First published in *Psychological Review*, 63, 1956, 249-264.

Mediation Variables in Scientific Theory

It is now nearly a decade since MacCorquodale and Meehl's widely misunderstood paper,"On a Distinction Between Hypothetical Constructs and Intervening Variables"(1948), appeared in this Journal. Since then, a rash of rebuttals has broken out in the psychological literature, some attempting to redefine the distinction along dimensions other than the one proposed by MacCorquodale and Meehl (e.g., Ginsberg, 1954; Marx, 1951; Seward, 1955), and others repudiating the distinction entirely (Bergmann, 1951; Kendler, 1952), until a point has been reached where writers on methodological issues no longer treat the two notions as separate concepts, but merely mention in a passing footnote that a distinction may possibly exist.

If this present conceptual chaos were merely a dispute over the naming of formal tools whose usage were nonetheless well understood, the interests of the hypothetical construct (HC)–intervening variable (IV) controversy would best be served by silence. Unfortunately, however, the HC and IV are fundamentally very different in the bases of their introduction into a scientific system, in their confirmation relevant to pre-existent scientific knowledge, and in their potency and functioning as methodological devices. Failure to distinguish between them inevitably leads to the construction of formal models hopelessly at odds with empirical reality, structurally inadequate, inflexible, and presuming precisely the concepts and relations most in need of experimental or semantic clarification. It is advisable, therefore, to re-examine the differences between HC's and IV's in special reference to a number of popular misconceptions that seem to have arisen despite MacCorquodale and Meehl's lucid original analysis.

Introductory Considerations

Preliminary to the major analysis, it is convenient to undertake a terse explication of the notions of "variable" and "empirical law," and a statement of three lemmas in light of which the differences between HC's and IV's will be examined.

Variables. A "scientific variable" is a propositional function—i.e., an incomplete sentence. The kind of variable basic to a scientific system is the "elementary scientific variable." (Other varieties exist, but are either irrelevant for systemic usage or are of a higher logical type.) An elementary scientific variable is an incomplete sentence whose blanks are of two kinds: (*a*) the "reference component," a blank which is filled from a domain of object or event names; and (b) the "predicate component," a blank filled from some domain of adjectives or sentence particles such that when the reference component is filled with the name of a specific object, there is one and only one adjective (excluding synonyms) which, when inserted in the predicate component, makes a true sentence.[1] For example, "o weighs x gms." is a weight variable where o ranges over the set of objects under consideration (men, cargoes, molecules, or whatever) and x covers the domain of positive real numbers. Propositions which are either true or false about an object also fall under this rubric; "o is alive" can be reinterpreted as "It is x that o is alive," which defines an elementary variable when x ranges over the two-valued domain, "true, false." In short, what a variable does is to predicate of each object in the system one of a set of mutually exclusive attