \,.\, d(x, r, q) > 0]]$

This is the definition of "fixation". It reads: 'x is fixated on p as a means to q' means 'x demands p as a means to q and there is no r which has been confirmed by x to be the same as not p and which is also demanded by x as a means to q'. In effect, this definition shows the outright "stubbornness" or rigidity of the fixated organism; drastic means are needed in order to change his course.

64. In this definition, I have departed from Tolman's writings, but, I think, not from his real intentions. First of all Tolman's discussion of "fixation", "cognition" and "means-end readiness (or expectation)" in Purposive Behaviorism leads to an inconsistency in his system. Consider the following statements.

(a) Something is a cognition if and only if it is a means-end-readiness (or expectation).[25]

(b) No means-end-readiness is cognitive, if it is fixated.[26]

---

25. Ibid., p. 440.
26. Ibid., p. 94.

(c) Some means-end-readinesses are fixated.[27]

From this set of propositions it is rather easy to deduce the contradiction of (c), that is the statement which reads: No means-end-readiness is fixated. As can be seen in the above set of statements, the problem consists in being unable to determine whether or not there is such a thing as a fixated means-end-readiness. Tolman seems to say that "you can" and "you can't" have such a state-of-affairs (cf. especially his discussion on pages 29-31 in Purposive Behaviorism). But, as the above argument shows, this situation is absurd. I have therefore taken the liberty of rejecting the view that when an animal is fixated he has no cognition. This amounts to a denial of the statement (b) above. Hence, I arrive at the already proposed definition of fixation. From this definition we can conclude that

$$F(x, p, q) \supset (\exists q)\,(d(x, p, q) \neq 0),$$

that is, when an organism is fixated on p, he demands p as a means to another state-of-affairs q. That this is Tolman's real intention is supported, in the following passage on page 31 of Purposive Behaviorism.

> It must be pointed out finally, however, that unless the rats would show a tendency, on further training, finally to be docile to this fact that the larger seeds are really no better, and, if anything, worse than the smaller seeds, then the above described sub-

---

27. Ibid., p. 31.

ordinate readiness for the larger seeds would after all be not truly judgmental. In this latter case it would have to be conceived, rather, as of the nature of what we might call a means-end-fixation.

TI. 1.4 $\quad pSx =_{Df} (\exists\varphi)\ [r(x, \varphi, p) > 0]$

This is the definition of "stimulus". It reads: 'p is a stimulus of x' means 'There is a φ such that x tends to respond by it to p].

CHAPTER III

POSTULATES AND THEOREMS OF SYSTEM TI

1. This chapter presents the postulates and theorems of TI. The postulates are distinguished from the definitions and theorems by the prefix,

TI. 2.N.

The whole number which replaces N in the above prefix is the number of the particular postulate. An analogous convention is adopted for the theorems. That convention will be discussed when we come to the presentation of the theorems.

TI. 2.0 $[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0] \supset e(x, p \supset r) > 0$

TI. 2.0 reads: if x has confirmed p and expects that p implies q and also expects that q implies r, then he expects that p implies r.

2. TI. 2.0 affirms the "restricted" transitivity of implicative expectation. It permits the development of an expectation chain. The addition of 'xCp' to the antecedent is a necessary restriction. For, to use Tolman's words, expectation chains cannot be developed "in vacuo"; they require "supports". Again, TI. 2.0 suggests a certain provisional character to all chains of expectations. For example, consider the case of an organism who has built up the following means-end expectations. "I expect that when

I am put in the starting box, I can turn right. I also expect that if I turn right I will get to food". Now the question is this: under the supposed circumstances will the organism expect that if he is put in the starting box then he will get to food? The answer is: it depends. He will if he has some evidence for the truth of p; that is, if in some sense, he has "taken p to be true" - if he has confirmed that p is true. For example, in the above situation, it would be quite possible on, say, the second or third trial, where the animal is not put in the starting box, but rather is put on an open stand in the same position as the original starting box, that the animal would not expect that being put in the starting box was (or is) a means to getting to food.

TI. 2.1 $[d(x, q) > 0 \,.\, e(x, p \supset q) > 0] \supset d(x, p) > 0$

TI. 2.1 reads: if x demands q and expects that if p then q, then x demands p.

3. There are four points to be considered in the discussion of TI. 2.1 (1) TI. 2.1 suggests that the development of means-objects moves backward. Essentially, this is Tolman's contention, that states-of-affairs which are expected as means to others may become demanded objects in their own right. That is, that a state-of-affairs, p, which is expected as a means to another, q, is itself a demanded object is contingent upon the presence of a demand for q.[1]

---

1. Tolman explains this principle and gives the evidence for it on pages 148-151 of <u>Purposive Behaviorism</u>.

(2) The opposite case, that is,

$$[d(x, p) > 0 \; . \; e(x, p \supset q) > 0] \supset d(x, q) > 0,$$

which reads, "if x demands p and expects that p is a means to q, then x demands q", is sometimes false. This is evident upon the substitution of the propositions 'x crosses the maze' for 'p' and 'x will be shocked severely' for 'q'. (3) 'p' in TI. 2.1 is a means-object. A theorem later in this chapter will demonstrate this point. In this sense, TI. 2.1 differs from the next postulate which is the "correlate" of TI. 2.1 for avoidance objects. For in the case of the "correlate" of TI. 2.1, '~p' cannot be shown to be a means-object. (4) TI. 2.1 describes the typical approach situation of ordinary psychological discussion and investigation.

TI. 2.2 $\quad [d(x, \sim q) > 0 \; . \; e(x, p \supset q) > 0] \supset d(x, \sim p) > 0$

TI. 2.2 reads: if x demands ~q and expects that p implies q, then he demands ~p. Another way of reading this postulate is: if q is an avoidance object and x expects that p leads to q, then p is an avoidance object.

4. TI. 2.2 is the correlate of TI. 2.1; it expresses the typical avoidance situation. Remarks (1) and (2) in the discussion of TI. 2.1 apply equally well in the case of TI. 2.2.

TI. 2.3 $\quad [d(x, p) > 0 \; . \; d(x, \sim p) > 0] \supset pMx$

TI. 2.3 reads: if x demands p and also ~p, then p is a means-object of x.

5. The antecedent of TI. 2.3 is not to be construed as a conflict-of-demand situation. In paragraph 10. of chapter II a typical conflict situation was described. Notice that the two conflicting cases described in paragraph 10. not only showed x as demanding both p and ~p but also that x demanded these contradictory states-of-affairs with <u>the same</u> strength. The antecedent of TI. 2.3 requires only that the two demands are greater than 0. It is not difficult to find a case in which an organism can demand contradictory states-of-affairs at different strengths and still not be in conflict. For example, suppose an animal to have built up expectations that two and only two mutually exclusive alleys, one longer than the other, lead to food. If he has a demand for hunger alleviation, then in accordance with postulate TI. 2.1, he will have a certain demand strength for running down either alley. Let us call the longer alley 'A' and the shorter alley 'B'. The demand strength for running down alley A will be less than for running down alley B (in accordance with the law of least effort). But notice that a demand for running down alley B satisfies the proposition that x demands not running down alley A. Where M>N>0, this situation may be described in symbols as follows.

(1) $d(x, \sim A) = M$,

that is, 'x's demand for not running down alley A is equal to M'.

(2) $d(x, A) = N$,

that is, 'x's demand for running down A is equal to N'. The conjunction of (1) and (2) satisfy the antecedent of TI. 2.3, but clearly x does not have conflicting demands here; that is, he is satisfied with this arrangement. The only case where demands are in conflict is where the demands for two contradictory states-of-affairs are at the same, or nearly the same, strength.

6. Nevertheless, we shall want TI. 2.3 as a postulate for the reason that it allows us to prove, that if p is an ultimate goal-object then if x demands p he does not demand ~p. In other words, if p is an ultimate goal-object of x the contradictory of p cannot be demanded (nor indeed can it be an ultimate goal-object). This conforms with the interpretation of goal-object given in TI. 1.2. These statements will be proven as theorems later in this chapter.

TI. 2.4 $[xCp \,.\, e(x, p \supset q) > 0] \supset e(x, q) > 0$

TI. 2.4 reads: if x has confirmed that p is true and also expects that if p then q, then x expects q.

7. We have seen in the discussion of confirmation that if an expectation chain is to be "fired", at least confirmation (in the weak sense in which we employ it in this ssay, that is, as an indirect "taking to be true") of the irst term in that chain is required. This is why Tolman's

use of the word "support" in the quotation cited in that discussion is so appropriate here; an expectation cannot be "released" unless the chain in which that expectation is a member is supported. Anthropomorphically, we do not permit our expectations free rein without grounds or at least certain convictions. TI. 2.4 thus helps guarantee the forward pointing character of expectations; and hence of demands.

TI. 2.5 $[e(x, p) > 0 \ . \ xC(q \supset \sim p)] \supset \sim e(x, q) > 0$

TI. 2.5 reads: when x expects p and when he has confirmed that q implies not p, then he does not expect q.

8. This postulate guards against an overly rationalistic interpretation of expectation. For example, suppose an organism to expect that he will do something at a certain time. Suppose that time is an hour later. Now it seems cogent to suppose that the same organism might also expect to do something else incompatible with the first action at the same time. Of course, if he realized or confirmed that the two actions were incompatible, it seems fair to suppose that he would reject one or the other. It is exactly this kind of situation that TI. 2.5 allows. In other words, an organism might have conflicting expectations so long as he does not verify that the expected states-of-affairs are incompatible. Another and perhaps more intuitive way of showing this restriction can be shown in the following theorem which is deducible from

TI. 2.5, namely,

$$(e(x, p) > 0 \;.\; e(x, \sim p) > 0) \supset \sim xC\ (\sim p \supset \sim p)$$

which, in effect, says that if x has conflicting expectations relative to p, then he has not confirmed that p is incompatible with not p.

TI. 2.6 $[d(x, q) > 0 \;.\; e(x, p \supset (\varphi x \supset q)) > 0 \;.\; pR^{e}x] \supset$
$r(x, \varphi, p) > 0$

TI. 2.6 reads: when x demands q and expects that if p then if φx then q and when furthermore p is sensorily received by x, then x has a tendency to respond by φ to p.

9. Perhaps, it would be helpful to give an interpretation of TI. 2.6 in English. Suppose we have a typical discrimination learning situation. Substitute 'x gets to the food-box in back of the white card' for 'q', 'x jumps to the white card' for 'φx' and 'there is a white card on the left' for 'p'. Then, in this case, TI. 2.6 affirms that when an organism demands that he get to the food-box in back of the white card and furthermore expects that if there is a white card on the left, so that if he jumps to it, he gets to the stand in back of it, and finally sensorily receives that there is a white card on the left, he tends to respond by jumping to the white card in relation to the state-of-affairs that there is a white card on the left.

10. TI. 2.6, more than any other postulate in TI, marks

the difference between Tolman's theory of learning and Hull's s-r reinforcement theory of learning. TI. 2.6 suggests that not every state-of-affairs p which is merely sensorily received is a stimulating state-of-affairs.[2] In addition, p must be "attended to", that is, expected as a means to some other state-of-affairs, if it is to be a stimulus. Hull, on the other hand, denies this contention. To put it briefly, and thus a bit loosely, Hull apparently conceives of every state-of-affairs which is sensorily received as a stimulus.[3] Epigramatically, the issue might be put as follows: Hull believes an organism learns something about everything; Tolman believes that the organism learns something about only something.

11. TI. 2.6 expresses the conditions under which an organism will tend to respond to a given state-of-affairs. However, it is not being claimed that these are the only conditions under which an organism will tend to respond. Under certain circumstances, - for example, when a gun is fired behind an unsuspecting subject's head, -an organism will tend to respond to a state-of-affairs though he does not demand or expect anything relative to that state-of-affairs. Notice, however, that such a stimulus situation, -call it a compulsive stimulus situation, - does not falsi-

---

2. Purposive Behaviorism, pp. 35-36.
3. Koch, S., "Clark L. Hull", Modern Learning Theory, Appleton Century Crofts, 1954 (Cf. especially, p. 9).

fy 1.7.[4] In other words the "promise" that when an organism demands a state-of-affairs q, when he expects that p leads to φx which in turn leads to q and finally when he sensorily receives p, then he tends to respond by φ to p is not broken when a compulsive stimulus situation occurs. Furthermore, the psychologist is really concerned with the non-compulsive stimulus situation. TI. 2.6 is one way of stating the conditions for response in a non-compulsive or "normal" stimulus situation. Clearly the investigator assumes "normality" of the situation in his laws pertaining to stimulation and learning. Furthermore, the fact that the compulsive stimulus situation is not the "normal" kind of stimulus situation makes it a much more rare and hence a much less scientifically interesting situation --at least, as far as learning is concerned.

12. Finally, evidence that this postulate is a reasonable translation of Tolman's views is given in the following passages. On page 35 of Purposive Behaviorism Tolman writes

> Finally, we must now note a situation as regards both rats and other organisms which "stimulus-response" psychologies, as well as "behaviorisms" proper, seem largely to have overlooked. It is a situation which orthodox mentalisms were well aware of. The latter sought to care for it by their doctrines of attention and apperception. It is the fact that rats and men have hundreds, not to say thousands, of stimuli impinging upon them every instant of their waking lives; and yet to by far

---

4. A very good example of a compulsive stimulus situation is Tolman's "initiating physiological state".

> the majority of these stimuli they do not, at the given moment, respond. But in order now, in our system, to explain this choosiness as to stimuli, we have merely to refer to these facts of superordinate and subordinate demands and means-end-readinesses as just outlined.
>
> Consider the case of food-stimuli. The satiated rat "pays no attention" to food. He even lies down and goes to sleep in its presence and so, also, does the satiated human being. The reason the hungry rat is responsive is, we would assert, (a) because he is demanding hunger-satiation, and (b) because he is provided with a means-end-readiness (innate or acquired, "judgmental" or "fixated"), to the effect that commerce with the type of food, presented by the given stimuli, will lead on to satiation.

His most direct statement on this matter appears on page 407 in the same source.

> The organism responds to the given stimuli only, by virtue of an initiating physiological state which, given his innate or acquired means-end-readinesses, gives rise to demands, (superordinate or subordinate) --one or more of which leads him to respond to the given S(timulus) as presenting an appropriate means-object. These depending demands control the whole line of the S(timulus)-- R(esponse) process.

This list of quotations is concluded with a concrete example of Tolman's views concerning stimulation. The passage is found on page 330 of Purposive Behaviorism.

> But it appears obvious, to us at any rate, that the conditioned "approach" response made to S_2, the white side of the discrimination-box, involves different supports, i.e., expectations of, and commerces with different discriminanda and manipulanda, from those involved in the original unconditioned response made to S_1, the food. The response which comes to be made to S_2 is, in short, not only, as we pointed out in the preceding sections, appropriate to the sign-character of this immediate object; --it is not only appropriate to the fact that this object (the white side of the box) is a sign for leading-on-ness to the food (the object presented by the unconditioned

stimulus S_1) -- but it is also appropriate to the immediate discriminanda and manipulanda characters of this immediate S_2 - object.

To sum up, we note: (a) that the conditioned response is a response to the sign-relationship S_2---S_1. And this means that it goes off only so long as that sign-relationship is "believed" in by the organism.

TI. 2.7 $\quad r(x, \varphi, p) > 0 \supset pR^e x$

TI. 2.7 reads: if x tends to respond by φ to p then p is sensorily received by x.

13. This postulate is important. One might think that a given state-of-affairs must exist in order to be a stimulus. But this contention is false. For p might be an illusory state-of-affairs. Nevertheless, the illusory state-of-affairs is sensorily received, that is, it is taken to be true by sense-inspection. Hence, the postulate TI. 2.7.

TI. 2.8 $\quad t^A(x, p, q) > 0 \supset d(x, p, q) > 0$

TI. 2.8 reads: if x has been taught that p is a means to q then he demands p as a means to q. In other words if an organism has no demand he cannot be taught.

14. TI. 2.8 employs one of the meanings of docility discussed in the last chapter. In the discussion of the primitive ideas of "taughtness" and "teachability" it was pointed out that "teachability" could be defined in terms of "taughtness" if an adequate analysis of what "capability" means was at hand. However, also, it was pointed out that

such an analysis was not immediately forthcoming. Nevertheless an "operational definition" of teachability can be given in TI. For it is provable that

$$t^A(x, p, q) > 0 \supset [t(x, p, q) > 0 \equiv d(x, p, q) > 0];$$

that is, if x has been taught that p leads to q then x is teachable with respect to p as a means to q is equivalent to x demands p as a means to q. This theorem is important because Tolman speaks of teachability as that which "objectively defines demand". "Teachability" is the "mark" of demand in virtue of the fact that x cannot be taught that p is a means to q unless x is teachable with respect to p as a means and also demands p as a means to q. Hence, we may say that if x has been taught that p is a means to q then he is teachable with respect to p as a means to q. And since teachability is the "mark" of demand, then it follows that taughtness is a "mark" of demand. The point is this: one must get it out of his thought that teachability "causes" or "entails" demand. But since Tolman's test for demand is via taughtness, teachability may be regarded as the mark of demand in the sense that if x is not teachable one need not go on to test for demand, because it follows that x cannot be taught. Hence, one may regard teachability as objectively defining demand; for unless x is teachable, demand cannot be tested for.

15. Not only is "teachability" (and hence "taughtness") a mark of demand (purpose), it is also a mark of expecta-

tion.[5] For it follows directly from TI. 2.8 that

$$t^A(x, p, q) > 0 \supset e(x, p \supset q) > 0$$

16. The evidence that Tolman himself believes taughtness implies expectation is deducible from the fact that taughtness implies demand is given in the following quotation.[6]

> Finally, it is to be noted that purposiveness and cognitiveness seem to go together, so that if we conceive behavior as purposive we *pari passu* conceive it also as cognitive.

TI. 2.9 $\quad t^A(x, p, q) > 0 \supset t(x, p, q) > 0$

TI. 2.9 reads: If x is taught that p leads to q, then x is teachable with respect to p as means to q. This postulate makes explicit the relationship between "taughtness" and "teachableness" discussed in paragraph 56. in chapter II.

TI. 2.10 $\quad (\varphi x)\ddot{U}x \supset \sim r(x, \varphi, p) > 0$

TI. 2.10 reads: If φx is an ultimate goal-object of x, then it is false that x tends to respond by φ to p. This postulate merely affirms that no ultimate goal-object is a response to anything.

TI. 2.11 $\quad d(x, p, q) > 0 \supset [t^A(x, p, q) > 0 \vee F(x, p, q)]$

TI. 2.11 reads: If x demands p as a means to q, then x

5. *Purposive Behaviorism*, pp. 14-17. See "Glossary" for equivalence of "demand" and "purpose".
6. *Purposive Behaviorism*, p. 13.

has been taught that p leads to q or he is fixated on p as a means to q. This postulate is consonant with the remarks concerning "fixation" in the discussions of the definition of "fixation" and the concept of teachability in preceding chapter.

TI. 2.12 $pR^e x \supset x^C p$

TI. 2.12 reads: if p is sensorily received by x, then x has confirmed p.

17. This postulate reaffirms a suggestion made in the discussion of "confirmation" in the last chapter (Cf. paragraph 44). Since '$x^C p$' contains as part of its meaning 'is taken to be true', it does not follow that

$$x^C p \supset p;$$

that is, that if x has confirmed p, then p is true. For not everything which is taken to be true is true. The converse also holds. Hence, it is not the case that

$$pR^e x \supset p;$$

that is, that if p is sensorily received by x then p is true. Again, the converse is also true for sensory reception.

18. There are borderline cases which are hard to deal with in the case of TI. 2.12. This is because the difference between '$pR^e x$' and '$x^C p$' is a matter of degree. Nevertheless, we are able in a more or less vague way to say something about the extremes of that continuum. TI. 2.12 is really a statement which holds when we are willing to say definitely

p is sensorily received by x or x has confirmed p.

A Replacement Rule in TI

19. The system of *Principia Mathematica* is called a truth-functional system. A system is a truth-functional system when every propositional function in it is a truth-function of its constituent propositions. For example, the truth of 'p ≡ q' is a function of the truth of its constituents 'p' and 'q'. 'p ≡ q' is true when both 'p' and 'q' are true or when they are both false; in all other cases 'p ≡ q' is false. The present system (TI and TII) is not a truth functional system. This can be shown in the following way. If the present system is truth functional then every proposition in it should be a truth function of its components. Hence, if 'p' and 'q' are the only propositional components of 'e(x, p) > 0' and 'e(x, q) > 0' and if they are true, then 'e(x, p) > 0' and 'e(x, q) > 0' should be both true or both false. Indeed, any relation obtaining between 'p' and 'q' should also obtain between 'e(x, p) > 0' and 'e(x, q) > 0'. For example,

(1) $p \equiv q \supset [e(x, p) > 0 \equiv e(x, q) > 0]$,

which reads: if the propositions 'p' and 'q' are equivalent then 'e(x, p) > 0' is equivalent to 'e(x, q) > 0'. (1) can be shown to be false. For example, substitute the propositions 'Lambert has decided to buy a car' for 'p' and 'Eisenhower will run for reelection' for 'q'. Both 'p' and 'q' are true. Hence 'p ≡ q' is true. But 'e(x, p) > 0' is not equivalent to '$e(x, q) > 0$' where 'x' is 'Eisenhower' because Eisenhower does not know Lambert. Hence 'Eisenhower expects he will run for reelection' is undoubtedly true but 'Eisenhower expects that Lambert will buy a car' is surely false.

20. The present system is not a modal logic either. A logic is modal if it contains an operator '$\Diamond$' which means "is possible", that is, "is not logically inconsistent". Now, to say that p and q are mutually deducible, that is

(2) $p = q$,

is to say that it is not possible that p is true and q is false and also it is not possible that q is true and p is false, that is,

(3) $\sim\Diamond(p \cdot \sim q) \cdot \sim\Diamond(q \cdot \sim p)$.

It has been shown that the truth functional connective '$\equiv$', when in main position in a formula in Principia, can be "strengthened" to '$=$', that is, "equivalence" can be "strengthened" to "mutual deducibility". For example, the Principia formula

(4) $p \equiv p$

can be "strengthened" to

(5) $p = p$.

This can be done because all the formulae of Principia are analytic and hence are not logically inconsistent. There are some cases where $p = q$ does not materially imply that $e(x, p) > 0 = e(x, q) > 0$, that is,

(6) $p = q \supset [e(x, p) > 0 = e(x, q) > 0]$

is, in some instances, false. For example, let 'p' be '$(p \cdot q) \supset r$' and 'q' be '$p \supset (q \supset r)$'. In Principia the following law holds:

$$(7) \quad [(p \cdot q) \supset r] \equiv [p \supset (q \supset r)]$$

According to the principle of strengthening (7) may be rewritten as

$$(8) \quad [(p \cdot q) \supset r] = [(p \supset (q \supset r)].$$

Thus, under the above substitutions, the antecedent of (6) is true. But there is good reason to believe that the present consequence is sometimes false. For though it is true that

$$(9) \quad e(x, p \supset (q \supset r)) > 0 \rightarrow e(x, (p \cdot q) \supset r) > 0$$

the converse of (9), that is,

$$(10) \quad e(x, (p \cdot q) \supset r) > 0 \rightarrow e(x, p \supset (q \supset r)) > 0$$

does not always hold.[7] For example, let 'q' be a stimulus situation, say, 'There is a white card before x', 'p' be a condition consequent on the occurrence of 'q', say, 'x jumps to the white card', and let 'r' be a situation postdating the occurrence of 'p', say, 'x gets to food'. Then (10) reads: "x expects that if x jumps to white card then if there is a white card before him, he gets to food" is deducible from "x expects that if both he jumps to the white card and there is a white card before him, then he gets to food". The trouble here is that the "if, then"

---

7. '$p \rightarrow q$' is usually defined as '$\sim\Diamond(p \cdot \sim q)$', hence '$p \rightarrow q$' means 'p is deducible from q'; or more commonly 'p strictly implies q'.

relation in the consequent of (10) does not seem to capture the feature of <u>consecutivity</u> present in Tolman's analysis of expectation. This is more dramatically seen if in the above application of (10) one rephrases that application using "is a means to" rather than "if, then". Of course, this leads to the conclusion that the material (or truth functional) '$\supset$', that is, "if, then", is too weak a relation to express the meaning of "is a means to" in the notion of means-end expectation. This point will be considered again in the last chapter of this essay.

21. Since (10) can be false, then it is possible for the present consequent of (6) to be false. Hence we conclude that the present system is not a modal system.

22. In <u>Principia</u>, if a law of the form

$$p \equiv q$$

holds, then if 'p' occurs in any other formula, 'q' may replace 'p' in that formula. This is the <u>rule of biconditional replacement</u>. But the results in the preceding paragraphs will not allow us to employ this rule in TI without restriction <u>within</u> the contexts of expectation, demand, confirmation, etc. Let us call such contexts <u>propositional attitude contexts</u>. Our revised rule then reads: if a <u>Principia</u> formula of the form

$$p \equiv q$$

also holds when governed by a given propositional attitude, for example,

$$e(x, p) > 0 \equiv e(x, q) > 0,$$

then if 'p' occurs in any formula governed by that propositional attitude, 'q' may replace 'p' and vice versa. For example, the following proposition is true in TI.

$$(11) \quad d(x, p, q) > 0 \supset [d(x, p) > 0 \,.\, d(x, q) > 0]$$

that is, if x demands p as a means to q, then x demands p and x demands q. The Principia law

$$p \equiv \sim(\sim p)$$

also holds within the propositional attitude context of demand. That is,

$$(12) \quad d(x, p) > 0 \equiv d(x, \sim(\sim p)) > 0$$

holds in TI. By our rule of restricted biconditional replacement, we get from (11), in accordance with (12),

$$(13) \quad d(x, \sim(\sim p), q) > 0 \supset [d(x, \sim(\sim p)) > 0 \,.\, d(x, q) > 0].$$

This rule is used tacitly in TI.

Proof Procedures in TI

23. Two proof procedures are employed in TI. The first is a rather full proof and is quite similar to the proof technique in Principia. The second is an abbreviated procedure which, in effect, makes every line in a proof a theorem. Let me illustrate.

TI. 3.0.00 $\quad d(x, p, q) > 0 \supset d(x, q) > 0$

$d(x, p, q) > 0 \supset [d(x, q) > 0 \,.\, e(x, p \supset q) > 0]$: (Id$[d(\underset{p}{\underline{x,p,q}}) > 0]$, and TI. 1.0) (1)

$[d(x, p, q) > 0 \supset d(x, q) > 0] \,.\, [d(x, p, q) > 0 \supset e(x, p \supset q) > 0]$

(1) and *4.76 $[d(\underset{p}{\underline{x, p, q}}) > 0,\ d(\underset{q}{\underline{x, q}}) > 0,\ e(\underset{r}{\underline{x, p \supset q}}) > 0]$ (2

$$d(x, p, q) > 0 \supset d(x, q) > 0 \quad (2) \text{ and Simp } [d(x, p, q) > 0 \underset{p}{\supset} d(x, q) > 0, d(x, p, q) > 0 \underset{q}{\supset e} (x, p \supset q) > 0]$$

The first line in the above example is not a line in the proof; it is the theorem to be proven.

24. If the proof is rather difficult, then the above is the form the proof will take. In this example the numbers and abbreviations to the right hand of any line indicate the laws which sanction that line <u>or</u> the number of a line in the proof; I mention no methods in any proof in TI. The numbers beginning with the asterisk and the abbreviations, for example, "Simp", are the expressions used in <u>Principia</u> to designate certain laws in that system. Any other number designates a postulate, theorem, or definition in TI --or a line in the proof. The definitions are represented by TI. 1.N while the postulates, are denoted by the prefix TI. 2.N. Thus the first line in the above proof cites a law from <u>Principia</u> and a definition in TI. The number of any line in a proof in written immediately after the justification for that line.

25. '[]' after a sanctioning law indicate a substitution into the law immediately preceding the brackets. Hence, in the first line in the above proof they indicate a substitution instance in the <u>Principia</u> law "Id." In general, however, the substitutions will only be noted in the more difficult cases.

26. In some cases the last line, but not the laws which sanction it, may be replaced by "prop." in accordance with the conventions of Principia. For example, the last line may have been written as; prop. (2) and Simp. [same substitutions], which indicates that the result of (2) and Simp is the desired theorem.

27. However, many of the proofs in TI are the result of only two or three steps since, in general, what normally would be taken as a line in a proof is proven as a theorem. Thus the abbreviated proof procedure in TI is illustrated by the following example

$d(x, p, q) > 0 \supset [d(x, q) > 0 \,.\, e(x, p \supset q) > 0]$ TI. 3.0.0

$d(x, p, q) > 0 \supset d(x, q) > 0$ TI. 3.0.0 and Simp: TI. 3.0.00

28. Notice (1) that TI. 3.0.0, since it precedes TI. 3.0.00 in logical order, constitutes a line in the proof of TI. 3.0.00; (2) that the numbers of the laws sanctioning TI. 3.0.00 represent a "telescoped" justification of TI. 3.0.00 --for example, the extended proof of TI. 3.0.00 in paragraph 23. contains reference to ✱4.76 in its second line. Telescoping, is done in cases which seem to follow obviously; (3) that the number of a line is now the number of a theorem; (4) that the sanction of the line is given with the theorem to be proven; and (5) that there is no mention here (and in most cases there will not be) of the requisite substitutions. (Cf. paragraph 25)

Numeration of the Theorems of TI

29. The theorems of TI fall into two classes; (1) those which follow from the definitions alone and (2) those which follow from the postulates alone or in conjunction with the definitions. Those which fall in class (1) are designated by the prefix

TI. 3.M.N;

those falling in class (2) are designated by the prefix

TI. 4.M.N.

30. All theorems in TI are subdivided into theorems and corollaries of theorems. For example, consider the following set of theorem-numbers:

(A) TI. 4.0.0
(B) TI. 4.0.00
(C) TI. 4.0.01
(D) TI. 4.1.0

In this set, the number replacing 'M' in the prefix

TI. 4.M.N

identifies a given subset of theorems. If the number which replaces 'N' in the prefix in

TI. M.N

is '0' then the prefix designates the first theorem in the given subset. If the number which replaces 'N' is '00', then the prefix designates the first corollary of the first theorem in the given subset. For example, in the above set

TI. 4.0.0

identifies the first subset of theorems which depend upon the postulates in TII and also the first theorem in that subset. The number

TI. 4.0.00

designates the first corollary of the first theorem in the first subset of theorems which depend upon the postulates in TI. Finally,

TI. 4.1.0

designates the first theorem in the second subset of theorems which depend upon the postulates of TI.

31. There can be corollaries of corollaries. For example,

TI. 4.0.010

designates the first corollary of the second corollary of the first theorem in the first subset of theorems which depend upon the postulates in TI. All of these remarks also hold true for those theorems following from the definitions.

The Theorems of TI

Group A: Consequences of the Definitions

32. The theorems in Group A are deduced solely from the definitions without any appeal to the postulates. They are arranged in subgroups, the first subgroup depending exclusively on the first definition, the second subgroup on the second definition, and so on (with slight modifications which will be explained when they arise).

33. The following theorems depend on definition TI. 1.0.

TI. 3.0.0: $d(x, p, q) > 0 \supset [d(x, q) > 0 \;.\; e(x, p \supset q) > 0]$

Id and TI. 1.0

TI. 3.0.0 reads: if x demands p as a means to q, then x demands q and expects that if p then q.

TI. 3.0.00: $SA \supset d(x, q) > 0$ TI. 3.0.0 and Simp[8]

TI. 3.0.00 reads: if x demands p as a means to q, then x demands q.

TI. 3.0.01: $SA \supset e(x, p \supset q) > 0$ TI. 3.0.0 and Simp

TI. 3.0.01 reads: if x demands p as a means to q, then x expects that if p then q.

TI. 3.0.1: $[d(x, q) > 0 \,.\, e(x, p \supset q) > 0] \supset d(x, p, q) > 0$
Id and TI. 1.0

TI. 3.0.1 reads: if x demands q and also expects that if p then q, then x demands p as a means to q.

34. The following theorems are derived from TI. 1.1 either alone or in conjunction with TI. 1.0

TI. 3.1.0: $pMx \supset (\exists q)(d(x, p, q) > 0)$ TI. 1.1, Id and Simp

TI. 3.1.0 reads: if p is a means-object of x, then there is a q such that x demands p as a means to q.

---

8. In theorem TI. 3.0.00, there occurs an abbreviation which will be regularly used throughout the remainder of this work. "SA" ("same antecedent") stands in the position of antecedent in any theorem (or line of a proof) to indicate that the antecedent of the theorem (or line) is the same as that in the immediately preceding theorem (or line).

TI. 3.1.00: $pMx \supset (\exists q)(d(x, q) > 0)$

Proof: $(q)[d(x, p, q) > 0 \supset d(x, q) > 0]$	TI. 3.0.00 and *9.13	(1)
$(\exists q)(d(x, p, q) > 0) \supset (\exists q)(d(x, q) > 0)$	(1) and *10.28	(2)
$pMx \supset (\exists q)(d(x, q) > 0)$	(2) and TI. 3.1.0	

TI. 3.1.00 reads: if p is a means object of x, then x demands some state-of-affairs.

TI. 3.1.000: $(q)(d(x, q) = 0) \supset \sim(pMx)$ — TI. 3.1.00, Transp, *10.252 and RN[9]

TI. 3.1.000 reads: for every q if x does not demand q, then it is false that p is a means-object of x.

TI. 3.1.01: $pMx \supset (\exists q)(e(x, p \supset q) > 0)$ — Proof is similar to proof of TI. 3.1.00.

TI. 3.1.01 reads: if p is a means-object of x, then there is a q such that x expects that if p then q.

TI. 3.1.1: $(\exists q)(d(x, p, q) > 0) \supset pMx$ — TI. 1.1, Id and Simp.

TI. 3.1.1 reads: if there is a q such that x demands p as a means to q, then p is a means-object of x.

TI. 3.1.10: $d(x, p, q) > 0 \supset pMx$ — TI. 3.1.1, *10.23 and *10.1.

TI. 3.1.10 reads: if x demands p as a means to q, then p is a means-object of x.

TI. 3.1.100: $[d(x, q) > 0 \,.\, e(x, p \supset q) > 0] \supset pMx$ — TI. 3.1.10 and TI. 3.0.1

---

9. All references to the laws of <u>Real Numbers</u> will be referred to as simply "RN".

TI. 3.1.100 reads: if x demands q and expects that p is a means to q, then p is a means-object of x.

35. The following theorems are derived from TI. 1.2 either alone or in conjunction with TI. 1.0 or TI. 1.1 (or both). Hereafter, except in the most difficult cases, English language translations of the theorems will be omitted.

TI. 3.2.0: $pUx \supset [d(x, p) > 0 \,.\, \sim(\exists q)(d(x, p, q) > 0]$ TI. 1.2, Id and Simp

TI. 3.2.00: $pUx \supset d(x, p) > 0$ TI. 3.2.0 and Simp

TI. 3.2.000: $pUx \supset (\exists q)(d(x, q) > 0)$ TI. 3.3.00 and *10.24

TI. 3.2.01: $pUx \supset \sim(\exists q)(d(x, p, q) > 0)$ TI. 3.2.0 and Simp

TI. 3.2.010: $pUx \supset \sim d(x, p, q) > 0$ TI. 3.2.01, *10.01 and *10.1

TI. 3.2.0100: $pUx \supset [d(x, q) > 0 \supset \sim e(x, p \supset q) > 0]$ TI., 3.2.010, TI. 3.0.1 and Transp.

TI. 3.2.01000: $[e(x, p \supset q) > 0 \,.\, pUx] \supset d(x, q) = 0$ TI. 3.2.01000, Transp, Exp. and RM

TI. 3.2.011: $pUx \supset \sim(pMx)$ TI. 3.2.01, TI. 3.1.0 and Transp.

TI. 3.2.0110: $pMx \supset \sim(pUx)$ TI. 3.2.011 and Transp.

TI. 3.2.012: $pUx \supset (\exists q)(\sim e(x, p \supset q) > 0)$

Proof: $\sim(\exists q)(d(x, p, q) > 0) \supset [d(x, q) > 0 \supset \sim e(x, p \supset q) > 0]$ TI. 3.0.1, *10.24 and Transp (1)

$pUx \supset [d(x, q) > 0 \supset \sim e(x, p \supset q) > 0]$ (1) and TI. 3.2.01 (2)

$(\exists q)(d(x, q) > 0) \supset [pUx \supset (\exists q)(\sim e(x, p \supset q) > 0)]$ (2), *9.13, *10.5 and *4.87 (3)

prop. (3), TI. 3.2.000 and Taut.

TI. 3.2.1: $[d(x, p) > 0 \,.\, \sim(\exists q)(d(x, p, q) > 0] \supset pUx$ TI. 1.2, Id and Simp

TI. 3.2.10: $d(x, p) > 0 \supset [\sim(pMx) \supset pUx]$

Proof: $[\sim(pMx) \,.\, d(x, p) > 0] \supset pUx$ TI. 3.2.1, TI. 3.1.1 and Trans (1)

prop. (1), Comm and Exp.

TI. 3.2.100: $d(x, p) > 0 \supset [pMx \lor pUx]$ TI. 3.2.10 and *4.6

TI. 3.2.101: $d(x, p) > 0 \supset [pUx \equiv \sim(pMx)]$ TI. 3.2.10 and TI. 3.2.011

36. The following theorems are derived from TI. 1.3 in conjunction with TI. 1.0, TI. 1.1, and/or TI. 1.2.

TI. 3.3.0: $F(x, p, q) \supset [d(x, p, q) > 0 \,.\, \sim(\exists r)(xC(r=\sim p) \,.$

$d(x, r, q) > 0)]$ TI. 1.3, Id and Simp

TI. 3.3.00: $F(x, p, q) \supset d(x, p, q) > 0$ TI. 3.3.0 and Simp

TI. 3.3.000: $F(x, p, q) \supset [d(x, q) > 0 \,.\, e(x, p \supset q) > 0]$ TI. 3.3.00 and TI. 3.0.0

TI. 3.3.0000: $F(x, p, q) \supset d(x, q) > 0$ TI. 3.3.000 and Simp

TI. 3.3.00000: $F(x, p, q) \supset [qUx \lor qMx]$ TI. 3.3.0000 and TI. 3.2.100

TI. 3.3.00001: $F(x, p, q) \supset [qUx \equiv \sim(qMx)]$ TI. 3.3.0000 and TI. 3.2.101

TI. 3.3.001: $F(x, p, q) \supset pMx$ TI. 3.3.00 and TI. 3.1.10

TI. 3.3.0010: $F(x, p, q) \supset \sim(pUx)$ TI. 3.3.001 and TI. 3.2.0110

TI. 3.3.01: $F(x, p, q) \supset [\sim(\exists r)(xC(r=\sim p) \,.\, d(x, r, q) > 0)]$ TI. 3.3.0 and Simp

TI. 3.3.010: $SA \supset \sim[xC(r=\sim p) \,.\, d(x, r, q) > 0]$ TI. 3.3.01 and *10.1

TI. 3.3.0100: $[xC(r=\sim p) \,.\, d(x, r, q) > 0] \supset \sim F(x, p, q)$ TI. 3.3.010 and Transp.

TI. 3.3.1: $[d(x, p, q) > 0 \,.\, \sim(\exists r)(xC(r = \sim p) \,.\, d(x, r, q) > 0)]$
$\supset F(x, p, q)$ TI. 1.3, Id and Simp

TI. 3.3.10: $\sim F(x, p, q) \supset [d(x, p, q) > 0 \supset (\exists r)(xC(r = \sim p) \,.$
$d(x, r, q) > 0)]$ TI. 3.3.1 and Transp

TI. 3.3.100: $[\sim F(x, p, q) \,.\, d(x, p, q) > 0] \supset (\exists r)(xC(r = \sim p))$ TI. 3.3.10, Imp, *10.28 and Simp

37. The following theorems are derived from TI. 1.4 in conjunction with TI. 1.0, TI. 1.1, TI. 1.2 and/or TI. 1.3.

TI. 3.4.0 $pSx \supset (\exists \varphi)(r(x, \varphi, p) > 0)$ TI. 1.4, Id and Simp

TI. 3.4.1 $(\exists \varphi)(r(x, \varphi, p) > 0) \supset pSx$ TI. 1.4, Id and Simp

TI. 3.4.10 $r(x, \varphi, p) > 0 \supset pSx$ TI. 3.4.1, *10.23 and *10.1

<u>Group B: Consequences of the Postulates</u>

38. The theorems in Group B are deduced from the postulates in conjunction with other postulates, definitions or theorems resulting both from other postulates and the definitions. The theorems in this section are not subdivided into groups depending upon any given postulate.

TI. 4.0.0: $[d(x, \psi x) > 0 \,.\, pR^e x \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset$
$(\varphi x \supset \psi x)) > 0] \supset r(x, \varphi, p) > 0$

Proof: $[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset (\varphi x \supset \psi x)) > 0] \supset e(x, p \supset$
$(\varphi x \supset \psi x)) > 0$ TI. 2.0 $\left[\dfrac{\varphi x \supset \psi x}{r}\right]$ (1)

$[d(x, \psi x) > 0 \,.\, e(x, p \supset (\varphi x \supset \psi x)) > 0 \,.\, pR^e x] \supset r(x, \varphi, p) > 0$
TI. 2.6 $\left[\dfrac{\psi x}{q}\right]$ (2)

$xCp \supset ([d(x, \psi x) > 0 \,.\, pR^e x \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset (\varphi x \supset \psi x)) > 0]$
$\supset r(x, \varphi, p) > 0)$ (1), (2) and Exp. (3)

prop: (3), TI. 2.12, Imp and Taut.

39. This theorem allows us to assert that x has a tendency to respond by φ to a state-of-affairs p when x demands ψx, has sensorily received p and when x expects that p leads to φx via the intermediate event q which intervenes between p and φx.

TI. 4.0.00: $[(\psi x)Ux \,.\, pR^{e}x \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset (\varphi x \supset \psi x)) > 0] \supset r(x, \varphi, p) > 0$ TI. 4.0.0 and TI. 3.2.00 $[\frac{\psi x}{p}]$

TI. 4.0.000: $[(\psi x)Ux \,.\, pR^{e}x \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset(\varphi x \supset \psi x)) > 0] \supset \sim(\varphi x)Ux$ TI. 4.0.00, TI. 2.10 and Transp

40. This theorem shows that, under the same conditions prevailing in TI. 4.0.0 -except that 'x's demand for ψx' is replaced by 'ψx is an ultimate goal-object of x'- φx is not an ultimate goal-object of x. This can be seen more clearly if one compares φx which we may allow to mean 'x eats bran mash cubes' with ψx which we may allow to mean 'x is (hunger) satiated'.

TI. 4.0.01: $[d(x, \psi x) > 0 \,.\, pR^{e}x \,.\, t^{A}(x, p, q) > 0 \,.\, t^{A}(x, q, (\varphi x \supset \psi x)) > 0] \supset r(x, \varphi, p) > 0$ TI. 4.0.0 and TI.2.8

TI. 4.0.02: $[d(x, \psi x) > 0 \,.\, pR^{e}x \,.\, F(x, p, q) \,.\, F(x, q, \varphi x \supset \psi x)] \supset r(x, \varphi, p) > 0$ TI. 4.0.0 and TI. 3.3.00

TI. 4.0.020: $[(d(x, \psi x) > 0 \,.\, pR^{e}x) \,.\, (t^{A}(x, p, q) > 0 \,.\, t^{A}(x, q, (\varphi x \supset \psi x)) > 0) \vee (F(x, p, q) \,.\, F(x, q, \varphi x \supset \psi x))] \supset r(x, \varphi, p) > 0$ TI. 4.0.02 and TI. 4.0.01

41. This theorem shows that when an organism demands that ψx, -say, 'x be picked up' --, has sensorily received that p,

--say, 'There is a white card on the left'--, and, finally, when he exhibits a chain of taughtness or a chain of fixation both of which involve p, q, -say, 'x moves, on the stand, toward the white card'-, φx, -say, 'x jumps to the white card--, and ψx, x has a tendency to respond to the fact that 'there is a white card on the left' by 'jumping to the white card'.

TI. 4.1.0: $[xCp \;.\; d(x, p, q) > 0 \;.\; d(x, p, r) > 0] \supset [d(x, r) > 0 \;.\; e(x, p \supset r) > 0]$

Proof: $[d(x, q) > 0 \;.\; d(x, r) > 0] \supset d(x, r) > 0$ Simp (1)

$[(xCp \;.\; e(x, p \supset q) > 0 \;.\; e(x, q \supset r) > 0) \;.\; d(x, q) > 0 \;.$ $d(x, r) > 0] \supset [d(x, r) > 0 \;.\; e(x, p \supset r) > 0]$ (1), TI. 2.0 and *3.47 (2)

prop. (2), *4.32, Comm and TI. 1.0

TI. 4.1.00: SA $\supset d(x, r) > 0$ TI. 4.1.0 and Simp

TI. 4.1.01: SA $\supset e(x, p \supset r) > 0$ TI. 4.1.0 and Simp

TI. 4.1.02: SA $\supset d(x, p, r) > 0$ TI. 4.1.0 and TI. 3.0.1

42. Theorem TI. 4.1.02 expresses the restricted transitivity of means-end demand. That is, it expresses a condition under which a demand chain is "set off" --to use Tolman's suggestive terminology.

TI. 4.1.020: SA $\supset$ pMx TI. 4.1.02 and TI. 3.1.10

TI. 4.1.0200: SA $\supset$ [pMx . qMx] TI. 4.1.020, TI. 3.1.10 and *3.47.

43. TI. 4.1.0200 shows that p and q, because they are demanded as means to other goal objects in the demand chain,

are, therefore, proximate or means-objects. In this connection it is interesting to note that a similar law for <u>expectation</u> chains cannot be derived. This is important. For where 'p is a sign object of x' means

$$[pR^e x \,.\, (\exists q)(e(x, p \supset q) > 0)]$$

the failure to deduce such a law conforms with Tolman's requirement that a sign object becomes a means-object when and only when it is demanded.[10]

TI. 4.2.0: $[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, d(x, r) > 0] \supset d(x, p) > 0$ TI. 2.0 and TI. 2.1

44. This theorem shows that an approach situation can be extended along an expectation chain.

TI. 4.2.00: $[pR^e x \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, d(x, r) \;\; 0] \supset d(x, p) \;\; 0$ TI. 4.2.0 and TI. 2.12

TI. 4.2.000: $[r(x, \varphi, p) > 0 \,.\, e(x, p \supset \varphi x) > 0 \,.\, e(x, \varphi x \supset r) > 0 \,.\, d(x, r) > 0] \supset d(x, p) > 0$ TI. 4.2.00 $\left[\frac{\varphi x}{q}\right]$ and TI. 2.7

45. This theorem reads: if x has the tendency to respond by φ to p and expects that p is a means to φx and expects that φx is a means to r and demands r, then x demands p. This theorem shows a circumstance under which one can infer a demand from a certain combination of response and expectation conditions.

TI. 4.2.1: $[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, d(x, {\sim}r) > 0] \supset d(x, {\sim}p) > 0$ TI. 2.0 and TI. 2.2

---

10. Cf. <u>Purposive Behaviorism</u>, Chapter X

46. This theorem shows that an avoidance situation can be extended along an expectation chain.

TI. 4.2.10 $[pR^ex \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, d(x, \sim r) > 0] \supset d(x, \sim p) > 0$ TI. 4.2.1 and TI. 2.12

TI. 4.2.100 $[r(x, \varphi, p) > 0 \,.\, e(xp \supset \varphi x) > 0 \,.\, e(x, \varphi x \supset r) > 0 \,.\, d(x, \sim r) > 0] \supset d(x, \sim p) > 0$ TI. 4.2.10 and TI. 2.7

TI. 4.2.2: $[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, d(x, r) = d(x, \sim r) = M > 0] \supset [d(x, p) > 0 \,.\, d(x, \sim p) > 0]$

Proof: $[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, d(x, r) > 0 \,.\, d(x, \sim r) > 0] \supset [d(x, p) > 0 \,.\, d(x, \sim p) > 0]$ TI. 4.2.0 and TI. 4.2.1 (1)

$[d(x, r) = M \,.\, M > 0 \,.\, d(x, \sim r) = M.M > 0 \,.\, xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0] \supset [d(x, p) > 0 \,.\, d(x, \sim p) > 0]$ (1) and *13.13 (2)

$[d(x, r) = d(x, \sim r) = M.M > 0 \,.\, xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0] \supset [d(x, p) > 0 \,.\, d(x, \sim p) > 0]$ (2) and *13.03

$[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, d(x, r) = d(x, \sim r) = M > 0] \supset [d(x, p) > 0 \,.\, d(x, \sim p) > 0]$ (3), RM and Comm.

47. TI. 4.2.2 is a very interesting theorem. Notice that the expression "$d(x, r) = d(x, \sim r) = M > 0$" depicts the typical conflict-in-demands situation. Hence, TI. 4.2.2, in effect, says that if an organism has conflicting demands with respect to the last member of an expectation chain, then he demands both the occurrence and the non-occurrence of the initial goal-object. Notice that one cannot infer that the organism is in conflict with respect to the initial goal-objects. This comports with ordinary experience; for sometimes one is and sometimes one is not in conflict with respect to the initial goal-object under the above conditions.

TI. 4.2.3 $[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0] \supset e(x, r) > 0$ TI. 2.0 and TI. 2.4

TI. 4.2.30 $[xCp \,.\, e(x, p \supset q) > 0 \,.\, xC(q \supset \sim r)] \supset \sim e(x, q \supset r) > 0$

Proof: $[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, xC(q \supset \sim r)] \supset \sim e(x, q) > 0$ TI. 4.2.3 and TI. 2.5 $[\frac{r}{p}]$ (1)

$[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, xC(q \supset \sim r)] \supset [e(x, q) > 0 \,.\, \sim e(x, q) > 0]$ (1), TI. 2.4 *3.47 (2)

$[xCp \,.\, e(x, p \supset q) > 0 \,.\, e(x, q \supset r) > 0] \supset \sim xC(q \supset \sim r)$ (2), Transp and Id (3)

Prop. (3), Imp, Transp and Exp

48. This theorem, TI. 4.2.30, affirms that when x has confirmed that q implies ~r and, also, has confirmed p and expects that p leads to q, then the chain of expectations is "broken" with respect to the goal-object r -that is, then x does not expect that q leads to r.

TI. 4.3.0 $d(x, p, q) > 0 \supset d(x, p) > 0$ TI. 2.1 and TI. 3.01

TI. 4.3.00 $d(x, p, q) > 0 \supset [d(x, p) > 0 \,.\, d(x, q) > 0]$ TI. 4.3.0 and TI. 3.0.00

49. TI. 4.3.00 asserts that if x demands p as a means to q then, he demands p and he demands q.

TI. 4.3.000 $[xCp \,.\, d(x, p, q) > 0 \,.\, d(x, q, r) > 0] \supset [d(x, p) > 0 \,.\, d(x, q) > 0 \,.\, d(x, r) > 0]$ TI. 4.3.00 $[\frac{r}{q}]$, TI.4.1.00, *3.47

50. The above theorem claims that all of the goal-objects in a demand chain are individually demanded.

TI. 4.3.001: $t^A(x, p, q) > 0 \supset [d(x, p) > 0 \,.\, d(x, q) > 0]$ TI. 4.3.00 and TI. 2.8

TI. 4.3.002: $F(x, p, q) \supset [d(x, p) > 0 \,.\, d(x, q) > 0]$
TI. 4.3.00 and TI. 3.3.00

TI. 4.3.003: $[d(x, p) > 0 \supset \sim d(x, \sim p) > 0] \supset \sim d(x, p, \sim p) > 0$
TI. 4.3.00 $\left[\frac{\sim p}{q}\right]$ and Transp

51. TI. 4.3.003 affirms that if x's demand for p implies that he does not demand not p, then it is false that he demands p as a means to not p.

TI. 4.3.01: $pMx \supset d(x, p) > 0$ TI. 4.3.0 and TI. 3.1.0

TI. 4.3.010: $[pMx \vee p\bar{U}x] \supset d(x, p) > 0$ TI. 4.3.01 and TI. 3.2.00

TI. 4.3.0100: $[p\bar{U}x \equiv \sim(pMx)] \supset d(x, p) > 0$ TI. 4.3.010 and TI. 3.2.011

TI. 4.3.01000: $d(x, p) > 0 \equiv [p\bar{U}x \equiv \sim(pMx)]$ TI. 4.3.01000 and TI. 3.2.101

52. This theorem, TI. 4.3.01000, affirms that x's having a demand for p is equivalent to p is an ultimate goal-object of x if and only if p is not a proximate goal-object of x.

TI. 4.3.011: $[qMx \vee q\bar{U}x \,.\, e(x, p \supset q) > 0] \supset d(x, p) > 0$
TI. 4.3.010 and TI. 2.1

TI. 4.3.012: $[qMx \vee q\bar{U}x \,.\, e(x, p \supset q) > 0] \supset pMx$ TI. 4.3.010, TI. 3.0.1 and TI. 3.1.10

53. TI. 4.3.012 asserts when q is a proximate goal-object or an ultimate goal-object and when x expects that p leads to q, then p is a proximate goal-object. However, this is not the case when we are dealing with avoidance situations. That is, we cannot prove that

(A) $[(\sim q)Mx \vee (\sim q)Ux \,.\, e(x, p \supset q) > 0] \supset (\sim p)Mx$,

though it is easy to prove that

(B) $[(\sim q)Mx \vee (\sim q)Ux \,.\, e(x, p \supset q) > 0] \supset d(x, \sim p) > 0$.

And indeed we want this result in our system. The point is, that under the conditions specified in (A) and (B), ~p is demanded by x, not with the <u>expectation</u> that it will lead to ~q (an expectation required if (A) is to be true) but, so to speak, with the <u>hope</u> it will lead to ~q. (To <u>expect</u> ~p to lead to ~q under the indicated conditions is, so to speak, to commit the logical fallacy of denying the antecedent).

TI. 4.4.0: $[e(x, p \supset \sim q) > 0 \,.\, e(x, q \supset r) > 0 \,.\, d(x, r) > 0] \supset d(x, \sim p) > 0$. TI. 2.1 $[\frac{q}{p}, \frac{r}{q}]$ and TI. 2.2 $[\frac{\sim q}{q}]$

TI. 4.4.1: $[e(x, p \supset q) \;\; 0 \,.\, e(x, \sim q \supset r) \;\; 0 \,.\, d(x, \sim r) \;\; 0] \supset d(x, p) \;\; 0$ TI. 2.1 and TI. 2.2

§4. These two theorems are interesting. The first describes an approach situation relative to q in a means-end expectation chain. The second describes an avoidance situation under the same conditions. If we substitute 'x turns to the left' for 'p', 'x will find a closed door' for 'q', and 'x gets to food' for 'r', the first theorem reads, if x expects that if he turns to the left, then he will find a closed door and expects that if he doesn't find a closed door, then he can get to food and demands that he get to food, then he demands that he not turn left. The avoidance situation represented by the second theorem can be illustrated by substituting 'x turns to the left' for 'p', 'x can walk along a wooden path over the electric grid' for 'q', and 'x gets severely shocked' for 'r'.

TI. 4.5.0: $[pR^e x \,.\, e(x, p \supset q) > 0] \supset e(x, q) > 0$ TI. 2.4 and TI. 2.12

TI. 4.5.00: $[r(x, \varphi, p) > 0 \,.\, e(x, p \supset q) > 0] \supset e(x, q) > 0$ TI. 4.5.0 and TI. 2.7

TI. 4.5.000: $[r(x, \varphi, p) > 0 \,.\, e(x, p \supset \varphi x) > 0] \supset e(x, \varphi x) > 0$ TI. 4.5.00 $\left[\frac{\varphi x}{q}\right]$

55. These three theorems, TI. 4.5.0 - 000, describe various ways in which a member of an expectation chain can be released. The first theorem shows that '$e(x, q) > 0$' is released when p is sensorily received, the second, when p is a stimulus, and the third shows that when 'φx' is a realizable state-of-affairs, it is released as an expectation if x has a tendency to respond by φ to p. (Cf. paragraph 47 in chapter II) These various ways of releasing an expectation find their justification, in general, in chapter III through V of Purposive Behaviorism.

TI. 4.6.0: $[(\sim q)Mx \vee (\sim q)Ux \,.\, e(x, p \supset q) > 0] \supset d(x, \sim p) > 0$ TI. 2.2 and TI. 4.3.010

56. This theorem is the avoidance situation correlate of TI. 4.3.011. Like TI. 4.3.011 it is important for this reason: it shows quite clearly that x's demand that ~p is a consequence of his expectation that p leads to q when the end object in this expectation -namely, q- is either an

ultimate or means avoidance object.

TI. 4.7.0: $[d(x, \sim q) > 0 \,.\, e(x, p \supset \psi x)) > 0 \,.\, e(x, \sim\psi x \supset q) > 0 \,.\, pR^{e}x] \supset r(x, \varphi, p) > 0$

Proof: $[d(x, \sim q) > 0 \,.\, e(x, \sim\psi x \supset q) > 0] \supset d(x, \psi x) > 0$ TI. 2.2 $\left[\frac{\sim\psi x}{p}\right]$ (1)

prop. (1) and TI. 2.6 $\left[\frac{\psi x}{q}\right]$

57. TI. 4.7.0 asserts that when x demands ~q, both expects that p leads to φx which leads to ψx and expects that ~ψx leads to q and, finally, when p is sensorily received by x, then his tendency to respond by φ to p is greater than 0.

TI. 4.8.0: $[d(x, p) > 0 \,.\, d(x, \sim p) > 0] \supset [(\sim p)Mx \equiv pMx]$

Proof: $[d(x, p) > 0 \,.\, d(x, \sim p) > 0] \supset [(\sim p)Mx \,.\, pMx]$ TI. 2.3 $\left[\frac{\sim p}{p}\right]$ and TI. 2.3 (1)

prop. (1) and *5.1

58. TI. 4.8.0 asserts that in a situation where x demands both p and ~p, the means object '(~p)Mx' is equivalent to the means object 'pMx'.

TI. 4.8.1 $p\bar{U}x \supset \sim d(x, \sim p) > 0$

Proof: $p\bar{U}x \supset [d(x, \sim p) > 0 \supset pMx]$ TI. 2.3 and TI. 3.2.00 (1)

$p\bar{U}x \supset [p\bar{U}x \supset \sim d(x, \sim p) > 0]$ (1), Transp and TI. 3.2.011 (2)

prop. (2), Imp and Taut

TI. 4.8.10 $p\bar{U}x \supset \sim(\sim p)Mx$ TI. 4.8.1, TI. 4.3.01 $\left[\frac{\sim p}{p}\right]$ and Transp

TI. 4.8.11 $p\bar{U}x \supset \sim(\sim p)\bar{U}x$

Proof: $pUx \supset \sim d(x, \sim p) > 0$ TI. 8.1 (1)

$\sim d(x, \sim p) > 0 \supset \sim((\sim p)Ux)$ TI. 3.2.00 and Transp (2)

prop: (1) and (2)

59. TI. 4.8.11 affirms that when p is an ultimate goal-object its negate is not. Indeed the two theorems TI. 4.8.10 and TI. 4.8.11, permit us to assert the following very strong condition:

TI. 4.8.100 $pUx \supset [\sim(\sim p)Ux \,.\, \sim(\sim p)Mx]$ TI. 4.8.10 and TI. 4.8.11

TI. 4.8.2: $pUx \supset \sim[d(x, p) = d(x, \sim p) = M > 0]$

Proof: $[d(x, p) = M.M > 0 \,.\, d(x, \sim p) = M.M > 0] \supset pMx$ TI. 2.3.4 *13.13 (1)

$[d(x, p) = d(x, \sim p) = M.M > 0] \supset pMx$ (1) and *13.03 (2)

$[d(x, p) = d(x, \sim p) = M > 0] \supset pMx$ (2) and RM (3)

$\sim(pMx) \supset \sim[d(x, p) = d(x, \sim p) = M > 0]$ (3) and Transp (4)

prop. (4) and TI. 3.2.011

60. TI. 4.8.2 affirms that if p is an ultimate object of x then x does not have conflicting demands with respect to p.

TI. 4.9.0: $[r(x, \varphi, p) > 0 \,.\, (\varphi x)R^e x \,.\, d(x, \chi x) > 0 \,.\, e(x, p \supset (\varphi x \supset (\psi x \supset \chi x))) > 0] \supset r(x, \psi, \varphi x) > 0$

Proof: $[xCp \,.\, e(x, p \supset (\varphi x \supset (\psi x \supset \chi x))) > 0] \supset e(x, \varphi x \supset (\psi x \supset \chi x)) > 0$ TI. 2.4 $[\varphi x \supset (\underset{q}{\underline{\psi x \supset \chi x}})]$ (1

$[d(x, \chi x) > 0 \,.\, e(x, \varphi x \supset (\psi x \supset \chi x)) > 0 \,.\, (\varphi x)R^e x] \supset r(x, \psi x, \varphi x) > 0$ TI. 2.6 $[\underset{p}{\underline{\varphi x}}, \underset{\varphi}{\underline{\psi}}, \underset{q}{\underline{\chi x}}]$ (2

$[xCp \,.\, (\varphi x)R^e x \,.\, d(x, \chi x) > 0 \,.\, e(x, p \supset (\varphi x \supset (\psi x \supset \chi x))) > 0] \supset r(x, \psi, \varphi x) > 0$ (1) and (2) (3

$[pR^e x \,.\, (\varphi x)R^e x \,.\, d(x, \chi x) > 0 \,.\, e(x, p \supset (\varphi x \supset (\psi x \supset \chi x))) > 0] \supset r(x, \psi x, \varphi x) > 0$ (3) and TI. 2.12 (4

prop. (4) and TI. 2.7

61. TI. 4.9.0 shows how the tendency to respond moves along an expectation chain. It is more dramatically represented in the next theorem.

TI. 4.9.00 $[r(x, \varphi, p) > 0 \,.\, (\varphi x)Sx \,.\, d(x, \chi x) > 0 \,.\, (e(x, p \supset (\varphi x \supset (\psi x \supset \chi x))) > 0)] \supset r(x, \psi, \varphi x) > 0$ TI. 4.8.0, TI. 2.7 and TI. 3.4

62. TI. 4.9.00 reads: if x tends to respond by φ to p such that φx is a stimulus to x and x demands χx and expects that (the initial stimulus condition) p leads to (the consequent stimulus condition) φx which in turn is a means to ψx which leads to χx, then x has the tendency to respond by ψ to the stimulus situation φx. This theorem shows how a "response" becomes a stimulus to another "response" whose order and direction is guided by an expectation chain.

TI. 4.10.0: $xC(q \supset \sim q)] \supset \sim e(x, q) > 0$

Proof: $[e(x, q) > 0 \,.\, xC(q \supset \sim q)] \supset \sim e(x, q) > 0$ TI. 2.5 $\left[\frac{q}{p}\right]$ (1)

$xC(q \supset \sim q) \supset [e(x, q) > 0 \supset \sim e(x, q) > 0]$ (1), Comm and Exp (2)

prop. (2), *4.62 and Taut

TI. 4.10.00: $[xCp \,.\, xC(q \supset \sim q)] \supset \sim e(x, p \supset q) > 0$ TI. 4.10.00, TI. 2.4 and Transp

TI. 4.10.000: $[pR^e x \,.\, xC(q \supset \sim q)] \supset \sim e(x, p \supset q) > 0$ TI. 4.9.00 and TI. 2.12

TI. 4.10.0000: $[r(x, \varphi, p) > 0 \,.\, xC(q \supset \sim q)] \supset \sim e(x, p \supset q) > 0$ TI. 4.10.000 and TI. 2.7

TI. 4.11.0 : $[xCp \,.\, d(x, p, q) > 0] \supset e(x, q) > 0$ TI. 2.4 and TI. 3.0.01

TI. 4.11.00: $[xCp \,.\, t^A(x, p, q) > 0] \supset e(x, q) > 0$ TI. 4.11.0 and TI. 2.8

TI. 4.11.01: $[xCp \,.\, F(x, p, q)] \supset e(x, q) > 0$ TI. 4.6.0 and TI. 3.3.00

TI. 4.12.0: $[d(x, q) > 0 \,.\, e(x, p \supset (\varphi x \supset q)) > 0] \supset [r(x, \varphi, p) > 0 \equiv pR^e x]$

Proof: $[d(x, q) > 0 \,.\, e(x, p \supset (\varphi x \supset q)) > 0] \supset [r(x, \varphi, p) > 0 \supset pR^e x]$ TI. 2.7 and Simp (1)

prop. (1) and TI. 2.6

63. This theorem, TI. 4.12.0, asserts that x's having a tendency to respond by φ to p is equivalent to x's having sensorily received p provided that he demands q and expects that p leads to φx which leads to ψx. In general, this theorem gives an "operational definition" -though incomplete- of the notion of response tendency by φ to p. It provides the basis for the organisms' action in a given situation.

TI. 4.12.00: $[d(x, q) > 0 \,.\, d(x, p, \varphi x \supset q) > 0] \supset [r(x, \varphi, p) > 0 \equiv pR^e x]$ TI. 4.12.0 and TI. 3.0.0

TI. 4.12.000: $[d(x, q) > 0 \,.\, t^A(x, p, (\varphi x \supset q)) > 0] \supset [r(x, \varphi, p) > 0 \equiv pR^e x]$ TI. 4.11.00 and TI. 2.8

64. TI. 4.12.000 seems to express a fundamental contention of Purposive Behaviorism, namely that x's tendency to respond by φ to p is equivalent to x's having sensorily received p provided that x demands q and has been taught with respect to p relative to the means-end situation $\varphi x \supset q$. In general, if an organism has been taught with respect to a potential stimulus-response pair, he will actualize that stimulus-response situation if and only if he has sensorily received what turns out to be stimulating state of affairs.

TI. 4.12.0000: $[r(x, \varphi, p) > 0 \,.\, d(x, q) > 0] \supset [t^A(x, p, (\varphi x \supset q)) > 0 \supset pR^e x]$ TI. 2.7, *3.41 and Expt.

TI. 4.12.0001: $[pR^e x \,.\, d(x, q) > 0] \supset [t^A(x, p, (\varphi x \supset q)) > 0 \supset r(x, \varphi, q) > 0]$ TI. 4.12.000 Simp and Imp

Theorems TI. 4.12.0000 and TI. 4.12.0001 may be taken as a pair of Carnapian reduction sentences for the term '$t^A(x, p, (\varphi x \supset q))$' in the proposition '$t^A(x, p, (\varphi x \supset q)) > 0$'.

TI. 4.12.00010: $[pR^e x \,.\, d(x, q) > 0] \supset [t^A(x, p \supset (\varphi x \supset q)) > 0 \supset pSx]$ TI. 4.12.0001 and TI. 3.4.10

TI. 4.12.00010 affirms that when p is sensorily received by x and x demands q, then if x has been taught with respect to p relative to the means-end situation $\varphi x \supset q$, p is a stimulus to x.

TI. 4.13.0: $[t^A(x, p, q) > 0 \vee F(x, p, q)] \equiv d(x, p, q) > 0$ TI. 2.8, TI. 3.3.00 and TI. 2.11

TI. 4.13.1: $[F(x, p, q) \,.\, xC(r = \sim p)] \supset \sim t^A(x, r, q) > 0$

Proof: $[F(x, p, q) \,.\, xC(r = \sim p)] \supset \sim d(x, r, q) > 0$ TI. 3.3.010, *4.51, *4.62 and Imp (1)

prop. (1) and TI. 2.8

TI. 4.13.1 and its corollary require discussion. TI. 4.13.1 asserts that if x is fixated on p -say, 'running down the left alley' as a means to the end q and he has confirmed that 'running down the right alley' is the same as 'not running down the left alley', then he has not been taught with respect to 'running down the right alley as a means to end q. This theorem suggests that despite the fact that other routes may lead to the end q, if x is fixated on p he ignores the others as means to q -superior or inferior though they may be.

TI. 4.14.0: $t^A(x, p, q) > 0 \supset [t(x, p, q) > 0 \equiv d(x, p, q) > 0]$
TI. 2.8 and TI. 2.9

TI. 4.14.0 pictures teachability as the "mark" of demand and hence of expectation. (Cf paragraph 14. in this chapter). This concept, that is, "teachability", will be employed in TII rather than its non-dispositional correlate taughtness. Hence every place "teachability" is said to imply something or other it follows that "taughtness" will also. We gain generality by using teachability.

CHAPTER IV

SYSTEM TII

1. The purpose of TII is to present a part of Tolman's working system. It is mainly concerned with the production of certain laws appearing in Chapters III and IV of *Purposive Behaviorism*. It also presents the deduction of some versions of the latent learning principle. In some cases the postulates are merely elaborations of the postulates in TI. Finally, the postulates of TII are so constructed that they allow the deduction of certain laws concerning fundamental relations such as demand, expectation, response tendency and so on in virtue of their various strengths over a given period of time. TII, in other words, is concerned with the "variability" character of the intervening variables.

Time Arguments in the Formulae in TII

2. Throughout TI, the major emphasis was on the existence or non-existence of demands and expectations: What conditions give rise to a demand or expectation, how the existence of demands or expectations are related to another, and what are the consequences of there being or not being a demand for or an expectation of some state-of-affairs. In TII, on the other hand, we are concerned to note that not all demands or expectations exist with equal strength and to investigate the consequences of these variations in strength. Thus, while in TI, it was suffi-

cient to write such things as "$d(x, p) > 0$", here in §II the formulas will compare the strength of one demand (or expectation) with the strength of another. Hence, we should expect to find such formulas as "$d(x, p) > d(x, q)$" more nearly what will be required.

3. When we first consider comparing strengths of demands, we find there are several basic ways in which a comparison might be made. Let us use some English paradigms to represent these types of comparison:

i. x is more hungry than he is thirsty.

ii. x is more hungry than is y.

iii. x is more hungry now than he was an hour ago.

Still other modes of comparison will occur, for example, x is more hungry than y is thirsty, or x is more hungry than y was an hour ago. But all of these other modes of comparison can be handled by some combination of the methods necessary to handle the three listed modes of comparison. Let us take up one after the other the special problems emerging as we try to symbolize the three typical modes of comparison cited above.

4. For situation i, the symbolism already suggested would be adequate: "$d(x, p) > d(x, q)$". There are, to be sure, circumstances in which a more flexible symbolism would be useful; but let us leave consideration of that until the other cases have been examined.

5. In the case of situation ii, it is easy to see that the <u>first</u> arguments to the demand functor will vary from "x" to "y", as they did not in case i, above: "d(x,) $>$ d(y,)". But perhaps it is less easy to see that the second arguments must also be different one from the other. Nevertheless, they must be. If x is hungry, he is demanding, not food (except as a means), nor merely food satiation; he is demanding his food satiation. That is, if "d(x, p)" represents x's hunger, that is, the strength x's demand for food satiation, the "p" is representing that state of affairs which we could express in English by "x is food satiated". Now, keeping this meaning for "p", if we were to write "d(x, p) $>$ d(y, p)" this would mean that x's demand that x be food satiated is greater than y's demand that <u>x be food satiated</u>. This might compare x's "food-egotism" with y's "food altruism". But it does not compare their hunger drives.

6. Thus it is plain that we cannot compare x's hunger to y's by keeping the second arguments to the demand function identical. On the other hand, to write "d(x, p) $>$ d(y, q)" misses the whole point that we are comparing x and y on a single drive. It would as well represent "x is hungrier than y is thirsty". The solution lies in representing the character of the drive by a <u>predicate</u> variable, and replace "p" and "q" in the above formula by appropriate propositional functions:

$$d(x, \varphi x) > d(y, \varphi y)$$

(Incidentally, we could now even represent the greater "food altruism" of the mother who sacrifices her meal to the child: "$d(x, \varphi y) > d(y, \varphi x)$"!) The more elaborate symbolism required for case ii, can be useful for case i:

$$d(x, \varphi x) > d(x, \psi x)$$

7. The symbolism adequate for case ii is, however, inadequate for situation iii. For in situation iii, were we to use the symbolism given above, both sides of the inequation would be written in the same manner:

$$d(x, \varphi x) > d(x, \varphi x).$$

What is requisite is that the symbolism includes some sign to mark the difference in time at which different demands of x for food satiation occurred.

8. The purposes of TII will be met by a somewhat crude but simple "dating" device. It is the organism-at-a-certain-time which has the demand. Hence we can discriminate the two sides of the inequation by discriminating the organisms-at-a-time:

$$d(x_{t1}, \varphi x) > d(x_{t2}, \varphi x)$$

That is, the demand of x-at-time-t_1, that x be food-satiated is greater than the demand of x-at-time-t_2. No other subscripts than time subscripts will be made on the organism arguments to our various functors. Hence the "t" may be dropped to give

$$d(x_1, \varphi x) > d(x_2, \varphi x).$$

9. Further consideration of the kind of time involvement in the postulates and theorems of TII, together with the necessity of keeping the symbolism as simple as possible, lead to the adoption of certain conventions for interpreting the time notation. (In the concluding chapter of this essay, certain inadequacies of this notation will be pointed out and discussed). Generally speaking, the system is oriented toward utilization to represent the laboratory situation of the experimental psychologist. Here, the details of a date (e.g. January 4, 1956) on which he ran an experiment, are immaterial. Also, it is immaterial that he cannot run all of the rats involved in an experiment simultaneously. What is important is the <u>internal</u>, <u>relative</u> dating of a <u>series</u> of situations, for example, trial runs of a given maze. Thus, in general, the dating subscript refers to the trial run in a temporal series of trial runs. For example,

$$d(x_2, \varphi x) > d(x_1, \varphi x),$$

says that on x's second trial he was hungrier than on his first, whereas

$$d(x_1, \varphi x) = d(y_1, \varphi y)$$

says that x's hunger on his first trial run equalled y's hunger on his first trial run. (Perhaps measured by 24 hours food deprivation each.) It does <u>not</u> mean that x and y were equally hungry at 10 A.M. Perhaps x is run regularly at 10 A.M. and y at 10:15. Then it would mean that x's hunger at 10 A.M. equalled y's at 10:15.

10. A typical sort of experimental situation arises when the experimenter keeps a drive constant through one part of a series of trials and then varies it to a new constant. (One or several animals might be involved, perhaps a control animal or group with the drive kept constant and another animal for which the drive is varied.) Representations of these varieties of situation can be achieved by various manipulations of the following device.

$$(i)\ \ [c \leq i \leq k \supset d(x_i,\ \varphi x) = M].$$

The above says that on every trial from the cth to the kth, the strength of x's demand for φx equals M. Variations on this device, such as,

$$(i)\ \ [c \leq i \leq k \supset d(x_i,\ \varphi x) = d(y_i,\ \varphi y],$$

will be self-explanatory or explained as they occur.

11. Four further time conventions must be explained. (1) The numerical constants, "1", "2", "3", etc. for times represent consecutive trials, with "1" representing the first trial. (2) Also alphabetical symbols (variables) for times "j", "k", etc. represent <u>earlier</u> and <u>later</u> trials according to their alphabetical order, but alphabetically consecutive letters do not necessarily represent consecutive trials. To represent the trial next after j we write "j+1", rather than "k". j+1 may be identical with or earlier than k. (3) The expression "(i) $(c \leq i \leq k \supset \ldots)$" occurs so frequently and is so cumbersome, that we have abbreviated it to "$(i_c{}^k)$ (...)". This symbolism is more or less parallel

to a similar symbolism in mathematics, for example,

$$\Sigma_1^k\ (fx) \text{ and } \Pi_1^k\ (fx).$$

Notice that no essential distortion is produced by this convention. The normal logical operations on any universal quantifier are applicable here. For example, from '$(i_1^k)(\varphi x_i)$' we may conclude, by $*10.1$, 'φx_k'. Carried out in terms of the expression '$(i)(c \leq i \leq k \supset \varphi x_i)$' the operation is as follows: by $*10.1$

$$c \leq k \leq k \supset \varphi x_k.$$

The antecedent of this conditional is true (that is, it is analytic). Hence by a rule of inference we get

$$\varphi x_k$$

which is the same result as in the case of the abbreviated procedure. (4) Lastly, only organism variables in the formulae of TII are subscripted. Strictly speaking, every variable should be dated. However, in TII, there is no gain in such a procedure. We thus adopt the more abbreviated form of subscription. For example,

$$\varphi(x_i,\ a)$$

and

$$d(x_i,\ \varphi x_i) > 0$$

and so on.

Primitive Ideas of TII

12. The primitive ideas of TII include all of those of TI. Two of these primitive ideas, namely,

$$t^A(x,\ p,\ q)$$

that is, "taughtness", and

$$xCp$$

that is, "confirmation", due to space and time considerations, are either not used in the present version TII or are restricted in their importance as compared with their use in TI.

13. TII contains only one additional primitive idea: it is the idea of greater effort. Greater effort will be symbolized by '$E\!\!>$'. It will appear in contexts like

$$\varphi x_i \; E\!\!> \; \psi y_j$$

which means 'φx at time i requires greater effort than ψy at time j';

$$\varphi x_i \; E\!\!> \; \psi x_j$$

which means 'φx at time i requires more effort than ψx at time j';

$$\varphi x_i \; E\!\!> \; \varphi y_j$$

which means 'φx at time i requires greater effort than φy at time j'. Tolman construes this concept as meaning a "greater expenditure of energy".[1] It is, of course, the analogue of "greater action" in physics. There is only one point to be considered here; greater effort makes no reference to final goals. For example, we can say that a newsboy walking up two flights of stairs to deliver a newspaper engages in less effort than an old man going to his office two flights up. Effort is somehow measured by reference

---

1. Purposive Behaviorism, p. 48.

to the agent's capacity. Thus the newsboy going up two flights expends less effort than the old man, <u>not</u> because the newsboy weighs less than the old man, so that less weight was hoisted (although that might be a contributing factor), but because he has more " zip and go " than the old man: he is less completely used up by the expenditure of effort than was the old man.

14. A more complete analysis of Tolman's working system would also include two other concepts as primitive; namely, the <u>incentive value</u> (or valence) of a goal-object and the <u>efficiency</u> of goal-objects.[2] These concepts are, however, not symbolized in System III. Tolman considers them, nevertheless, quite important. It may be considered objectionable to separate greater effort from greater efficiency on the grounds that one is reducible to the other. The defense here is that it is Tolman who makes the separation.[3] However, it is the present author's belief that the separation is a sound one. This point will be taken up again in the last chapter.

15. As a result of the above omission of valence and

---

2. <u>Purposive Behaviorism</u>, p. 14.
3. Compare the remarks on efficiency on p. 14 of <u>Purposive Behaviorism</u> with those on pages 110 and 179 having to do with effort. It will be argued in the final chapter that those behavior patterns which tend to get the animal more easily and quickly to a given goal are not necessarily those which involve lesser effort.

efficiency from TII, all of the postulates and theorems of TII are to be understood as holding true _provided_ these factors of valence and efficiency are held constant. On this convention we may drop consideration of efficiency and valence. This is done, in part, because too little is said about these concepts or, at least, their interrelationships in _Purposive Behaviorism_.[4] Hence, TII does contain certain speculation concerning the properties of '$E\rangle$'. The "properties" of this concept and its relation to the notion of demand and expectation will be made clear in the discussion of the postulates.

Procedures and New Symbols in TII

16. The replacement rules, proof procedures and numeration of the formulae in TII are the same as those in TI. There is one difference. In TII, we do not include theorems deduced from the definitions alone under a separate heading. For example, in TI theorems following from the definitions alone were denoted by the prefix

TI. 3.M.N.

Hence, in TII, the prefix

TII. 1.N.

denotes a definition,

---

4. For example, the discussion of "least action" (that is least effort) in relation to shortness-easiness preference covers in its entirety, at most, two pages (cf. pp. 110 and 179 of _Purposive Behaviorism_). However, Tolman has a lot to say about efficiency throughout his book.

TII.2.N

a postulate, and

TII.3.M.N

theorems which follow from the postulates either alone or in conjunction with other postulates, definitions and/or theorems.

17. As has already been indicated, we shall make liberal use of <u>predicate</u> variables in TII. According to the conventions of TI, Greek letters like, 'φ', 'ψ' and 'χ' are predicate variables. However, in TII, we may use as many as eight predicate variables in a single formula. Hence, in TII we supplement the Greek letters with English capital letters from the beginning of the alphabet. We shall avoid any duplication of an English capital letter which represents a constant in either TI or TII. Hence, the English capital letter 'C' will not be used because it is used in the primitive idea of confirmation in TI (that is, 'xCp').

<u>The Definitions of TII</u>

TII. 1.0 $(\varphi, \psi, \ldots)\ A(x, \chi) =_{Df} [\varphi x \supset \chi x \,.\, \psi x \supset \chi x, \ldots]$

This definition reads: to say that φ, ψ ..., are available to x as means to χ is to say that φx implies χx and ψx implies χx and ....,.

18. Tolman's system requires an idea like means-end availability. For it is in terms of such a condition as is expressed by the concept of means-end availability that we are able to impute to the organism expectations of the elements

in a given problem situation. Expectations are personalized affairs. We assume that the organism expects something in a situation on the basis of that something's availability to the organism as a means to something else. The experimenter, by constructing the problem situation in such and such a way, makes certain things available to the organism. Then he imputes to the organism expectations of these things after the organism has had a certain amount of acquaintance with them in the problem situation. In erecting this concept we are perhaps stretching the use of the sign '$\supset$' (that is, truth-functional implication) a bit. For what '$\varphi x \supset \chi x$' is meant to portray is a situation of the following sort: if x <u>can make</u> φx true then <u>indirectly</u> x <u>can make</u> χx true. In other words, it is via x's being able to make φx true that he is <u>indirectly</u> able to make χx true.

19. This notion of <u>realizability</u> in relation to the values of the variables is not entirely new in TII. It was tacitly present in TI: when we wrote $e(x, p \supset (\varphi x \supset q)) > 0$ the only appropriate substituends for "p" were sensorily received (by x) propositions and for 'φx' and 'q' sensorily receivable (by x) or realizable (by x) propositions.[5]

---

5. For example, consider the postulate TI.2.6. There 'p' is sensorily received and φx is sensorily receivable or realizable (in the sense of "realizable" explained in paragraph 47 of Chapter II). Hence, φx could not replace p because this would have conflicted with p's being sensorily received. For though something which is realiz<u>ed</u> is realiz<u>able</u>, the converse is not necessarily true.

however, in TI those restrictions on propositions substitutable for "p" were limited to certain argument positions for "e", "d" and definitionally derived concepts. Here in TII we must consider the hypothesis that something is objectively available to x.

20. Strictly speaking, we have here two unexpressed primitive ideas: $\varphi A^{v}x$: φx is directly realizable by x; pNq: the world being what it is, p's being true would be sufficient to bring about the truth of q.
Were these PI's formulated means-end availability as defined in TII.1.0 would be the result of a definitional chain such as is represented below.

i $\varphi A(x, \chi) =_{Df} [\varphi A^{v}x \,.\, \varphi x N (\chi A^{v}x)]$,

that is, 'φ is available to x as a means to χ' means 'φx is directly realizable by x' and φx would be sufficient to bring about the state-of-affairs (necessarily implies) that χx is directly realizable by x.

ii $(\varphi, \psi)\, A\, (x, \chi) =_{Df} [\varphi A(x, \chi) \,.\, \psi A\, (x, \chi)]$

Finally, we get to TII.1.0 in the following form

iii $(\varphi, \psi, \ldots)\, A\, (x, \chi) =_{Df} [\varphi A(x, \chi)\ \psi A(x, \chi), \ldots]$

21. Since $(pNq) \supset (p \supset q)$ is valid, it is true that

$$\varphi A(x, \chi) \supset [\varphi x \supset \chi A^{v}x]$$

But since "$\chi A^{v}x$" is only unofficial the best that can be written in TII, on the basis of $(pNq) \supset (p \supset q)$, is

$$\varphi A(x, \chi) \supset (\varphi x \supset \chi x)$$

While the converse does not in general hold, we can and do undertake to reduce the complexity of TII by not making replacements of the definiens by the definiendum in accordance with TI. 2.0 unless the informal conditions suggested in the above unofficial definitions hold.

TI. 1.1 $(\varphi)\ R^e(x, y) =_{Df} [\varphi x R^e x\ .\ \varphi y R^e y]$

TI. 1.1 reads: 'φ is sensorily received by x and y'means 'φx is sensorily received by x and so is φy by y'.

22. TI. 1.1, like the next four definitions is primarily abbreviatory. However, the expression '$\varphi x R^e x$' bears some discussion. The expression 'φx' may depict x as either an agent or merely a location point. For example, let 'φ' be 'jumps to the white card'. Then '$\varphi x R^e x$' means 'x has sensorily received that he jumps to the white card'. Here x is an agent; he sensorily experiences his jumping to the white card when he does it. This is essentially a response condition [but which may be a taking on of properties as a stimulus to something else, that is, 'x jumps to the white card' might be a potential stimulus]. The second case: let'φ' be 'to the left of the white card'. Then '$\varphi x R^e x$' reads 'x has sensorily received that x is to the left of the white card'. Here x is a location point. This latter is more descriptive of merely a stimulus situation. Again, these ideas are not new in TII. For example, consider the expressions '$pR^e x$' and '$(\varphi x)R^e$' in the 4th line of the proof of theorem TI. 4.9.0. 'p' is merely a stimulus

condition; 'φx' is best interpreted as a response condition getting ready to become a stimulus. Indeed, this very point is proven in the next theorem. But the point is 'p' could have taken as a substitution instance 'x is to the left of the white card' where x is a location point or 'p' could have been 'x jumps to the white card' where x is an agent. The same might be said for 'φx' in '$(\varphi x)R^e x$' in TI. 4.9.0.

TII. 1.2 S_x^{1-h-k} $[(B, D, G), (\varphi, \chi), (\psi, H)] =_{Df} [(i_1^h)((r(x_i, B, Gx_i) > 0 \;.\; Bx_i) \supset \varphi x_i \;.\; \varphi x_i \supset \psi x_i \;.\; (r(x_i, D, Gx_i) > 0 \;.\; Dx_i) \supset \chi x_i \;.\; \chi x_i \supset Hx_i) \;.\; (i_{h+1}^k)(r(x_i, B, Gx_i) > 0 \;.\; Bx_i) \supset \chi x_i \;.\; \chi x_i \supset Hx_i \;.\; (r(x_i, D, Gx_i) > 0 \;.\; Dx_i) \supset \varphi x_i \;.\; \varphi x_i \supset \psi x_i)]$

TI. 1.2 defines 'at trial h+1 in the series 1-h-k there is substituted with respect to (organism) x and (his responses by) B and D relative to (the stimulus) G, (the goal-objects) φ for χ and (the goal-objects) ψ for H' as 'during each of the trials from the 1st to the h^{th} (1) x's response by B to the state of affairs Gx implies φx and φ is available to x as a means to ψ and (2) x's response by D to the state-of-affairs Gx implies χx and χ is available to x as a means to H and during each of the trials from the $h+1^{st}$ to the k^{th} (3) x's response by B to the state-of-affairs Gx implies χx and χ is available to x as a means to H and (4) x's response by D to the state-of-affairs Gx implies φx and φ is available to x as a means to ψ'.

23. This definition is thus a definition of substitution

as is described by Tolman in chapters II through V of Purposive Behaviorism. Notice that expressions of the form '$r(x_i, B, Gx_i) > 0 \cdot Ex_i$) occur in it. This is the notion of response in the sense of R_1 discussed in paragraph 47. of chapter II of this essay. Hence the definition suggests, in part, by means of this expression that the animal has been trained. Again, according to the substitution conventions adopted in the discussion of means-end availability the substituends of 'φx_i', 'ψx_i', 'χx_i' and 'Hx_i' are only realizable states-of-affairs.

24. A good way to picture the above definition is this. Think of a Y maze with different goal-boxes at each end. Substitution occurs when the "rewards" (e.g., bran mash and electric shock) in the two goal-boxes are interchanged on trial h+1 and kept in their new locations through trial k.

The conventions of TII and RR make the following biconditional valid.

TII. 1.3 $(i_c^k)(\varphi x_i) \equiv [(i_c^h)(\varphi x_i) \cdot (i_{h+1}^k)(\varphi x_i)]$

TII. 1.3 reads: for every trial i from c to k, φx_i is true if and only if for every i from c to h φx_i is true and for every i from h+1 to k φx_i is true.

25. TII. 1.3 deserves further comment. First of all, according to the conventions of TII (as represented in paragraphs 10 and 11 in this chapter), $c \leq h \leq k$. Suppose now

that c=k. Then h=c; also h=k. Hence, h+1 = k+1. Under these conditions the left hand member of the above biconditional would be equivalent to

$$\varphi x_c,$$

but the right hand member would be equivalent to

$$\varphi x_c \cdot \varphi x_{c+1}$$

and thus the right hand member would not be implied by the left hand member. Accordingly, the following convention is adopted. TII. 1.3 is always used under the <u>tacit</u> assumption that $c < k$; an assumption in practice satisfied in all of the series of trials constituting psychological experiments of the sort undertaken by experimental psychologists. Secondly, TII. 1.3 is not strictly a definition, but its use is primarily abbreviative, hence like that of an abbreviative definition. It is for this reason that TII. 1.3 is listed among the definitions. It <u>really</u> belongs to the conventions of the symbolism and the presupposed system of arithmetic, that is, <u>RN</u>. When a replacement in accordance with it is required in a proof, it is more illuminating to appeal to TII. 1.3 than to merely "RN".

<u>The Postulates of TII</u>

TII.2.0 $\varphi x_i \ \mathrm{E}\!> \psi x_j \supset \sim(\psi x_j \ \mathrm{E}\!> \varphi x_i)$

TII. 2.0 reads: if φx at time i requires greater effort than ψx at time j, then the converse is not true.

26. The truth of this postulate is relatively obvious. It expresses the asymmetry of greater effort. From it we may also deduce that greater effort is not reflexive; that is, substituting 'φ' for 'ψ', and j for i we may deduce

$$\sim(\varphi x_j \mathrm{E}\!\!> \varphi x_j)$$

which again is self-explanatory.

TII. 2.1 $e(x_i, (\varphi x_i \mathrm{E}\!\!> \psi x_j)) > 0 \supset e(x_i, (\psi x_j \mathrm{E}\!\!> \varphi x_i)) = 0$

TII. 2.1 reads: if x at time i expects that φx_i requires more effort than ψx_j, then he does not expect at time i that ψx_j requires more effort than φx_i.

27. This postulate, like the preceding one, is self-explanatory. It states the asymmetry of x's expectation (at a given moment) of greater effort. Again, from TII. 2.1, we may deduce the irreflexivity of x's expectation (at a given moment) of great effort. Substituting 'φ' for 'ψ' and j for i, we get

$$e(x_j, (\varphi x_j \mathrm{E}\!\!> \varphi x_j)) = 0.$$

Notice that, in TII. 2.0 and TII. 2.1, we need not consider the time variable over a given series because these two laws are true for all times.

TII. 2.2 (i_1^k) $[d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0 \;.\; e(x_i, (\varphi x_i \mathrm{E}\!\!> \chi x_i)) > e(y_i, (\varphi y_i \mathrm{E}\!\!> \mathrm{B} y_i)) \;.\; (\varphi, \chi)\ \mathrm{A}\ (x_i, \psi) \;.\; (\varphi, \mathrm{B})\ \mathrm{A}\ (y_i, \psi)] \supset [e(x_k, (\chi x_k \supset \psi x_k)) > e(y_k, (\mathrm{B} y_k \supset \psi y_k))]$

TII. 2.2 reads: if during each of the first k trials, (1) the strength of x's demand for ψx is kept equal to that of y's demand for ψy at a value above 0, (2) x's expectation (or belief) that φx is more effortful than χx is greater than y's belief that φy is more effortful than Ey and (3) φ and χ are available to x as means to ψ and φ and E are available to y as means to ψ, then on the k^{th} trial x's belief that χx leads to ψx will be stronger than y's belief that Ey leads to ψy.

28. This postulate might be called "the postulate of discrimination of effort". It might be represented by the following kind of situation: when there are two animals (or two groups of animals) run under the same type and degree of drive and so trained that one group of animals is better able to discriminate the difference in effort between two actions A and B (where both A and B are means to satisfying the drive) than is the other group with respect to the actions A and C (where both A and C are means to satisfying the drive) then the group with the better discrimination will have a clearer cognition that B leads to satisfying their individual drives than will the other group that C leads to satisfying their individual drives.

TII. 2.3 $[d(x_i, \psi x_i) = d(y_i, \chi y_i) > 0 \;.\; e(x_i, Gx_i \supset (Ex_i \supset \psi x_i)) > e(y_i, Gy_i \supset (\varphi y_i \supset \chi y_i)) \supset (d(x_i, Ex_i) > d(y_i, \varphi y_i))]$

TII. 2.3 reads: if the strength of x's demand for ψx at time i is kept equal to that of y's demand at time i for χy at a value above 0 and x's belief at time i that if Gx then Ex is a means to ψx is stronger than y's belief at time i that if Gy then φy is a means to χy, then x's demand at time i for Ex will be stronger than y's demand at time i for φy.

29. This postulate is an analogue of the postulate TI. 2.1. It characterizes the greater tendency of an animal x to approach a goal-object in view of his greater knowledge when compared with another animal y. TII. 2.3 thus describes a kind of approach situation under the conditions of the same drive but greater knowledge, that is, stronger vs. weaker beliefs. This postulate finds justification in part in chapters III and IV of Purposive Behaviorism.

TII. 2.4 $(i_1^k)\ [d(x_i, \psi x_i) = d(y_i, Ey_i) > 0\ .\ \varphi x_i\ E{>}\ \chi x_i\ .\ \sim(\varphi y_i\ E{>}\ Dy_i)\ .\ \sim(\varphi x_i\ E{>}\ \varphi y_i)\ .\ \sim(\varphi y_i\ E{>}\ \varphi x_i)\ .\ (\varphi, \psi)\ A\ (x_i, \psi)\ .\ (\varphi, D)\ A\ (y_i, E)] \supset [e(x_k, (\varphi x_k\ E{>}\ \chi x_k)) > e(y_k, (\varphi y_k\ E{>}\ Dy_k))]$

TII. 2.4 reads: if during each of the first k trials, (1) the strength of x's demand for ψx is kept equal to that of y's for Ey at a positive value above 0, (2) φx requires more effort than χx, φy does not require more effort than Dy, φx does not require more effort than φy, φy does not require more effort than φx, (3) both φ and χ are available to x as means to ψ and both φ and D are available to y as means to E, then on the k^{th} trial, x's expectation that

φx is more effortful than χx is greater than y's expectation that φy is more effortful than Dy. The point of TII. 2.4 is this: if the difference in effortfulness between φx and χx is greater than that between φy and Dy, then x will acquire in a given series of trials stronger anticipation of the fact of difference.

30. It should be noted in this postulate (and indeed, in <u>all</u> of the postulates which prescribe a training period for k trials and a consequence of that training at the k^{th} trial) that if k is not high enough the difference predicted in the conclusion will not be discernible even though it is there. For example, consider a maze in which x (and y) cannot retrace on one trial so as to examine both alternatives on one trial. Then the predicted difference will not even <u>exist</u> unless k is high enough to allow exploration of both alternatives. Notice again that the greater the disparity between φx and χx as compared to that between φy and Ey, the lower k may be and still allow the predicted difference to be discernible by the experimenter, for example, in terms of the frequency with which x chooses ψx rather than φx as compared with the frequency with which y chooses Dy rather than φy. What this means is that the postulates and theorems should be written, if complete strictness were to be observed, in the form

$$(\exists k)\ [\text{Postulate}]$$

In general, if a postulate holds for k then it holds also for k+1; that is, if k is high enough that is all one need worry about; one can't get k too high. But this is not always true. For example in postulate TII. 2.7 to follow, if k is too high the postulate will not hold.

TII. 2.5 $[(d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0 \;.\; e(x_i, Ex_i \supset (\chi x_i \supset \psi x_i)) > e(y_i, Ey_i \supset (\varphi y_i \supset \psi y_i)) \;.\; (E)R^e(x_i, y_i)) \supset (r(x_i, \chi, Ex_i) > r(y_i, \varphi, Ey_i))]$

TII. 2.5 reads: if x's demand at time i for ψx is equal to y's demand at time i for ψy and x's expectation at time i (or anticipation) that if Ex then if χx then ψx is greater than y's expectation at time i that if Ey then if φy then ψy and E is sensorily received by both x at time i and y at time i, then x's tendency to respond at time i by χ to Ex is greater than y's at time i by φ to Ey. The point of this postulate amounts to this: the greater the expectation that upon the occurrence of a certain stimulus if a certain response is made then a certain goal will be attained, the greater the tendency is to make that response to that stimulus.

31. It will be seen that this postulate is no more than an extension of postulate TI. 2.6 in the preceding chapter. Here, again, we are comparing different animals with different strengths of anticipation. Hence, we shall refer to TII. 2.5 as the response postulate.

TII. 2.6 $(i_1^k)\ [d(x_i, \chi x_i, \psi x_i) > d(y_i, \varphi y_i, By_i) \,.\, d(x_i, \psi x_i) = d(y_i, By_i) > 0 \,.\, [(r(x_i, D, Gx_i) > 0 \,.\, Dx_i) \supset \chi x_i] \,.\, [(r(y_i, D, Gy_i) > 0 \,.\, Dy_i) \supset \varphi y_i]] \supset [e(x_k, Gx_k \supset (Dx_k \supset \psi x_k)) > e(y_k, Gy_k \supset (Dy_k \supset By_k)]$

TII. 2.6 reads: if during each of the trials from 1 to k (1) x's demand for χx as a means to ψx is greater than y's demand for φy as a means to By, (2) x's demand for ψx is kept equal to y's for By which is greater than 0 and (3) x's response by D to Gx implies χx and y's response by D to Gy implies φy, then on the k^{th} trial x's anticipation that if Gx then if Dx then ψx will be stronger than y's that if Gy then if Dy then By. This postulate provides the inferential basis of Tolman's contention that routes to more demanded or "better goal-objects" are chosen more often than routes to less demanded or "poorer goal-objects" as means to the alleviation of some drive state (cf. pp. 71-77 in Purposive Behaviorism).

32. Attention is directed to the k in the quantifier; k is a variable taking as values any number down to and including 1. The point is this: though we don't know how many, some trials are needed to build up the expectations hypothecated in the consequent of TII. 2.6. Indeed if k = 2, then these expectations could have been built up in one trial. Hence, we allow in this postulate the possibility for one trial learning. (cf. Purposive Behaviorism, p. 73.)

TII. 2.7 $[(i_1^k)\,[E \neq D \,.\, \varphi \neq \chi \,.\, d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0 \,.$
$(r(y_i, E, Gy_i) > 0 \,.\, Ey_i) \supset \varphi y_i \,.\, \varphi y_i \supset \psi y_i \,.\, (r(y_i, D, Gy_i) > 0$
$.\, Dy_i) \supset \chi y_i . \chi y_i \supset Hy_i] \,.\, [S_x^{1-h-k}\,[(E, D, G), (\varphi, \chi), (\psi, H)]]$
$\supset (i_{h+1}^k)\,[e(y_i, Gy_i \supset (Ey_i \supset \psi y_i)) > e(x_i, Gx_i \supset (Ey_i \supset \psi x_i))]$

TII. 2.7 reads: Throughout each of k trials if (1) E is not identical with D and φ is not identical with χ, (2) x's demand for ψx is kept equal to y's for ψy which is above O, (3) y's response by E to Gy implies φy and φ is available to y as a means to ψ, and y's response by D to Gy implies χy and χ is available to y as a means to H and (4) there is substituted for x at trial h+1 in the trial series 1-h-k with respect to E and D relative to G, φ for χ and ψ for H, then for every trial i from h+1 to k y's anticipation that if Gy then if Ey then ψy is stronger than x's expectation that if Gx then if Ex then ψx.

33. Quite briefly this postulate says that where two animals are compared under the same drive of the same strength and where one of these groups experiences a substitution and the other no substitution, then the animal experiencing the substitution will, at some time after the substitution but before the end of the experiment, show less a tendency to expect that the original path leads to the old goal-object than will the animal who did not experience substitution. This postulate will be called the postulate of "substitution". Its justification may be found in chapter IV of <u>Purposive Behaviorism</u>.

34. This postulate would not hold without the stipulation of the differences between the responses B and D and the goal-objects φ and χ. For were they allowed, in some cases, to be the same there would be no substitution and the predicted consequence would not follow. This is important because the consequent of this postulate shows a certain disruptive character in the behavior of the animal getting the substitution. (cf. Chapter V, Purposive Behaviorism). Tolman is quite explicit about pointing out that disruption effects are only observed in substitution situations. One would not find disruption effects in latent learning because there is no substitution taking place in these studies.

TII. 2.8 $[(\varphi x_i\ E > \chi x_i \cdot \varphi y_i\ E > \chi y_i \cdot \varphi y_i\ E > \chi x_i \cdot \chi y_i \supset \psi y_i \cdot d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0 \cdot t(x_i, \chi x_i, \psi x_i) > 0) \supset (d(x_i, \chi x_i, \psi x_i) > d(y_i, \varphi y_i, \psi y_i)]$

TII. 2.8 reads: where φx at time i requires more effort than χx at time i, φy at time i requires more effort than χy at time i and φy at time i involves more effort than χx at time i and χy at time i leads to ψy at time i and x's demand at time i for ψx is equal to y's at time i for ψy, where x at time i is teachable with respect to χx as a means to ψx, then x's demand at time i for χx as a means to ψx is greater than y's at time i for φy as a means to ψy. In effect, this postulate is that the less effortful goal-object for animal x is a better goal-object for x than is the more effortful goal-object for y provided that x is teachable with respect to better goal-object.

35. Notice that this postulate is so "rigged" that its truth depends upon the valence and efficiency conditions being held equal. In other words, the less effortful goal-object (where there are only two worthwhile possibilities) is the better goal-object provided the conditions of valency and efficiency are held constant. This postulate is useful in the present version of TII as a means to getting an operational definition of docility.

TII. 2.9 $[(i_1^k)[(B \neq D \,.\, \varphi \neq \chi \,.\, d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0$
$.\, \varphi x_i \; E{>} \; \chi x_i \,.\, \varphi y_i \; E{>} \; \chi y_i \,.\, \varphi y_i \; E{>} \; \chi x_i \,.\, (r(y_i, B, Gy_i) > 0$
$.\, By_i) \supset \varphi y_i \,.\, \varphi y_i \supset \psi y_i \,.\, (r(y_i, D, Gy_i) > 0 \,.\, Dy_i) \supset$
$\chi y_i \,.\, \chi y_i \supset Hy_i) \,.\, S_x^{1-h-k} [(B, D, G)(\varphi, \chi), (\psi, H)]]] \supset$
$[(r(y_j, B, Gy_j) > r(x_j, B, Gx_j) \,.\, r(x_k, B, Gx_k) > r(y_k, B, Gy_k))$
$\supset (i_1^k)(t(x_i, \chi x_i, \psi x_i) > 0)]$

TII. 2.9 reads: if during each of the first k trials (1) B is not identical with D and φ is not identical with χ, (2) x's demand for ψx is kept equal to y's for ψy which is above 0, (3) φx is more effortful than χx, φy is more effortful than χy and φy is more effortful than χx, (4) y's response by B to Gy implies φy and φy implies ψy, (5) y's response by D to Gy implies χy and χy implies Hy and (6) there is substituted for x at trial h+1 in the series 1-h-k with respect to B and D relative to G, φ for χ and ψ for H, then if on the j^{th} trial y's tendency to respond by B to Gy is greater than x's tendency to respond by B to Gx and on the k^{th} trial x's tendency to respond by B to Gx is greater than y's tendency to respond by B to Gy, then throughout every trial from 1 to k x is teachable with respect to χx as a means to ψx.

36. TII. 2.9 presents certain conditions under which it is possible to infer the teachability of x with respect to a given means-object. It is therefore called the "teachability postulate". It amounts to this: given disruption in x's behavior on the j^{th} trial preceding the k^{th} trial if recovery by x on the k^{th} trial is made in terms of the least effortful object which was substituted at trial h+1 in the trial series 1-h-k, then x is teachable throughout that series with respect to that least effortful object. The justification for this postulate may be found on pages 14, 74, 442 and 443 of Purposive Behaviorism.

37. One final point. This postulate shows why we require response in the sense of R_1 described in chapter II of this essay, that is, why we need an expression having the form of

$$(r(x_i, B, Gx_i) > 0 \;.\; Rx_i).$$

For we must be able to claim that the animal had experience with the goal-objects during the early part of the experiment. Notice that we could not use the notion of response in the sense of R_2, that is,

$$r(x_i, B, Rx_i) = 1.$$

For this notion says that the frequency of response (in the sense R_1) is equal to 1. In the above postulate both animals would then be able to proceed along the various appropriate paths, say, of a maze, only when the frequency of response reached 1. But then the consequent which shows x having a greater response tendency on the k^{th} trial than y for the appropriate path would be false because y's response tendency and x's response tendency for the appropriate path already equalled 1 (the highest possible number) during the training. Response in the sense of R_1 avoids this consequence. It says that x has the tendency to respond by E to Gx and Ex is true. Hence x (and y) will begin to accumulate knowledge of the appropriate paths of the maze only when both of these conditions hold true. In general, both of these conditions would not be met on every trial. Hence the difference from

$$r(x_i, B, Gx_i) = 1.$$

TII. 2.10 $[(i_1^k)(z)[d(z_i, Hz_i) = 0 \supset d(z_i, \chi z_i) = r > 0 \,.$
$\varphi \neq \psi \,.\, D \neq G \,.\, Ez_i \supset ((r(z_i, \varphi, Ez_i) > 0 \,.\, \varphi z_i) \supset Dz_i, a) \,.$
$Ez_i \supset ((r(z_i, \psi, Ez_i) > 0 \,.\, \psi z_i) \supset Gz_i, b) \,.$
$(Ez_i \,.\, \varphi z_i) \supset {\sim}Gz_i, b \,.\, (Ez_i \,.\, \psi z_i) \supset {\sim}Dz_i, a] \,.$
$(i_1^h)\ [Ex_i \,.\, {\sim}\, Ey_i \,.\, d(x_i, Mx_i) = d(y_i, My_i) = t < r \,.$
$t^A(x_i, \varphi x_i, \chi x_i) > 0] \,.\, (i_{h+1}^{h+2})\ [Ja] \,.\, (i_{h+2}^{k})\ [Ex_i \,.$
$Ey_i \,.\, d(x_i, Mx_i) = d(y_i, My_i) = s > r \,.\, Ja \,.\, (Ja \,.\, Dx_i, a) \supset$
$Mx_i \,.\, (Ja \,.\, Dy_i, a) \supset My_i]] \supset [e(x_k, Ex_k \supset (\varphi x_k \supset Mx_k)) >$
$e(y_k, Ey_k \supset (\varphi y_k \supset My_k))]$

TII. 2.10 reads: if (I) during the first k trials, for every z it is true that: (1) z does not demand Hz implies that z's demand for χz is equal to r which is greater than

0 (2) φ is not identical with ψ and D is not identical with G, (3) Ez implies that if z responds by φ to Ez then Dz,a and Ez also implies that if z responds by ψ to Ez then Gz,b, (4) Ez and φz implies not Gz,b and Ez and ψz implies not Dz,a, and if (II) during the first h of those k trials Ex and not Ey and x's demand for Mx is equal to y's demand for My, which is equal to t which is less than r, and x has been taught that φx is a means to χx and if (III) during the trials h+1 and h+2, Ja, and if finally (IV) during the trials h+2 to k, (1) both Ex_1 and Ey_1 and (2) x's demand for Mx equals y's demand for My which is equal to s which is greater than r and (3) Ja and (4) Ja and Dx,a implies Mx and Ja and Dy,a implies My, <u>then</u> on the k^{th} trial x's anticipation that if Ex then φx leads to Mx is greater than y's expectation that if Ey then φy leads to My.

39. This postulate is called the latent learning principle. What it says is briefly this: given a situation in which one group of animals has prior training under some drive and another group does not under that drive, then when both groups are introduced into the same situation (but with the drive changed) the group which had the prior training will show transference and hence their performance will be better than the non-trained group. This claim is supported by the following illustration of TII. 2.10. Let 'Hz' be 'z alleviates

his thirst', 'χz' be 'z satisfies his curiosity', 'Ez' be 'z is in the starting box', 'φ' be 'taking the right route', 'Dz,a' be 'z gets into food-box a', 'ψ' be 'taking the left route', 'Gz,b' be 'z gets to water-box b', 'Ja' be 'a contains food', and 'Mx (and y)' be 'x (and y) alleviates his hunger'. Under these conditions, TII. 2.10 reads: if (I) during the first k trials, it is true for every z that (1) z does not demand that he alleviate his thirst implies that his demand that he satisfy his curiosity equals r (which is greater than 0) (2) taking the right route is not the same response as taking the left route and getting into the food-box is not the same response as getting into the water-box, (3) z is in the starting-box implies that if he responds by taking the right route to the fact that he is in the starting-box, then he gets into the food-box a and z is in the starting-box also implies that if he responds by taking the left route to the fact that he is in the starting-box, then he gets into the water-box b and (4) when z is in the starting-box and takes the right route, he does not get into water-box b and also when he is in the starting-box and takes the left route, then he does not get into food-box a, and if (II) during the first h of those k trials (1) x (an experimental animal) is in the starting-box but y (a control animal) is not (2) both x's and y's demands that their hunger be alleviated is equal to t (which is less than r) and (3) x has been taught that x takes the right route leads to x satisfying his curiosity, and if (III) during the trials h+1 and h+2 (food-box) a contains food, and if finally (IV) during the trials h+2 to k, (1) both x and y are in the starting-box (2) their demands for hunger-alleviation are equal to s (which is greater than r) (3) (food-box) a contains food and (4) if when (food-box) a contains food and x gets into the food-box a then x's hunger is alleviated and if when (food-box) a contains food and y gets into the food-box a, then y's hunger is alleviated, then on the k^{th} trial x's anticipation that if he is in the starting-box then if he takes the right route, his hunger will be alleviated is greater than y's anticipation that if he is in the starting-box then if he takes the right route, his hunger will be alleviated.

40. Because of its importance for Tolman's theory, it would be well to discuss this postulate a bit more. First of all the latent learning principle as presented in III is typically illustrated on a maze which has two distinct goal-boxes and whose mutually exclusive routes to the goal-box are so arranged that (1) retracing can be made possible or not possible and (2) the amount of effort involved in running one route is the same as the amount of effort involved in running the other. (But this latter is not included in the postulate because the influence of the effort factor may be either non-existent or negligible in this

kind of learning situation. At any rate, the postulate is neutral on this matter and the illustrative conditions are in cautious language.) Examples of mazes which would satisfy the above design are simple T mazes, simple Y mazes, rectangular mazes whose routes are the same length, and so on. Secondly, equality of taughtness for either route is measured by equal amounts of training on those routes, for example, by 50% of the runs on one route. (It should be noted, however, that equality of taughtness on either route could be measured by a certain criterion of accuracy for picking the route leading to food, and so on.) Thirdly, we are supposing the animals in both the control and experimental groups to be animals of the same size so that no inequality of effort is established between groups through some physical advantage in one group as opposed to the other (but cf. the first point above). Fourthly, the control group in this version -which is represented by the variable 'y'- is always the group which does not receive preliminary training. Fifthly, we are assuming equal intelligence for all the animals. (Indeed, one <u>may</u> regard the Buxton version of latent learning -where 50% of the animals never mastered the maze perhaps as the result of their relative stupidity- as an "intelligence test" for the animals.) Sixthly, as usual, we are assuming that the factors of valence and efficiency are being held constant.

41. Now in the light of these conditions let us examine the above postulate. That part of the postulate prefixed by (I) describes the common situation into which both the experimental animals, represented by 'x', and the control animals are introduced. The part of the postulate prefixed by (II) describes the experimental group's pre-training in the common situation. The part of the postulate prefixed by (III) indicates the introduction of something into the goal-box (or goal-boxes) and the last part of the postulate (IV) describes the test series where both control and experimental animals are introduced into the common situation. The advantage predicted for x is based on his ability (due to prior training) to discriminate the different routes to the goal-boxes; the control group has not had a chance to make these discriminations prior to introduction in the maze at trial h+2. The point to notice is that the postulate employs certain concepts in Tolman's system and, furthermore, appears to be justified by the remarks (especially) on pages 343-344 and in chapter III of _Purposive Behaviorism_. If these claims are justified it is easy to prove that latent learning in at least three of its versions (as classified by Thistlethwaite) can be deduced from Tolman's theory. The Blodgett version is not so deduced; it will be discussed at an appropriate time in the section on "The Theorems of TII".

TII. 2.11 $\varphi x_i \text{ E}\!> \psi y_j \,.\, \psi y_j \text{ E}\!> \chi z_k \supset \varphi x_i \text{ E}\!> \chi z_k$

TII. 2.11 reads: if φx at time i is more effortful than ψy at time j and ψy at time j is more effortful than χz at time k, then φx at time i is more effortful than χz at time k. This postulate states the transivity of greater effort.

The Theorems of TII

The following theorems are derived from the postulates either alone or in conjunction with other postulates, definitions and/or theorems.

TII. 3.0 $\sim(\varphi x_i \text{ E}\!> \varphi x_i)$ TII. 2.0 $[\frac{\varphi}{\psi}, \frac{i}{j}]$, *1.01 and Taut.

TII. 3.1 $e(x_i, \varphi x_i \text{ E}\!> \varphi x_i) = 0$ TII. 2.1 $[\frac{\varphi}{\psi}, \frac{i}{j}]$, *1.01, Taut and RE

42. TII. 3.0 and TII. 3.1 present, respectively, elementary properties of the ideas of "greater effort" and "the expectation of greater effort".

TII. 3.0 asserts that φx at time i is not more effortful than itself.

TII. 3.1 asserts that x at time i does not expect that φx requires more effort than φx.

TII 3.2 $(i_1^k)[(d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0 \,.\, \varphi x_i \text{ E}\!> \chi x_i \,.\, \sim(\varphi y_i \text{ E}\!> Dy_i) \,.\, \sim(\varphi x_i \text{ E}\!> \varphi y_i) \,.\, \sim(\varphi y_i \text{ E}\!> \varphi x_i) \,.\, Dy_i \text{ E}\!> \chi x_i \,.\, (\varphi,\chi) \text{ A } (x_i, \psi) \,.\, (\varphi, D) \text{ A } (y_i, \psi)] \supset [e(x_k, (\varphi x_k \text{ E}\!> \chi x_k)) > e(y_k, (\varphi y_k \text{ E}\!> Dy_k))]$

Proof: TII. 2.4, TII. 2.11 $[\frac{i}{J},\frac{i}{K}]$ $[\frac{\psi}{E}]$ prop.

TII. 3.2 affirms that if during each of the first k trials if x's demand that ψx is equal in strength to y's demand that ψy which is greater than 0, and φx involves more effort than χx, and φy does not involve more effort than Dy and φx does not involve more effort than φy and φy does not involve more effort than φx but nevertheless Dy involves more effort than χx, and φ and χ are available to x as means to ψ, and φ and D are available to y as means to ψ, then on the k^{th} trial x's knowledge (i.e., expectation) that φx involves more effort than χx is greater than y's knowledge (i.e., expectation) that φy involves more effort than Dy.

43. This principle, TII. 3.2, seems to express the Tolmanian version of the law of least effort. There can be no doubt that Tolman's system requires such a law, as is evidenced by Tolman's remarks in Purposive Behaviorism.[6] The distinctive mark of Tolman's version of least effort is seen in the consequent of the above theorem. It will be noticed that what is expected is the difference in effort between two goal-objects. In short, before the organism acts on an actual difference in effort between goal-objects he must learn (that is, acquire the knowledge) that there is such a difference. It should be

---

6. Cf. Discussion on page 110 of Purposive Behaviorism.

noticed that this law makes little sense whatever unless we remember the proviso stated at the beginning of this chapter, namely, that we are assuming that the valence conditions and the efficiency conditions with respect to the means objects are held constant whenever the antecedent of any theorem contains the expression 'E>'.

TII. 3.2.0 $(i_1^k)\ [d(x_i,\ \psi x_i) > 0\ .\ \varphi x_i\ E\!>\ \chi x_i\ .\ (\varphi,\ \chi)A\ (x_i,\ \psi)]\ \supset\ e(x_k,\ (\varphi x_k\ E\!>\ \chi x_k)) > 0$

The proof of this theorem will be more or less fully developed in order to show the general character of the proofs in TII.

Proof: $[(i_1^k)\ [d(x_i,\ \psi x_i) = d(x_i,\ \psi x_i) > 0\ .\ \varphi x_i\ E\!>\ \chi x_i\ .\ \sim(\varphi x_i\ E\!>\ \varphi x_i)\ .\ \sim(\varphi x_i\ E\!>\ \varphi x_i)\ .\ \sim(\varphi x_i\ E\!>\varphi x_i)\ .\ \varphi x_i\ E\!>\ \chi x_i\ .\ (\varphi,\ \chi)\ A\ (x_i,\ \psi)\ .\ (\varphi,\ \varphi)\ A\ (x_i,\ \psi)]]\ \supset\ [e(x_k,\ (\varphi x_k\ E\!>\ \chi x_k)) > e(x_k,\ (\varphi x_k\ E\!>\ \varphi x_k))]$ TII. 3.2 $[\frac{\varphi}{D}, \frac{x}{y}]$ (1)

$[(i_1^k)\ [d(x_i,\ \psi x_i) = d(x_i,\ \psi x_i) > 0]\ .\ (i_1^k)(\varphi x_i\ E\!>\ \chi x_i)\ .\ (i_1^k)\ [\sim(\varphi x_i\ E\!>\ \varphi x_i)\ .\ \sim(\varphi x_i\ E\!>\ \varphi x_i)\ .\ \sim\ (\varphi x_i\ E\!>\ \varphi x_i)]\ .\ (i_1^k)\ [(\varphi,\ \chi)\ A\ (x_i,\ \psi)\ .\ (\varphi,\ \varphi)\ A\ (x_i,\ \psi)]]\ \supset$ SC (1), *10.22 and Taut. (2)[7]

$[(i_1^k)\ [d(x_i,\ \psi x_i) = d(x_i,\ \psi x_i) > 0]\ .\ (i_1^k)\ [\varphi x_i\ E\!>\ \chi x_i]\ .\ (i_1^k)\ [(\varphi,\ \chi)\ A\ (x_i,\ \psi)\ .\ (\varphi,\ \varphi)A(x_i,\ \psi)]]\ \supset$ SC (2), TII. 3.0 and *9.13 (3)

$[(i_1^k)[d(x_i,\ \psi x_i) > 0]\ .\ (i_1^k)\ [\varphi x_i\ E\!>\ \chi x_i]\ .\ (i_1^k)\ [(\varphi,\ \chi)A\ (x_i,\ \psi)\ .\ (\varphi,\ \varphi)A(x_i,\ \psi)]]\ \supset$ SC (3), RN (4)

---

7. The expression "SC" means "same consequent as the preceding theorem or numbered theorem (or line in a proof)".

$[(i_1^k)[d(x_i, \psi x_i) > 0] \;.\; (i_1^k)[\varphi x_i \;\mathrm{E}\!> \chi x_i] \;.\; (i_1^k)[\varphi x_i \supset \psi x_i \;.$
$\chi x_i \supset \psi x_i \;.\; \varphi x_i \supset \psi x_i \;.\; \varphi x_i \supset \psi x_i]] \supset$ SC (4), TII. 1.0 (5)
$[(i_1^k)[d(x_i, \psi x_i) > 0] \;.\; (i_1^k)[\varphi x_i \;\mathrm{E}\!> \chi x_i] \;.\; (i_1^k)[(\varphi, \chi)\, A$
$(x_i, \psi)]] \supset$ SC (5), Taut and TII. 1.0 (6)

$[(i_1^k)[d(x_i, \psi x_i) > 0 \;.\; \varphi x_i \;\mathrm{E}\!> \psi x_i . (\varphi, \chi) A (x_i, \psi)]] \supset e(x_k,$
$(\varphi x_k \;\mathrm{E}\!> \chi x_k)) > 0 \;.$ (6), TII. 3.1, RR and $*10.22$

44. TII. 3.2.0 might be called the law of least effort for a single organism (or a single group of organisms). It is helpful in the sense that it shows clearly what elements are involved in that law without the added encumbrances of a control group. If we let 'ψx_i' represent, 'x alleviates his hunger', 'φx_i' represent 'x eats dry bran mash' and 'χx_i' represent 'x eats damp bran mash', the above law so interpreted reads: if during the first k trials x demands that he alleviate his hunger and x eats dry bran mash is more effortful for x than x eats damp bran mash and x eats dry bran mash and x eats wet bran mash are available to x as means to alleviating hunger, <u>then</u>, on the k^{th} trial x's knowledge (expectation) that x eats dry bran mash is more effortful than x eats damp bran mash is greater than 0.

TII. 3.2.1 $(i_1^k)[d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0 \;.\; \varphi x_i \;\mathrm{E}\!> \chi x_i \;.$
$\sim(\varphi y_i \;\mathrm{E}\!> Dy_i) \;.\; \sim(\varphi x_i \;\mathrm{E}\!> \varphi y_i) \;.\; \sim(\varphi y_i \;\mathrm{E}\!> \varphi x_i) \;.\; Dy_i \;\mathrm{E}\!> \chi x_i \;.$
$(\varphi, \chi)\, A\, (x_i, \psi) \;.\; (\varphi, D)\, A\, (x_i, \psi)] \supset [e(x_k, \chi x_k \supset \psi x_k) >$
$e(y_k, Dy_k \supset \psi y_k)]$

Proof: prop. TII. 3.2 and TII. 2.2

45. This theorem, TII. 3.2.1, seems to express part of the meaning of the following proposition stated by Tolman on pages 67-68 of Purposive Behaviorism

> If different groups of animals, but with the same physiological drive are run with different goal-objects, the groups run with certain goal-objects learn faster than the others.

I say "part of the meaning" because the expression "certain goal-objects" could conceivably refer to the different valences of the goal-objects, the difference in efficiency of one goal-object compared with another, or the difference in effortfulness of one goal-object compared with another. It will be noticed that TII. 3.2.1 is concerned with the difference in effort only. The proof sketch shows that this law is, in part, a direct product of the law of least effort. Consequently, it may be summarized as follows: Given the conditions prevailing in the antecedent of the law of least effort (TII. 3.2), then on the k^{th} trial x's knowledge that χx leads to ψx is greater than y's knowledge that Dy leads to ψy.

TII. 3.3 $[(i_1^k)[(B \neq D \, . \, \varphi \neq \chi \, . \, d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0$
$. \, (r(y_i, B, Gy_i) > 0 \, . \, By_i) \supset \varphi y_i \, . \, \varphi y_i \supset \psi y_i \, . \, (r(y_i, D, Gy_i)$
$> 0 \, . \, Dy_i) \supset \chi y_i \, . \, \chi y_i \supset \psi y_i] \, . \, S_x^{1-h-k} [(B, D, G), (\varphi, \chi),$
$(\psi)]] \supset [e(y_j, Gy_j \supset (By_j \supset \psi y_j)) > e(x_j, Gx_j \supset (Bx_j \supset \psi x_j)]$

TII. 2.7 $[\frac{\psi}{H}]$and RN

46. TII. 3.3 will be called the "disruption" theorem. It seems to capture Tolman's intentions concerning the notion of disruption as they are expressed in the following remarks quoted from page 76 of Purposive Behaviorism. Tolman writes:

> And we shall suppose in general that in behavior there is always immanent the expectation of some more or less specific type of goal-object. If such type of goal-object be not found, whether it be because a better, a worse, or merely a different, goal-object has been substituted for it, then the animal's behavior will show some sort of disruption such as hunting, startled speeding up, or what not.

TII. 3.3 describes the conditions under which disruption can be predicted where the means objects lead to the same end. If we take the typical hunger situation, then TII. 3.3 amounts to saying that where two organisms (or two groups of organisms) have the same strength of demand for the alleviation of hunger, if one organism or group of organisms, that is, the experimental group, is shifted from one goal-object to another such that either goal-object leads to hunger alleviation, while the other organism or group of organisms, that is, the control group, does not experience the shift in goal-objects though in fact either goal-object leads to hunger alleviation, then at some trial; after the shift (h+1) but before the end of the experiment (k), the experimental organism's knowledge that the path on which he was trained (which led to the original goal-object) leads to hunger alleviation will be less than the control organism's knowledge that that same path leads to hunger alleviation. Notice that this theorem shows the "substitution" as taking place only with respect to means objects. The experiment of Elliot's which Tolman cites on pages 72-74 of Purposive Behaviorism shows the change as taking place both with respect to means-objects and final goal-objects. This situation is easily deduced by substituting in postulate TII. 2.7 $[\frac{\sim\psi}{H}]$. Because of its similarity to TII. 3.3 the theorem will not be produced here.

TII. 3.3.0 $SA \supset [[d(x_j, \psi x_j) = d(y_j, \psi y_j) > 0 \,.\, (G)R^e (x_j, y_j))] \supset [r(y_j, B, Gy_j) > r(x_j, B, Gx_j)]]$ TII. 3.3 and TII. 2.5 $[\frac{G}{E}, \frac{E}{\chi}, \frac{B}{\varphi}, \frac{\chi}{y}, \frac{y}{x}]$, *10.1

TII. 3.3.00 $[(i_1^k) [B \neq D \,.\, \varphi \neq \chi \,.\, d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0 \,.\, (G)R^e(x_i, y_i) \,.\, (r(y_i, B, Gy_i) > 0 \,.\, By_i) \supset \varphi y_i \,.\, \varphi y_i \supset \psi y_i \,.\, (r(y_i, D, Gy_i) > 0 \,.\, Dy_i) \supset \chi y_i \,.\, \chi y_i \supset \psi y_i] \,.\, S_x^{1-h-k}[(B, D, G), (\varphi, \chi), (\psi)]] \supset [r(y_j, B, Gy_j) > r(x_j, B, Gx_j)]$

TII. 1.3, Imp. and Taut. TII.3.3.0

47. TII. 3.3.00 affirms nearly the same thing as TII. 3.3. The difference is that now (provided the stimulus is sensorily received) we are predicting the advantage of y's tendency to respond over x's tendency to respond on the j^{th} trial in the disruption situation and not merely the animal's (y's) greater knowledge (or expectation). In short, in this theorem we have gone from learning as acquisition to learning as exhibited in performance. We have gotten Tolman's learner

into "action". This should, in part, satisfy Guthrie's complaint that Tolman leaves the learner "buried in thought". Furthermore it seems appropriate to point out that the notion of response tendency used here seems consonant with Tolman's "Principles of Performance" which was published in the Psychological Review in September of 1955.

TII. 3.4 $[(i_1^k)\ [B \neq D\ .\ \varphi \neq \chi\ d(x_i,\ \psi x_i) = d(y_i,\ \psi y_i) > 0$
$.\ d(x_i,\ \chi x_i,\ \psi x_i) > d(y_i,\ \varphi y_i,\ \psi y_i)\ .\ (G)R^e(x_i,\ y_i)\ .$
$(r(y_i,\ E,\ Gy_i) > 0\ .\ By_i) \supset \varphi y_i\ .\ \varphi y_i \supset \psi y_i\ .\ (r(y_i,\ D,\ Gy_i) > 0$
$.\ Dy_i) \supset \chi y_i\ .$
$\chi y_i \supset \psi y_i]\ .\ S_x^{1-h-k}\ [(E,\ D,\ G),\ (\varphi,\ \chi),\ (\psi)]] \supset$
$[e(x_k,\ Gx_k \supset (Ex_k \supset \psi x_k)) > e(y_k,\ Gy_k \supset (Ey_k \supset \psi y_k))]$
TII. 2.6 $[\frac{\psi}{E},\frac{B}{D}]$, Simp, TII. 1.2, TII. 1.3, Taut and Exp.

TII. 3.4.0 $[(i_1^k)[B \neq D\ .\ \varphi \neq \chi . d(x_i,\ \psi x_i) = d(y_i,\ \psi y_i) > 0\ .$
$d(x_i,\ \chi x_i,\ \psi x_i) > d(y_i,\ \varphi y_i,\ \psi y_i)\ .\ (G)R^e(x_i,\ y_i)\ .\ (r(y_i,\ E,\ Gy_i)$
$> 0\ .\ By_i) \supset \varphi y_i\ .$
$\varphi y_i \supset \psi y_i\ .\ (r(y_i,\ D,\ Gy_i) > 0\ .\ Dy_i) \supset \chi y_i\ .\ \chi y_i \supset \psi y_i]\ .$
$(i_1^h)\ [(r(x_i,\ E,\ Gx_i) > 0\ .\ Ex_i) \supset \varphi x_i\ .\ \varphi x_i \supset \psi x_i\ .$
$(r(x_i,\ D,\ Gx_i) > 0\ .\ Dx_i) \supset \chi x_i\ .\ \chi x_i \supset \psi x_i]\ .\ (i_{h+1}^k)\ [(r(x_i,$
$E,\ Gx_i) > 0\ .\ Ex_i) \supset \chi x_i\ .\ \chi x_i \supset \psi x_i .$
$r(x_i,\ D,\ Gx_i) > 0\ .\ Dx_i) \supset \varphi x_i\ .\ \varphi x_i \supset \psi x_i]] \supset SC$ TII. 3.4 and TII. 2.2

TII. 3.4.00 $SA \supset [d(x_k,\ Ex_k) > d(y_k,\ By_k]$ TII. 3.4.0 and TII. 2.3 $[\frac{E}{\varphi},\frac{\psi}{\chi}]$

TII. 3.4.000 $[(i_1^k)\ [E \neq D \cdot \varphi \neq \chi \cdot d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0 \cdot d(x_i, \chi x_i, \psi x_i) > d(y_i, \varphi y_i, \psi y_i) \cdot (G)R^e(x_i, y_i) \cdot (r(y_i, E, Gy_i) > 0 \cdot Ey_i) \supset \varphi y_i \cdot \varphi y_i \supset \psi y_i \cdot r(y_i, D, Gy_i) > 0 \supset \sim Dy_i \cdot \sim \chi y_i] \cdot (i_1^h)[(r(x_i, E, Gx_i) > 0 \cdot Ex_i) \supset \varphi x_i \cdot \varphi x_i \supset \psi x_i \cdot r(x_i, D, Gx_i) > 0 \supset \sim Dx_i \cdot \sim \chi x_i] \cdot (i_{h+1}^k)[r(x_i, E, Gx_i) > 0 \cdot Ex_i) \supset \chi x_i \cdot \chi x_i \supset \psi x_i \cdot r(x_i, D, Gx_i) > 0 \supset \sim Dx_i \cdot \sim \varphi x_i]] \supset d(x_k, Ex_k) > d(y_k, Ey_k)$ TII. 3.4.00 and *2.21

48. TII. 3.4.000 seems to express Tolman's meaning in the following passage on page 68 of Purposive Behaviorism.

> (b) If a relatively "good" goal-object be substituted during the course of learning for a relatively "poor" goal-object, the rat's performance shows a sudden improvement. (c) Conversely, if a "bad" goal-object be substituted during the course of learning for a "good" one, the animal's performance shows a sudden degeneration.

The use of the word "sudden" in the above passage is extremely vague. It has therefore been ignored in theorem TII. 3.4.000. In a moment, I shall make certain remarks concerning the above quotation in relation to the "latent learning" experiment of Blodgett's. The above quotation expresses two sides of the same coin. Hence a single theorem seems to adequately express both (b) and (c) in that quotation. Now where we understand "better goal-object" to mean, essentially, that one is more demanded than the other -in the theorem, this is symbolized as '$d(x_i, \chi x_i, \psi x_i) > d(y_i, \varphi y_i, \psi y_i)$'- and also interpret "improvement in learning" to mean, in this case, an advantage in demand for the "route" leading to the

substituted better goal-object -which is the consequent of the theorem TII. 3.4.000- we may read the above theorem as follows: If during the trial series 1 to k, B is not identical with D and φ is not identical with χ and x and y have the same demand strength (above 0) for the same type of goal-object ψ and χx is a better goal-object for x than is φy for y and (the stimulus) G is sensorily received by both x and y and y's response by B to Gy always leads to φy which leads to ψy and if during the trial series 1-h-k, χx is substituted for φx at h+1 relative to x's response by B to Gx, then on the k^{th} trial x's demand for Ex is greater than y's demand for Ey.

49. The Blodgett latent learning experiment is simply a version of this principle.[8] However, Tolman's neglect of the vagueness of term "sudden" does not make the Blodgett study a very good test of the latent learning issue. The present author's objection to the vagueness found in the above quotation must not be taken as a denial that the Blodgett version of the latent learning principle is deducible from TII. On the contrary, if and when Tolman makes clear the limits of sudden improvement (or degeneration), a ramification of TII. 3.4.000 suited to the new requirements will automatically permit a deduction of the Blodgett latent learning principle.

---

8. The Blodgett experiment is discussed on pages 48-50 of <u>Purposive Behaviorism</u>.

TII. 3.4.1 SA (TII. 3.4) $\supset [r(x_k, E, Gx_k) > r(y_k, E, Gy_k)]$

RN TII. 3.4 and TII. 2.5

TII. 3.4.10 Same as TII. 3.4.1 except that '$d(x_k, \chi x_i, \psi x_i) > d(y_i, \varphi y_i, \psi y_i)$' in TII. 3.4.1 is replaced by '$(\varphi x_i\ E\!> \chi x_i\ .\ \varphi y_i\ E\!> \chi y_i\ .\ \varphi y_i\ E\!> \chi x_i\ .\ \chi y_i \supset \psi y_i\ .\ d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0\ t(x_i, \chi x_i, \psi x_i) > 0)$' . TII. 3.4.1 and TII 2.8

TII. 3.4.100 $[(i_1^k)\ [B \neq D\ .\ \varphi \neq \chi\ .\ d(x_i, \psi x_i) = d(y_i, \psi y_i) > 0\ .\ \varphi x_i\ E\!> \chi x_i\ .\ \varphi y_i\ E\!> \chi y_i\ .\ \varphi y_i\ E\!> \chi x_i\ .\ (G)R^e (x_i, y_i)\ .\ (r(y_i, E, Gy_i) > 0\ .\ Ey_i) \supset \varphi y_i\ .\ \varphi y_i \supset \psi y_i\ .\ (r(y_i, D, Gy_i) > 0\ .\ Dy_i) \supset \chi y_i\ .\ \chi y_i \supset \psi y_i]\ .\ S_x^{1-h-k}[(E, D, G), (\varphi, \chi), (\psi)]] \supset [(i_1^k)(t(x_i, \chi x_i, \psi x_i) > 0 \supset [r(y_j, E, Gy_j) > r(x_j, E, Gx_j)\ .\ r(x_k, E, Gx_k) > r(y_k, E, Gy_k)]$ TII. 3.4.10, Simp, TII. 3.3.00, RN and Taut.

TII. 3.4.1000 SA $\supset [(i_1^k)(t(x_i, \chi x_i, \psi x_i) > 0) \equiv [r(y_j, E, Gy_j) > r(x_j, E, Gx_j)\ .\ r(x_k, E, Gx_k) > r(y_k, E, Gy_k)]$ TII. 3.4.100 and TII. 2.9

50. This theorem, TII. 3.4.1000, has the form of Carnap's bilateral reduction sentence. It may thus be construed as an "operational definition" of docility, that is, "teachability". It possesses certain interesting features. The only intervening variable except that of demand -which here may be taken as "drive" (that is, $(\psi)U(x, y)$ -ψ is an ultimate goal-object of x and y-) is teachability. All other terms are either in what the psychologists call a "data language", for example, "$\varphi y_i \supset \psi y_i$", or are in a

language which is immediately reducible to the data language, for example, "$(r(x_i, B, Gx_i) > 0 \,.\, Ex_i) \supset \varphi x_i$". Again it is, in part, a consequence of the law of least effort. Now it is true that Tolman couches docility in terms of efficiency. But the above operational definition -what Tolman calls an "objective definition"- finds much support in *Purposive Behaviorism*, especially in chapter IV. For it will be noticed that teachability, in the above theorem, is, in part, couched in terms of disruption -or more precisely, recovery from disruption.

> The doctrine we here contend for is, in short, that wherever a response *shows docility* relative to some end -- wherever a response is *ready to break out into trial and error* and (b) to select gradually, or suddenly, the more efficient of such trials and errors with respect to getting to that end, such a response expresses .... a purpose.

This passage on page 14 of *Purposive Behaviorism* is made a bit more clear if we substitute the expression "an organism" for the expression "a response" in all but the last occurrence of the latter expression in the quoted passage. The point I wish to make is that the italicized phrase in the above passage describes a condition which is the product of disruption in precisely the sense in which the latter term is explained in TII. 3.3. Now Tolman, in the *Glossary* of *Purposive Behaviorism*, by no means limits the selection (which takes place before or after disruption) of "better" goal-objects to merely the more efficient ones-- where we may assume that efficiency and effort are not reducible one to the other.

Hence, it seems perfectly proper to couch one operational definition of docility partly in terms of effort. Given these considerations, what TII. 3.4.10.0 amounts to is this: If the organism has a drive and is presented with certain goal-objects (paths) one of which involves the least amount of effort in its execution (in relation to the others) and at some time in the experiment the least effortful goal-object is substituted for a more effortful goal-object as a means to alleviating the drive state, then throughout the experiment the organism is docile (is teachable) with respect to the "better" (i.e., least effortful) goal-object if and only if at some time after substitution but before the end of the experiment there is disruption relative to the means to getting to the substituted "better" goal-object and later there is recovery in favor of a greater selectivity for the means leading to the substituted "better" goal-object.

51. This theorem seems to express Tolman's meaning in the following passage on page 74 of Purposive Behaviorism.

> And it would undoubtedly also be found (b) that, even when the new goal-object was at first distinguishable and did at first cause disruption such a disruption would after enough experiences disappear. That is, it would undoubtedly be found that the behavior was docile with respect to the new goal-object. The rats would never, perhaps, run as well for the new goal, if it were less desirable, as they did for the old, but the disruption in their behavior, qua disruption, would disappear.

TII. 3.5 $[(i_1^k)(z)\,[\varphi \neq \psi \,.\, D \neq G \,.\, J \neq L \,.\, d(z_i, Hz_i) = 0$
$\supset d(z_i, \chi z_i) = r > 0 \,.\, Ez_i \supset ((r(z_i, \varphi, Ez_i) > 0 \,.\, \varphi z_i) \supset$
$Dz_i,a) \,.\, Ez_i \supset ((r(z_i, \psi, Ez_i) > 0 \,.\, \psi z_i) \supset Gz_i,b) \,.\, (Ez_i \,.$
$\varphi z_i) \supset {\sim}Gz_i,b \,.\, (Ez_i \,.\, \psi z_i) \supset {\sim}Dz_i,a] \,.\, (i_1^h)\,[Ex_i \,.\, {\sim}Ey_i \,.$
$d(x_i, Mx_i) = d(y_i, My_i) = t \,.\, t = 0 \,.\, Ja \,.\, L_b \,.\, t^A(x_i, \varphi x_i,$
$\chi x_i) > 0] \,.\, (i_{h+1}^{h+2})\,[Ja \,.\, L_b]\,(i_{h+2}^{k})\,[Ja \,.\, L_b \,.\, Ex_i \,.$
$Ey_i \,.\, d(x_i, Mx_i) = d(y_i, My_i) = s > r \,.\, (Ja \,.\, Dx_i,a) \supset Mx_i \,.$
$(Ja \,.\, Dy_i,a) \supset My_i]] \supset [e(x_k, Ex_k \supset (\varphi x_k \supset Mx_k)) > e(y_k, Ey_k$
$\supset (\varphi y_k \supset My_k))]$ TII. 2.10 Simp, AN and Comm.

TII. 3.5 expresses the weak drive, irrelevant incentive version of latent learning. Let 'Hz' be 'z alleviates his thirst', 'χz' be 'z satisfies his curiosity', 'Ez' be 'z is in the starting box', 'φ' be 'taking the right route', 'Dz,a' be 'z gets to goal-box a', 'ψ' be 'taking the left route', 'Gz,b' be 'z gets to goal-box b', 'Ja' be 'a contains food', 'L_b' be 'b contains water' and 'Mx (and y)' be 'x (and y) alleviates his hunger'. TII. 3.5 when thus applied, in effect, reads: during k trials for every z if the fact that z does not demand that he alleviate his thirst implies that he demands that he satisfy his curiosity and if z is in the starting box then if he responds by taking the right route to the fact that he is in the starting box then he gets to goal-box a and if he is in the starting box then if he responds by taking the left route to the fact that he is in the starting box, then he gets to goal-box b and when he is in the starting box and

takes the right route then he does not get to goal-box b and when he is in the starting box and takes the left route then he does not get to goal-box a and during the first h trials x is put in the starting box and y isn't and x's demand that he alleviate his hunger is equal to y's that y alleviate his hunger which is equal to t which in turn is equal to 0 and goal-box a contains food and goal-box b contains water and x has been taught taking the right alley leads to satisfying his curiosity and during the trials h+1 to h+2 goal-box a contains food and goal-box b contains water but from trial h+2 to k both x and y are put in the starting box and their drive for alleviating hunger is greater than r which is greater than 0 and if a contains food and x gets to a then x alleviates his hunger and the same for y, <u>then</u> on the k^{th} trial x's expectation that if he (x) is in the starting box then if he takes the right route, he alleviates his hunger, is greater than y's that if he (y) is in the starting box then if he takes the right route, he alleviates his hunger.

52. I shall make only one comment here concerning the proof of the above theorem. The elements which were introduced into TII. 2.10 so that TII. 3.5 is the result are: in (i_1^h), 't = 0', 'Ja' and 'L_b'; in (i_{h+1}^{h+1+1})'L_b'; in (i_{h+1+1}^{k}), 'L_b'; and finally in $(i_1^h)(z)$, '$J \neq L$'.

TII. 3.5.1 $[(i_1^k)(z)[\varphi \neq \psi . E \neq D . J \neq L . d(z_i, \chi z_i) = r > 0 . Ez_i \supset [(r(z_i, \varphi, Ez_i) > 0 . \varphi z_i) \supset Dz_i,a . (r(z_i, \psi, Ez_i) > 0 . \psi z_i) \supset Gz_i,b] . (Ez_i . \varphi z_i) \supset \sim Gz_i,b . (Ez_i . \psi z_i) \supset \sim Dz_i,a] . (i_1^h)[Ex_i . \sim Ey_i . d(x_i, Mx_i) = d(y_i, My_i) = t < r . \sim Ja . \sim L_b . t^A(x_i, \varphi x_i, \chi x_i) > 0] . (i_{h+1}^{h+2})[Ja . Px_i,a . \sim Py_i,a . \sim L_b . \sim Ex_i . \sim Ey_i] . (i_{h+2}^k)[Ex_i . Ey_i . \sim Px_i,a . \sim Py_i,a . d(x_i, Mx_i) = d(y_i, My_i) = s > r . Ja . \sim L_b . (Ja . Dx_i,a) \supset Mx_i . (Ja . Dy_i,a) \supset My_i]] \supset [SC$ (TII. 2.10)] TII. 2.10, *2.02, Simp, Taut.

TII. 3.5.1 represents a modified version of the Buxton-Seward free exploration type of latent learning. Given the replacements for the variables listed in the discussion of the preceding theorem with the addition of 'x is put in goal-box a' for 'Px_i,a', this claim may be supported under the appropriate interpretation for TII. 3.5.1.

TII. 3.5.2 $[(i_1^k)(z)$ [same as in TII. 3.5.1] $. (i_1^h)[Ex_i . \sim Ey_i . d(x_i, Mx_i) = d(y_i, My_i) = t . t = 0 . t < r . d(x_i, Hx_i) = d(y_i, Hy_i) = s > r . Ja . \sim L_b . t^A(x_i, \varphi x_i, \chi x_i) > 0] . (i_{h+1}^{h+2})[Ja . \sim L_b] . (i_{h+2}^k)[Ex_i . Ey_i . d(x_i, Mx_i) = d(y_i, My_i) = s > r . d(x_i, Hx_i) = d(y_i, Hy_i) = 0 . Ja . \sim L_b . (Ja . Dx_i,a) \supset Mx_i . (Ja . Dy_i,a) \supset My_i]] \supset [SC]$ (TII. 2.10)] TII. 2.10, *2.02, RN, Simp and Taut.

TII. 3.5.2 represents the strong drive, irrelevant incentive version of latent learning. Replacement of the variables by the values given in the discussion of the pre-

ceding theorem TII. 3.5 will confirm this claim.

53. One final note: the latent learning principle should probably be much more general in character. I have decided on this more specific handling for two reasons. First, the theorems represent fairly closely existing types of experimental designs. Secondly, I wish to show that the latent learning studies, as represented by the various types of experimental design in contemporary psychology, are deducible from Tolman's system and, hence, constitute a prediction from that system.

CHAPTER V

RE-EXAMINATION OF THE PRESENT SYSTEM

1. This final chapter has two main objectives. First, there will be put forth some general remarks concerning the system presented in this essay and some speculation on what can be done in a future development of Tolman's system in terms of the present symbolism. Secondly, there will be a discussion of that in Tolman's system which cannot be formulated in terms of the present symbolism. Here, again, some speculation as to the future development of Tolman's system in terms of a different but perhaps more appropriate symbolism will be set forth. In short, the first part deals with the strong points of the present symbolism; the second, with its weaknesses.

PART I

The "Core" System and the "Courtesy" System

2. The system presented in TI and TII is what I shall call the "core" system of purposive behaviorism. By "core system", I mean the part of the "System" called "Purposive Behaviorism" which constitutes the fundamental basis - the rock-bottom - of that system. The "courtesy" system is that part of "Purposive Behaviorism" which consists in certain statements which are merely repetitions, using slightly different terminology, of statements in the

core system and also in statements which can be abstracted away by definition from the core system. Clearly the determination of the core system is somewhat arbitrary; but not completely so. The core system in this essay, that is, the system comprising TI and TII, is partly determined by what has been presumed to be fundamental in Tolman's thought, by conditions of generality, and, of course, by certain external considerations. For example, the present author talks about expectations because most psychologists talk about expectations rather than sign-gestalt expectations. It should be clear that the system presented in this essay is only a part of what has been called the core system. To use a current psychological cliché, this enterprise has been "programmatic" --but, it is hoped, a little less so than current examples (for example Tolman's own recent Principles of Performance) of psychological theorizing which are often, in part, justified by this term. For example, the notion of expectation is dealt with only in a very general way in the present system. It has not considered its various particular manifestations, namely, perceptual expectation, memory expectation, and inferential expectation. It is assumed that however one construes these concepts inside Tolman's theory of learning, statements about them will not falsify any of the laws of expectation in TI and TII. Again, in TII, there are no laws having to do with the frequency and recency of stimulation presentation in relation to expectations and demand, and so on. It should be understood that these laws were left out because of time and space considerations and not because they are not formulable in the present system.

3. The core system in this essay is largely concerned with the two intervening variables of expectation and demand. It is the opinion of the present author that from this core system it may be possible to abstract away that entire part of Tolman's system having to do with sign-gestalt expectations, sign-objects, sign-significate relations, etc. (which comprises about one-third of Purposive Behaviorism). This supposition is based largely on the discussion of the concepts of readiness, expectation, sign-gestalt-readiness, means-object, etc. in the Glossary of Purposive Behaviorism. For example, Tolman defines "sign-gestalt-readiness" as "the same as means-end-readiness". Given this knowledge, plus the convention that expectation and readiness are to be treated in the same way in the present essay, then, in the language of TI, we can define a sign-object as:

$$pS^n x =_{Df} [pR^e x \,.\, (\exists q)(e(x, p \supset q) > 0)],$$

that is, 'p is a sign-object of x' means 'p is sensorily received and some q is the end toward which x expects that p leads'. We can define the signified-object as:

$$qS^d x =_{Df} (\exists p)[pR^e x \,.\, e(x, p \supset q) > 0],$$

that is, 'q is a signified-object of x' means 'some p is sensorily received and is the means which x expects leads to q'. (These definitions represent a fair translation of Tolman's meaning of sign-object, signified object, and sign gestalt readiness. It will be noticed that in these examples, the notion on the right hand side of the definition (the definiens) is a notion in the core system.) Again, in the opinion of the present author, Tolman's list of intervening variables may be significantly reduced by definition in terms of demand and expectation. For example, consider the notion of "appetite" (or, later on, "cathexis") which is a member in the list of intervening variables explicitly enunciated by Tolman in his 1937 presidential address before the American Psychological Association, but only implicitly enunciated in <u>Purposive Behaviorism</u>. Consider the <u>Glossary</u> characterization of the term appetite. Tolman writes:

> An appetite arises out of a cyclically appearing, metabolically conditioned, initiating physiological state, or excitement (q.v.). And it consists in a resultant demand for a certain complementary type of physiological quiescence (q.v.) plus a set of partly innate, partly acquired, more or less definite, means-end-readinesses (q.v.) as to the types of means-objects and goal-objects to be had commerce-with (q.v.) in order to reach this demanded quiescence. Typical appetites are hunger and sex.

Now let '$xA^{t}p$' mean 'x has an appetite for (or with respect to) p'. Then in accordance with the above quotation, we can define "appetite" as follows:

$$xA^{t}q =_{Df} (\exists p)(qUx \,.\, e(x, p \supset q) > 0)$$

that is, 'x has an appetite for q' means 'x expects (or has a readiness) that some p leads to q where q is an ultimate goal-object'. Similarly, we might define "aversion" as follows. Let 'xA^Vq' mean 'x has an aversion for q'. Then we may write:

$$xA^Vq =_{Df} (\exists p)((\sim q)Ux \;.\; e(x, p \supset q) > 0)$$

Furthermore, it seems that concepts like "differentiation", "biases", "drive discrimination", etc. are capable of reduction by definition in terms of demand and expectation.[1]

4. It is one of the virtues of the present symbolism that it permits notions like "commerce-with" in its relations to things to be abstracted out, by definition, from the core system in II. The present system, therefore, presents a special set of laws which include certain laws pertaining to notions like "commerce with" as abstractable (and thus derivable) elements. For example, consider the notion of confirmation. Let 'p' mean 'x runs down alley A'. Now this proposition may be analyzed further in the following way. Let 'F' mean 'runs down alley A'. Then 'x runs down alley A' may be written as:

$$Fx \;.$$

This might be written in terms of confirmation as 'xC(Fx)'; 'xC(Fx)' is thus a substitution instance of 'xCp'. Again, more generally, 'xCp' might also be analyzed as 'xC(φx, y)' where 'φ' is a <u>predicate variable</u> and 'y' is an <u>argument</u>

---

1. To a certain extent, this opinion is also held by Meehl and MacCorquodale. cf. Modern Theories of Learning, Appleton-Century-Crafts, 1954, pp. 169-190.

variable taking "things" as values, for example, (alley A). Hence, 'xC(φ x,y)' can be taken as a substitution instance of 'xCp'; but, of course, the converse does not hold. Accordingly, we might easily define "commerce with" in terms of "confirmation". For example, where 'CW' means "commerce-with", we can write:

$$CW(x, y) =_{Df} (\exists\varphi) \,.\, xC(\varphi x, y).$$

that is, 'x has commerce-with y' means 'There is a φ such that x has confirmed that x stands in the relation φ to y'. If one substitutes 'alley A' for 'y', the above definition reads 'x has commerce-with alley A' means 'x has confirmed something with respect to alley A'. The same kind of thing can be done for the other primitive ideas. By this means, it would be possible to deduce a great number of laws about certain relations between organisms and things from TI. This discussion then demonstrates the greater power and generality of the core system in TI.

5. Other interesting courtesy system concepts which might be defined in terms of the ideas in TI are the notions of satisfaction, frustration and conflicting expectations. To say x is frustrated is to say he has conflicting demands. Hence, we might define frustration thusly. Let 'xF^rp' mean 'x is frustrated with respect to the state-of-affairs p'; then we get the definition:

$$xF^rp =_{Df} (\exists M)\ [d(x, p) = d(x, \sim p) = M > 0]$$

that is, 'x is frustrated with respect to p' means x demands both p and ~p at the same strength which is above 0. Now let '$xS^{A}p$' mean 'x is satisfied with respect to p'. Then:

$$xS^{A}p =_{Df} \sim xF^{r}p$$

that is, 'x is satisfied with respect to p' means 'x is not frustrated with respect to p'. Finally, conflicting expectations might be defined thusly. Let 'xEp' mean 'x has conflicting expectations relative to p'. Then:

$$xEp =_{Df} [e(x, p) > 0 \;.\; e(x, \sim p) > 0]$$

which means that he expects both p and not p. These definitions are interesting because they serve to make a connection - however small it may be - between Tolman's theory of learning and the area of personality. Their examination would be interesting and useful.

6. The examples of reduction and abstraction considered in paragraph 3 and 4 are to be found in the courtesy system. They reduce to (or are abstracted away from) the core system. There remains only the case of redundant statements in the courtesy system. I will now provide examples of this case. Consider what Tolman calls a "first-order drive". On page 28 of Purposive Behaviorism, he writes:

> It is such demanded physiological states of quiescence and disturbance which constitute the final goal-objects which the rat, and all other animals, are to be conceived as persisting to or from.

It is clear from this statement that a demand for an ultimate goal-object, for example, hunger-satiation, is a first-order drive. Let '$x1p$' mean 'x has a first order drive for p'. Then we may write the following definition:

$$x1p =_{Df} pUx \;.$$

But clearly the notion of first order drive is redundant. Or consider the case of goal-object. Let 'pGx' mean 'p is a goal-object of x', then we may write:

$$pGx =_{Df} d(x, p) \neq 0.$$

Here again, we have a good example of redundancy.[2]

7. The courtesy system is not useless. Undoubtedly it has certain psychological advantages in the sense that it perhaps conveys to the reader a more intuitive grasp of what the system as a whole is about. Indeed, this is part of the reason why it is called a courtesy system. But from a logical point of view, it is unnecessary. This discussion of the core system and the courtesy system of purposive behaviorism indicates one of the main advantages of the system presented in this essay, namely, that it is more economical than the original system in Purposive Behaviorism. It is more economical in the sense that it attempts to present only the indispensable, rock-bottom elements of purposive behaviorism. Surely progress toward this end has

---

2. A concept in the courtesy system based on the ideas in TII is "better goal-object". Let '$(\varphi, \psi)(x, y)$' mean 'φ is a better goal-object for x than ψ is for y'. Then we have the definition:

$$(\varphi, \psi)(x, y) =_{Df} (\exists\chi)[d(x, \varphi x, \chi x) > d(y, \psi y, \chi y)]$$

This is, 'φ is a better goal-object for x than ψ is for y' means 'there is a χ such that x's demand for φx as a means to χx is greater than y's demand for ψx as a means to χx'. From this definition, substituting x for y, we would obtain the concept that 'φ is a better goal-object than ψ for x'. That is, we could get the biconditional formula:

$$(\varphi, \psi)x \equiv (\exists\chi)[d(x, \varphi x, \chi x) > d(x, \psi x, \chi x)].$$

been achieved to some degree. We have thus a concrete example of one of the benefits which results from formal organization. Secondly, the system in this essay is simpler in the sense that it involves fewer and less complex "primitive" ideas than are found in the original "system" in Purposive Behaviorism. But let us not misunderstand this matter. It is not being claimed that a further reduction cannot be made in the list of primitive ideas (by definitions) nor that the postulate set is not susceptible of reduction to a smaller set. Again, it is a most alluring possibility that many of the postulates in TI may be deducible from the set found in TII. This speculation is based on the fact that, except for a more elaborate form, some of the postulates of TI are quite similar to certain postulates in TII. (Of course, this would mean abandonment of the propositional variable 'p' in TI in favor of the propositional function variable 'φx' because all of the postulates and theorems of TII are couched in terms of propositional function variables.) To recapitulate, the economy and simplicity of the present system is to be seen only in comparison with the original system in Purposive Behaviorism. It was not the purpose of this essay to find the smallest postulate set or the smallest possible list of primitive ideas; that is, it did not strive for logical economy and logical simplicity.

Further Advantages of the Present System

8. Meehl and MacCorquodale write:[3]

> Secondly, there is no evidence of formal derivation in Tolman's discussion of [the classes of experiments which purport to be support for Tolman's theory]. It does seem as if these results might flow as consequences from "some such theory",...

Again, in connection with the preliminary discussion of latent learning as offering support for Tolman's view, they write:[4]

> ....one could note that the actual "derivation" of latent learning phenomena <u>from</u> Tolman's theory here is simply not given.

9. The present system includes derivation of at least some versions of latent learning <u>from</u> Tolman's theory. Possibly, the critical comment here will be that something has been put into Tolman's system which was not there originally and which allows one to get, as theorems, certain kinds of latent learning. This objection is fostered by the vague use of the expression "the system" or "the theory". (In the above quotations Meehl and MacCorquodale are not at all clear as to what they mean by "Tolman's theory"). The expression "the system" may be used to refer to a certain explicitly stated set of statements -and no more. However, it may also be used in a way which is a bit more indefinite but does not lead to any inconsistency with the first use.

---

3. Op. cit., Modern Theories of Learning, p. 196.
4. Ibid., p. 195.

That is, it may refer to a certain explicitly stated set of statements plus any others which might be added but which do not serve to make the "system" inconsistent. The present objection construes the expression "the system" in the first way. This is an unwise use. For it prohibits the growth of a system. Again, one can seriously wonder whether there is, in fact, anything which can be called "the system" in this first sense of the expression. For, it is certainly true that psychological theorists, for example, Hull, are constantly expanding their theories by the addition of new postulates. However, suppose that the first sense of the expression "the system" has justification. I wish now to argue that even in this narrower sense of the expression, "the system" of Purposive Behaviorism permits the derivation of the latent learning theorems found in TII. This would confirm, in the strongest possible way, Meehl's and MacCorquodale's feeling that the variations of latent learning do "flow as consequences" from Tolman's theory -or, at least, from a close replica of it.

10. The final ground for latent learning lies in the discussion of transfer and docility found on pages 32 to 34 of Purposive Behaviorism.

Richter has shown that general exploratory activity of the rat occurs in cycles corresponding to the cycles in the contractive activity of the stomach. And, further, if there be attached to the rat's main living cage a small cage in which there is food, he found that it is at the height of each activity cycle that the rat passes into the food-cage and eats, after which the animal returns to the living cage, cleans himself and then subsides .....

To quote:

> Thus we see that the small contractions give rise to the diffuse activity in the large cage. The animal seems at first simply to be annoyed and becomes more and more restless as the contractions grow larger, until the 'main' contractions set in and the general discomfort becomes centralized in the hunger sensation. This stimulus dominates the behavior of the organism and it enters the food-box to eat. When its appetite has been satisfied, it passes into a period of quiescence which lasts until the stomach has become empty and the contractions have started up again.

It appears, in short, that it is the hungry, or satiation-demanding, rat who is the exploration-demanding rat. And further it also appears that at the height of hunger, the exploratoriness is specifically directed toward food.

The further point we now wish to make is that such exploratoriness will prove docile relative to the actual finding of food. We want to show that an animal's exploratoriness embodies a means-end-readiness, judgmental in character, to the effect that certain types of exploration (exploratory object) are more likely to lead to food than are others. And, in fact, general evidence of this is to be seen, at once, in certain general findings as regards maze-adapation. A "naive" rat, when first run in a maze, is quite as likely to try to push through impossible crevices or to run upside down on the wire cover as he is to run in the alleys proper. A "maze-wise" rat, on the other hand, has become ready for alley-explorations only, i.e., for those general types of exploration which he has actually found tend to lead to food.

Or again it is to be observed that if a rat has learned in a given simple maze always to take, say, a right-turn to a T, he will tend to prefer such a right-turn when transferred to a second maze which presents another T under somewhat, but by no means exactly, similar conditions to those of the first maze. A rat will "abstract" the goodness of right-turning from his first maze. This has recently been demonstrated specifically and very prettily by Gengerelli, who reports that the group of rats who most clearly carried over the "generalized habit", or what we are calling the means-end-readiness, from the one situation to the other were not acting in any reflex fashion. He describes one of the most striking cases of such "transfer" as follows:

> The rats by this time did not run the maze as if it were a stereotyped habit. The continual changing of maze patterns from day to day had caused them to adopt a more circumspect poise in their running. By this time there was very little, if any, bumping of noses at bifurcations and elbows. The animals had become more exploratory in their running attitude. They invariably slowed up or paused as they approached anything that looked

like a turn in the maze.

> Practically all of the animals, therefore, approached the bifurcation in the maze pattern used in this experiment slowly and deliberately. There was, accordingly, some hesitation at the cross-roads, and a great deal of looking from side to side before the choice was made.

And yet they carried the right (or the left) turning-readiness over from the previous training.

Indeed it is obvious that according to us all so-called "transfer" experiments would be evidence of the formation and carrying over of specific (judgmental) means-end-readinesses. Thus Vincent's demonstrations of the formation and transfer of the choice of white alleys rather than black from discrimination box to "white-black" maze and vice versa would similarly be a demonstration of an acquired and truly docile menas-end-readiness to the effect that white alleys are better than black for leading on to food.

11. What has been said here might be summed up, in part, in the following statement; any expectation acquired under one set of circumstances or drive is potentially usable in another set of circumstances or drive (even though this latter set of circumstances may involve dissimilar conditions) provided that the learner is docile. This statement taken in conjunction with the well-known Tolmanian view (a proviso emphasized again and again in <u>Purposive Behaviorism</u>) that an expectation may be acquired in the absence of reinforcement- a condition which is implicit in the postulate set in this essay- constitutes the theoretical basis for the latent learning principle. These conditions are summed up in postulate III. 2.10. The only problem, so far as the present author was concerned, was to construct a sym-

bolic translation of the above considerations; we did not have to *invent* the latent learning principle.[5]

12. Let us return to the question of the use of the expression "the system" (or "the theory"). Let us construe "the system" in the wider sense. This use of "the system" is analogous to construing "the corporation" as a class of individuals whose membership though it increases still does not change the meaning of the "the corporation". There are good reasons for this decision. For example, consider the possibility of deriving "new" principles or laws from the present system in this essay. In a personal communication, Professor Tolman writes:

> Actually, however, in my own thinking, I have merely used my words as springboards for covert phenomenologizing which led me to believe that such and such further experiments would be interesting. And I have a suspicion that no new experiments can really be derived from my system in a formal way that are not already implied in the initial definitions.
>
> I tend now to believe that the future of psychology lies in the finding of independent empirical facts (probably physiological) which will then lead to new implications. I do not believe in short that my system is a system in the true logical sense from which anything very new or interesting can be derived.

---

5. The present author does not claim that the latent learning principle, TII. 2.10, and its corollaries, as presented in III, are *beyond* criticism; nor that they may not be false. These are matters for the experimenter to decide - not the logician. It is claimed only that the versions of latent learning which are formulable by means of the present symbolism can be put in postulate-theorem order in Tolman's system; that is, it is claimed that the versions of latent learning as categorized by Thistlethwaite are deducible as theorems in Tolman's system.

13. First, Professor Tolman's use of the expression "my system" is ambiguous in the way in which this particular definite description has been shown to be ambiguous above. Let us consider his comment in relation to the narrower view of the expression "my system".

14. Concerning the "suspicion that no new experiments can really be derived from, [Tolman's system] in a formal way that are not already implied in the initial definitions", the following point can be made. Tolman is correct - in a certain sense. Logically speaking, one cannot derive "experimental" conclusions which are not implied by the definitions and the postulates. But notice this does not mean that <u>all</u> of these <u>implications</u> are known - nor that they have yet been teased out of the system. Hence, there is a very good chance that there are "new" experimental conclusions to be derived from Tolman's system. But now let us assume that Tolman's system - say, as it stands up until 1954 - has exhausted all of its implications for new laws. There is, nevertheless, another way in which new laws may be discovered. The system can be extended by the addition of new postulates, for example, those resulting from the "finding of independent empirical facts (probably physiological)" and so on, and new experimental conclusions can thus be deduced. And, in fact, Tolman does this very thing in his paper in the <u>Psychological Review</u>, Vol. 62, No. 5 - <u>The Principles of Performance</u>. On page 319, using almost the same termin-

ology as a pre-1955 Tolmanite with the addition of the notion of "performance vector", Tolman writes:

> The greater the valence of the expected food, the greater the food need-push, and the greater the expectancy that the food will result, the greater the magnitude of the performance vector toward actually pressing the lever.

In effect, this is a new postulate. That is, in the opinion of the present author, it is independent of the other postulates in TI and TII - because of the idea of "performance vector". In other words, neither it nor its negate is deducible from the postulates of TI and TII. And certainly its relationship with the other laws of the pre-1955 system have yet to be discovered. Hence the possibility of new laws and also new experimental issues. This behavior is appropriate to the wider use of the expression "the system".[6] Indeed, it seems to betray the fact that Tolman himself implicitly construes the term "my system" in the wider sense. If this is true, then the second sentence in the cited communication above, is a bit misleading. It is on such grounds as these that the present writer defends the wider use of the term "the system", that is, the use which permits one to construe "the system" as embodying a consistent set of postulates, stated and <u>unstated</u>, but which employ the terms expectation, demand,

---

6. The use of the term "the system" in this wider sense does have pitfalls. For example, imagine two people to supplement a shared set of postulates with one postulate each, but that their supplements are mutually inconsistent. Which person has "the system"?

and so on, in a manner consonant with Tolman's own use. System building is a dynamic process. The system of Newtonian physics as taught in contemporary textbooks is not merely the physics of Isaac Newton.

EXTENSION OF THE PRESENT SYSTEM

15. We conclude the first part of this final chapter by suggesting certain ways in which the present system (in the narrow sense of the expression "the system") may be extended. The concern here is to show what is possible in this line within the context of the adopted mode of symbolism. (The present suggestions are more concerned with TII than with TI. Previous considerations in this chapter were largely concerned with TI.) The first problem to be dealt with concerns the use of the time argument _i_ in the formulae of TII. Here the concern is with an elaboration of the present symbolism as a way of extending the system.

16. Consider the series $(i_1^h) \ldots (i_{h+1}^k)$ where 1 designates the first number of the series and k the end of the series. In accordance with the conventions adopted in TII, the above series indicates a series of trials. 'h' and 'k' are variables. They designate respectively a point in the series (i_1^k) and the end of that series. The length of the series is undetermined. Hence, if the series is 20 trials long, k = 20. But even then, we do not know what value h is; 15 trials, the 1st trial, and so on.

Again, h+1 only indicates the first trial after h whatever h may be. The present symbolism does not say what a trial is; that is, whether a trial is 10 minutes long or 15 minutes long; 2 days or 3 days long, and so on. This latter problem is more empirical than logical since it involves a problem of measurement. That is, it involves a question of assigning numbers to certain predetermined units of time called trials. Again, a given trial may, (in terms of time) be longer or shorter than another - especially where a trial is determined by completion of a certain activity. Hence, the present symbolism does not reflect differences in trial times and thereby fails to distinguish between longer and shorter trials.

17. Some of the above postulates and theorems in TII might be false because, in the context of the present rough dating apparatus, we may be able to infer certain false theorems. For example, suppose the difference in effort between two actions is very slight. One would expect that in certain circumstances, more than one "trial" of training would be required to enable the learner to build up an expectation of such differences. The postulate TII. 2.4 does not allow for such circumstances. In terms of the present handling of time, it would be possible to deduce from that postulate that the learner in one trial could have built up the appropriate expectations necessary to distinguish between the effortfulness of the two activities. This seems quite unlikely. Hence, it is important

to know what 'h' and 'k' are in any trial series. But again, these are matters which are more empirical than logical.

18. The handling of the time series in this essay does <u>not allow</u> for the beginning or the ending of a given trial (assuming that this is our unit of measure, that is, each positive whole number designates one trial). And, of course, this is a very important problem. To illustrate, consider the case of sudden improvement after substitution of a "better" goal object. Our units are again trials (however these may be determined). Suppose that Tolman claims, in a given situation, that improvement is sudden when it begins to appear exactly 2 trials after substitution, that is, say, after h+1. The present symbolism cannot record <u>this</u> fact: Improvement will "<u>begin to appear</u> exactly 2 trials after substitution". All the present symbolism records is that improvement occurs, in this case, at h+3. It does not record possible increasing improvement over a certain span of time, say from h+1 to h+6; it does not tell us whether h+3 is the beginning, middle, or end of the trial or series of trials in which improvement occurs. This problem is more a logical problem than a problem of measurement. A more detailed handling of the time series could eliminate this problem <u>within</u> the context of the present symbolism. This, of course, raises the question of over-lapping time series. These matters might be dealt with after the technique of J. H. Woodger

in his monograph on the "Technique of Theory Construction" in "The International Encyclopedia of Unified Science" series.

19. However, the most serious difficulty, with the present symbolism is its inability to deal with changes happening within the course of a trial, or a given series of trials. For example, it might be important to know that a rat turned his head left and right in the course of a trial, for example, as in the VTE experiments of Muenzinger. Again, it might be important to know how fast an animal moved to a specific part of a maze at a specific time in a given trial, or in a given series of trials, to know whether the experimenter, say, puts in (or takes out) hurdles, mirrors, or the like during a given trial or a given series of trials, and so on. These circumstances cannot be recorded in the present symbolism for the simple reason that there is no way of expressing the situation "at such and such a time, in or during trial k so and so happened". However, there is no reason why the present symbolism cannot be extended so that situations of the sort just described can be expressed. Summing up, those problems having to do with merely the <u>logical</u> character of time series might be handled, within the <u>context</u> of the present symbolism, by a more detailed account of time. It is not being suggested that the problems just cited in relation to the matter of time series are the only ones. But they are representative and are

the kind of problems which most immediately come to mind when one examines the present system. Hence, it is concluded that a more detailed analysis of time will result in a certain sense, in an extension of the present system; that is, in the sense that it will "correct" the present formulae and make certain implicit assumptions explicit.

20. It will be remembered that the convention was adopted (in TII) that all the laws involving the concept of least effort assumed that the conditions of valence and efficiency were held constant. Clearly another way of extending the present system (within the context of the present symbolism) is to develop certain laws about the interrelations between effort, efficiency and valence. We speculate here only on a possible relationship between effort and efficiency.

21. Tolman is very cautious in Purposive Behaviorism about assimilating efficiency to effort as has been done by certain psychologists. I think his caution is well founded, because there is an essential difference between effort and efficiency. The latter can only be determined in relation to a given end or goal; determination of the former requires no such restriction. Let me illustrate. Consider a maze with two alleys one longer and more circuitous than the other. The longer we shall call "L"; the shorter 'R'. R is a straight alley to a goal box. Both alleys lead to the same goal-box. Now it seems

plausible to say that running down L involves more effort than running down R because running down L requires a greater expenditure of energy. Now consider the case where animals are put in such a maze and move about under the impulsion of curiosity. It seems plausible to claim that the longer alley, L, is more efficient as a means to satisfying curiosity, that is, curiosity satiation, than is R. But suppose we consider L and R in relation to the hunger drive. Then the shorter alley R is more efficient than L as a means to hunger alleviation. Thus we see that in one case the more effortful alley L is more efficient than the less effortful alley R, but, in the other case having to do with hunger, the more effortful alley L is less efficient than the less effortful alley R.

22. Assuming that the distinction between effort and efficiency is legitimate, we are in a position to lay down certain laws concerning their interrelations. Tolman seems to suggest in Purposive Behaviorism, that given the two conditions of efficiency and effort (other conditions being equal) efficiency is dominant. As a result we may lay down the law that of any two actions φx and ψx where φx is more effortful then ψx but ψx is more efficient than φx, as a means to satisfying a given drive, then the organism will respond more frequently by ψ than he will by φ. Indeed, we might tentatively propose the following postulate (which has the form of an

operational definition) for efficiency. Let '$EF\succ$' mean 'more efficient than'.

$$(i_1^k)\ [d(x_i, \chi x_i) > 0 \ .\ (Ex_i \ .\ \varphi x_i) \supset \chi x_i \ .\ (Ex_i \ .\ \psi x_i) \supset \chi x_i \ .\ \sim F\ (x_i, \varphi x_i, \chi x_i)] \supset [(i_1^k)(\psi x_i\ EF\succ\ \varphi x_i, \chi x_i) \equiv r(x_k, \psi, Ex_k) > r(x_k, \varphi, Ex_k)]$$

This proposition reads: during each of k trials if x demands that χx and E and φ are available to x as means to χ and E and ψ are available to x as means to χ and *it is false that x is fixated on φx as a means to χx*, then throughout each of k trials ψx is more efficient than φx as a means to χx if and only if on the k^{th} trial x's tendency to respond by ψ to Ex is greater than his tendency to respond by φ to Ex. The italicized part in the above translation emphasizes that efficiency is only operative when the animal is not fixated on the poorer or what proves to be the less efficient action. For, as fixation is defined, in chapter II of this essay there is a rigid condition in the organism's behavior in the sense that there is no confirmation of possible alternatives; --which amounts to saying that there are no alternatives, better or poorer, for the fixated animal.

23. The relationship between efficiency and effort discussed at the beginning of the last paragraph can be deduced easily from the above postulate. Hence we get the following proposition:

$$(i_1^k)\ [d(x_i, \chi x_i) > 0 \ .\ (Ex_i \ .\ \varphi x_i) \supset \chi x_i \ .\ (Ex_i \ .\ \psi x_i) \supset \chi x_i \ .$$

$$\sim F(x_i, \varphi x_i, \chi x_i) \cdot \varphi x_i \overset{E}{>} \psi x_i \quad \cdot \psi x_i \overset{EF}{>} \varphi x_i, \chi x_i] \supset$$
$$r(x_k, \psi, Ex_k) > r(y_k, \varphi, Ey_k)$$

24. With respect to the above described situation involving alley L and alley R we might try to speculate on the implications of the concept of efficiency for latent learning. The supposition above is that when the organism is shifted from the exploration drive to the hunger drive there is a corresponding shift in efficiency factor, that is, the formerly less efficient alley R as a means to alleviation of curiosity becomes the more efficient alley under the hunger drive. We might use this hint in the confirmation of latent learning. For example, the latent learning principle claims that during the exploration period the organism is allowed to build up "knowledge" of the various routes to the goal-box. Then when the food drive is introduced we should expect, in the above situation, provided the organism is docile, that the organism, in accordance with the shift in efficiency to alley R, will pick that alley R more frequently than he would alley L. So far we have not differed too much from the customary experimental latent learning situation. But now suppose we block, at some point in alley R, the organism from going to food. Then it would follow from the avoidance principles in TI and the efficiency hypothesis presented above, that the organism would immediately stop running down alley R and pick the now more efficient alley L (which also is now the less effortful) as a means to

food.[7] And thus we have an example of how the present system may be extended in this latter sense. Finally, it might be noted that, in the same vein, Tolman's remarks on avoidance in his recent paper "The Principles of Performance" might be used to advantage in the extension of the present system.

PART II

Inadequacies of the Present Symbolism

25. This section has to do with the weaknesses of the present symbolism. It is directed to the question of what cannot be expressed in the context of the present symbolism. In general, the appraisal concerns itself with the notion of material or truth-functional implication which is expressed by the symbol '⊃' in this system. The main point to be made is this: though the logic of material implication -which philosophers call an extensional logic- is sufficient to express a great deal of what Tolman says in his system, it is too weak a relation to capture the full meaning of certain concepts. For example it is too weak a relation to capture the full meaning of means-end-expectation.

---

7. It is important to note that in the usual simple T maze type of apparatus used in the studies of latent learning, due to equalized training procedures, no effort factor is built up.

26. We have had a prelude to these difficulties in the preceding pages of this essay. The two most dramatic cases were, respectively, (1) the demonstration of the failure of the *principia* law

$$[(p \cdot q) \supset r] \supset [p \supset (q \supset r)]$$

when placed within the context of the propositional attitude of expectation, that is,

$$e(x, (p \cdot q) \supset r) > 0 \equiv e(x, p \supset (q \supset r)) > 0$$

does not hold; (2) the demonstration that the relationship between 'φx' and 'ψx' in means-end availability was too weak. The first case failed, that is

$$e(x, (p \cdot q) \supset r) > 0 \equiv e(x, p \supset (q \supset r)) > 0$$

failed, because the truth functional "if, then" does not capture a certain consecutivity factor in the conception means-end expectation. In case of means-end availability, we found that φx is a means to ψx was better described as 'The making true of φx is sufficient to bring about the realizability of ψx' rather than as merely 'If φx is true then so is ψx', that is, as merely a truth functional "if, then" relationship between φx and ψx.

27. The ensuing discussion is limited to the failures of the material "if, then" and the remedy thereof in the case of expectation. Reference will be made to the case of means-end availability as a means of comparison in certain instances. We begin with a discussion of the terms to the relation '$\supset$' within the context of the propositional attitude of expectation.

Expectation Re-examined

28. Consider the paradigm:

$$(1) \quad e(x, p \supset q) > 0$$

This may be read: the strength of x's expectation (readiness, anticipation, belief, knowledge) that if p then q is greater than 0. Suppose we substitute 'x runs down alley L' for 'p' and 'x eats food' for 'q'. Then (1) reads: x expects that if he runs down alley L is true then 'he eats food' is true. This example suggests a rather strained interpretation of means-end expectation. When an organism expects a certain state of affairs as a <u>means</u> to another state-of-affairs, he is not, in every case, expecting that if the first state of affairs is true then the second is also true. Rather he expects that if a state of affairs 'p' is true then it is possible that 'q'. For example, as I sit at my desk I expect that if I walk to the door, then I <u>can</u> open it.

29. Consider the following postulate which is in TI:

$$(2) \quad (xCp \,.\, e(x, p \supset q) > 0) \supset e(x, q) > 0$$

(2) says that (with appropriate substitutions), if x has confirmed that x runs down the alley L and he expects that if he runs down alley L then he <u>can</u> get to food, then he expects that he <u>can</u> get to food. Notice here that the word <u>can</u> makes all the difference. It is a translation of the "it is possible" used a few sentences back.

30. The use of the expression "it is possible" in the preceding discussion is perhaps unfortunate. To the logician it has a different meaning than the meaning ascribed to it in the preceding paragraph. "It is possible" means simply "it is not incompatible (or it is not inconsistent)" in contemporary logic. In its formal employment it is associated with those logicial systems which are called "modal" --and sometimes "intensional". After C. I. Lewis, it is usually represented by the diamond '$\Diamond$'. In the above discussion, we do <u>not</u> mean by "it is possible", "it is not inconsistent (or incompatible)". Rather by 'x expects that if p then it is possible that q', what is meant is that x expects that if p then he can (or is able to) make q true, or to put it another way, that x expects that if p is true then he can (or is able to) realize (or actualize) q --to use Carnap's phrase. In short, formula (1) can be read: x expects that if p is realized then q can be realized, that is, is realizable.[8]

---

8. It may seem arbitrary to delimit 'p'in 'e(x, p ⊃ q) > 0' to being a realized predicate. For example, this expression might with equal justification be read: x expects that if he is able to (or can) realize p then he is able to (or can) realize q. However, it will be remembered, that in the response postulate TI. 2.6 the expression 'pR^ex' occurred. Since the only values of p are realized states-of-affairs, it was accepted as a tacit substitution rule in this system that the values of the first member in a means-end expectation are realized states-of-affairs. Indeed, this comports with our ordinary notion of "stimulus situation" as an actual state-of-affairs (cf. footnote 5 on page 134 of this essay).

31. The problem of realizable states-of-affairs, that is, of states-of-affairs which can be made true, as values of the propositional variables also carries over to propositional functions. The postulate II. 2.6 -the response postulate- is expressed partly in terms of propositional functions. It will be remembered that there was adopted tacitly a substitution rule to the effect that the only values which the predicate 'φx' (in the part of that postulate which read '$e(x, p \supset (\varphi x \supset q)) > 0$') could take were realizable predicates. This re-examination requires that we give a more full analysis of expectation than has been given. For example, let '$K(z, \varphi)$' mean 'z can make φ true (or z is able to do φ)'. Then, in terms of the above appraisal, we might insert '$K(z, \varphi)$' in formula (1) thus:

$$(3) \quad e(x, p \supset K(z, \varphi)) > 0,$$

that is, x's expectation that if p then z is able to do φ is greater than 0. To illustrate: let 'p' be 'x brings apples to z' and 'φ' be 'make an apple pie'. Then (3) reads: x expects that if he brings apples to z then z can make an apple pie. Again by substituting 'x' for 'z' in (3) we get the typical expectation situation as viewed from the experimenter's position. Let 'p' be 'the white card is present before x' and φ 'jumps to the white card'. Then (3), with substitutions, reads: x expects that if the white card is present before him then he can jump to the white card.[9]

9. On these same matters see also paragraphs 18-21 in the discussion of means-end-availability in chapter IV of this essay.

32. Having dealt with the terms to the relation '⊃' within the context of expectation, we may now direct our attention to the relation itself. The translations of (1), (2) and (3) have all been put in the indicative form. However, these translations are a bit misleading in the sense that they do not capture a certain condition of necessity obtaining between 'p' and 'q' (or between 'p' and 'K(z, φ)'). Means-end expectations are often expressive of a more hypothetical or provisional character in the meaning of "is a means to". Let us consider the interplay of means-end-relationships in the rat who "debates" with himself at the choice point in a maze. The experimental phenomena which are the basis of this account are the so-called Vicarious Trial-and-Error phenomena discovered by Muenzinger and the "hypothesis" phenomena discovered by Krech(evsky). Again, support may be found in Tolman's account of inferential expectations (Cf., pages 97 and 140-141 of Purposive Behaviorism). This condition of necessity is much better expressed in the subjunctive rather than in the indicative. That is, when we say that an organism expects that something is a means to something else we might better translate this in the following way: x expects that if he were to run down alley L, then he would be able to get to food. Thus what was meant by saying that the rat was "debating" with himself at the choice point is simply that the rat is engaging in an expectation of this provisional sort. The

point is that subjunctive conditionals (such as that which comes after the word "expects" in the rephrased account of expectation above) are <u>not</u> truth functional.[10] Indeed, it has been suggested that there is a certain necessity factor involved in subjunctive conditionals which prohibits that conditional from being truth functional. This necessity factor is not the necessity which contrasts with possibility in the logical sense of possibility discussed above. What is intended here might better be called "natural necessity". As a result, we might again rephrase our notion of expectation; x expects that p naturally necessitates q (or p naturally necessitates K(z, φ)). Take our "apple pie" translation above. That situation is now rephrased as x expects that x's bringing home apples naturally necessitates that z can make an apple pie. That is, 'x's bringing home apples' contributes to his wife's being able to make an apple pie. This is just another meaning of "if, then". Quite generally, the revised version of expectation might read: x expects that p naturally necessitates the realizability of q. The remainder of this essay is concerned with a discussion of the concept of natural necessity and its relationship to expectation.

---

10. Quine, W. V., <u>Methods of Logic</u>, Holt, 1950, pp. 12-18.

The Nature of Natural Necessity

33. Consider the propositions

$$(4) \quad e(x, (p \,.\, q) \supset r) > 0 \equiv e(x, p \supset (q \supset r)) > 0$$

and

$$(5) \quad e(x, \sim p \supset (p \supset q)) > 0.$$

(4) and (5) are sometimes false. (4) is sometimes false because it does not capture a certain consecutivity characteristic of means-end expectation. But it is also false because it overlooks a certain stronger view of the conjunction 'p . q' in the antecedent of (4) than is presented by the ordinary truth functional conjunction 'p . q'. For example, substitute 'x lifts his paw' for 'p', 'x leans against the lever' for 'q' and 'x gets a food pellet' for 'r'. Under these substitutions, the antecedent of (4) reads: x expects that when he lifts his paw and leans against the lever then he gets to food. Suppose that the experiment is so rigged that the animal must do both actions in a certain consecutive manner before he does get the pellet of food. Under these conditions, the consequent of (4) under the above substitutions could be false. That is "x expects that when 'he lifts his paw' is true then it is also true that 'x leans against the bar' implies that 'x gets food' is true", could be false. The point is this: x's expectation of the conjunction of 'p' and 'q' implying 'r' is an expectation that the conjunction of 'p' and 'q'contributes to the bringing about of 'r' - 'r' is the result of the con-

junction of 'p' and 'q'. Indeed, it may be put more strongly; x expects that the world being what it is, 'p' and 'q' are sufficient to the bringing about of 'r'. The consequent of (4), so interpreted, fails precisely at this point. For it allows (1) that 'p' is sufficient to bring about the conditional '$q \supset r$' - which could be false - and (2) granting 'p' is sufficient to bring about '$q \supset r$', that 'q' is sufficient to bring about 'r' - which seems highly dubious in relation to x's expectation that it is the conjunction of 'p' and 'q', and not either alone, which is sufficient to the bringing about of 'r'. This characteristic of means-end expectations will be called the "efficacity" characteristic of expectation.

34. It is interesting to note that the same kind of relation obtains in the idea of means-end availability. This was recognized when we rewrote this relation as 'pMq' which means 'the world being what it is, the occurrence p is sufficient to bring about the occurrence of q'.[11] Our argument in the case of means-end availability proceeds as follows: 'When we affirm that certain states of affairs are available as means to others, are we merely saying that when a given state-of-affairs is true so is another given state-of-affairs?' The answer must be no. We are reluctant to claim that something is available as a means to something else unless we have perceived that in the past (or realize that in the present) a given state-of-affairs - (perhaps, most commonly, an action) leads to or contributes to the bring-

11. Cf. Chapter IV, paragraph 20.

ing about of an end. For example, the following two propositions are both true: Lambert scratched his ankle with his left hand at 9:15 on August 14, 1956, and Lambert finished writing a sentence with his right hand at 9:15 on August 14, 1956. The first proposition normally would not be considered as the kind of proposition which was available as a means to the second. The point is this: While it is a necessary condition that two propositions one of which is considered as available as a means to the other are both true or could be true, it is not a sufficient condition for one being available as a means to the other; not *any* old pair of true propositions can stand in an available-as-a-means-to relation one with the other. Hence, there is an important characteristic of what we mean by "availability-as-a-means-to" which is not expressible in terms of the material "if, then". This property is given in the account of "availability" in *Webster's New International Dictionary, Second Edition*. There "availability" implies that something is "*efficacious* of its object". That is, φx is efficacious of ψx; colloquially, φx has the "power to produce" ψx.

35. Consider proposition (5). (5) is sometimes false. For it proposes that x expects that a non existent state-of-affairs p is as a means to (that is, contributes to the bringing about of) an existent state-of-affairs if p then any old state-of-affairs q. But we noticed earlier in this essay that 'q', in general, had to be something

realizable by x - something which x can make true. But even if we adopt this convention the proposition would not be true. For organisms do not, generally, expect that non-existent or non-realizable states-of-affairs contribute to the bringing about of anything. For example, (5) under appropriate substitutions might read: x believes that if 'he does not light a match' is true then also the conditional state-of-affairs 'if he lights a match, the room will blow up' is true; that is, he believes that not lighting a match contributes to the existence of the above conditional. This proposition could easily be false.

36. It may have occurred to the reader that what is involved in the "if, then" relationship between 'p' and 'q' in the means-end expectation is more appropriately rendered by the verb "causes". This is not quite true. For consider the animal who debates with himself at the choice-point. His provisional expectations (or hypotheses) are better expressed in the subjunctive; that is, they have the form 'x expects that if x were ... then x would ...'. In such a case, the organism does not expect that the "if" condition exists and probably doesn't expect that the "then" condition exists. But the expectation that p causes q, that is, the non-provisional expectation, includes both x's expectation that p exists and p contributes to (is sufficient to, the world being what it is) the bringing about of q. For example, consider the follow-

ing proposition: (Experimenter) y unprovisionally expects that if (animal) x does not demand that x be hunger-alleviated, x expects that when he eats food, he can alleviate his hunger. Symbolically this is:

$$(6)\quad e[y, \sim d(x, Hx) > 0 \supset (d(x, Hx) > 0 \supset e(x, fx \supset K(x, H)) > 0] > 0$$

where 'Hx' means 'x alleviates his hunger', 'fx' means 'x eats food' and 'K(x, H)' means 'x is able to alleviate his hunger'. According the TI. 2.b, if

$$(7)\quad yC(\sim d(x, Hx) > 0)$$

then we get

$$(8)\quad e[y, d(x, Hx) > 0 \supset (e(x, fx \supset K(x, H)) > 0] > 0$$

Under conditions (6) and (7), it is the present author's belief that (8) is seldom, if ever drawn as a result of those conditions. The reason is this: the investigator construes (6) and (7) as causing (8). But (6) would surely be regarded as false where the material "if, then" is replaced by a causal "if, then". In an unprovisional expectation, the experimental investigator expects that the antecedent of a conditional statement exists because he conceives that it is the existence of that antecedent which "brings about", "influences", (or the like) the truth, that is, the existence, of the consequent.

37. The causal "if, then" like the subjunctive "if, then" contains the elements of consecutivity and efficacity. Hence it seems possible to define the causal relation in terms of the relation of natural necessity.

We shall discuss this point in a moment.

38. In general, where 'pNq' means the world being what it is, p is sufficient to bring about q,

$$pNq \supset p \supset q$$

is true. Hence, if the laws of TI and TII were couched in terms of the relation 'pNq' both inside and outside of propositional attitude contexts, the present laws would hold. Ramification of the present system in the above terms would "correct" the present system in the sense that laws of the following sort would not follow:

(9) $e(x, \sim p \supset (p \supset q)) > 0$

(10) $\sim p \supset (p \supset q)$

(11) $e(x, (p \cdot q) \supset r) > 0 \equiv e(x, p \supset (q \supset r)) > 0$

Hence, certain things we regard as always true (though suspect) in the present system in the ramified system could not be deduced. This ramification, in other words, would have somewhat the same effect on the present system as a more detailed analysis of time arguments would have.

39. Let us now consider the relations 'pNq' and 'p causes q' that is, 'pC^Aq'. 'pC^Aq' may be defined as follows

$$pC^Aq =_{Df} [p \cdot pNq]$$

that is 'pC^Aq' means 'p is true and p is sufficient to bring about q (or p naturally necessitates q)'. Hence, the provisional expectation (or hypothesis) may be written:

$$e(x, pNq) > 0;$$

the unprovisional expectation may be written:

$$e(x, \ p \,.\, pNq)) > 0;$$

and hence by the definition:

$$e(x, pCq)) > 0 \,.$$

40. The above discussion of natural necessity suggests that the following propositions are appropriate:

i $\sim[(\sim p \,.\, p)\, Nq]$

ii $[p \,.\, pNq] \supset q$

iii $pN(qNr) \supset (p \,.\, q)Nr$

iv $[pNr \,.\, qNr] \supset (p \,.\, q)Nr$

v $pN(q \,.\, r) \equiv [pNq \,.\, pNr]$

41. Proposition i affirms that it is false that $\sim p$ and p are sufficient to bring about the existence of q. Because of this law the antecedent of the true conditional

$$[(\sim p \,.\, p)\, Nq] \supset [\sim p \,.\, p) \supset q]$$

cannot be asserted and hence we cannot deduce

$$(\sim p \,.\, p) \supset q,$$

though, of course, it is true in its own right. The above conditional is the result of replacing 'p' in postulate ii by $(\sim p \,.\, p)$ and exporting on postulate ii. In accordance with the preceding remarks, we also find that

$$e(x, \sim[(\sim p \,.\, p)\, Nq]) > 0$$

is true.

42. Proposition ii reads: if p, and if p necessitates q, then q. Proposition ii and the definition of cause

allow us to assert

$$pC^{A}q \supset q,$$

and also,

$$pC^{A}q \supset (p \cdot q),$$

and hence,

$$pC^{A}q \supset (p \supset q).$$

This latter statement when imported gives us

$$(p \cdot pC^{A}q) \supset q$$

which is an analogue of proposition ii. Again, by the definition of cause, proposition i, and proposition ii we get

$$\sim[(\sim p \cdot p)\ C^{A}q].$$

These formulae comport with the remarks on causation above.

43. Proposition iii affirms that if p necessitates that q necessitates r, then p and q necessitate r. By this proposition and proposition i we get

$$\sim[\sim pN\ (pNr)]$$

hence we cannot infer the truth-functional

$$\sim p \supset (p \supset q),$$

though, of course, it is true in its own right. The point here is that we don't get this law where we don't want it, e.g., in expectations. The converse of this postulate does not hold. The converse is rejected on grounds found in the preceding discussion of proposition (4).

44. Proposition iv affirms that if p necessitates r and q necessitates r then the conjunction of p and q

necessitates r. From it we can prove, with the help of proposition i,

$$pNq \supset \sim(\sim pNq).$$

45. Proposition v reads: p necessitates the conjunction of q and r is equivalent to p necessitates q and p necessitates r. From it we get an interesting theorem (with the help of proposition ii) namely

$$pN(q \,.\, r) \supset (p \supset q \,.\, pNr].$$

46. The above propositions and their theorems help to give the meaning of 'pNq' and also 'pC^{A}q'. We conclude this essay with the following remarks. In general, given

$$[p \,.\, p \supset q] \supset q$$

where p is taken to be true, and p ⊃ q is a proposition in <u>Principia</u>, we found that we could not infer q because 'p ⊃ q' was false in certain propositional attitude contexts. We therefore could not categorically say that whenever a proposition in <u>Principia</u> was imbedded in a propositional attitude context it was true. However, given

$$[p \,.\, pNq] \supset q$$

where p is taken to be true and 'pNq' is one of the above postulates or a theorem deducible from the same, q can be deduced. We can say this because whenever such propositions of the form pNq are imbedded in propositional attitude contexts they are valid. These remarks are of necessity only preliminary. Much work and investigation is needed in order to determine the full meaning of 'pNq'.

BIBLIOGRAPHY

Note: The present bibliography lists only items that are specifically mentioned in the present essay.

Bennett, A. A. and Bayles, C. A. Formal Logic. Prentice-Hall Inc., 1939

Carnap, R. Logical Syntax of Language. Harcourt, Brace and Co., 1937

Carnap, R. Testability and Meaning. Yale University Press, 1950

Carnap, R. Foundations of Logic and Mathematics. International Encyclopedia of Unified Science. Vol. II, Number 7, 1952

Hilgard, E. Theories of Learning. Appleton-Century-Crofts, 1948

Koch, S. "Clark L. Hull" in Modern Learning Theory. Appleton-Century-Crofts, 1954

MacCorquodale, K. and Meehl, P. "Edward C. Tolman" in Modern Learning Theory. Appleton-Century-Crofts, 1954

Marx, M. Psychological Theories. MacMillan, 1951

Quine, W. V. Mathematical Logic. Norton, 1940

Quine, W. V. Methods of Logic. Holt, 1950

Reichanbach, H. Elements of Symbolic Logic. MacMillan, 1947

Russell, B. Human Knowledge; Its Scope and Limits. London, 1948

Tarski, A. Introduction to Logic. Oxford Press, 1941

Thistlethwaite, D. "Critical Review of Latent Learning and Related Experiments." Psychological Bulletin, 1951, 48, pp. 97-129

Tolman, E. C. Purposive Behavior in Animals and Men. University of California Press, 1932 (Reprinted in 1951)

Tolman, E. C. "The Determiners of Behavior at a Choice Point" Psychological Review, 1938, 45, pp. 1-41

Tolman, E. C. "The Principles of Performance." Psychological Review, 1951, 62, Pp. 315-325

Whitehead, A. and Russell, B. Principia Mathematica. Cambridge University Press, Second Edition, 1950

Woodger, J. H. The Technique of Theory Construction. International Encyclopedia of Unified Science. Vol. II, Number 5, 1947

Woodger, J. H. Biology and Language. Cambridge University Press, 1954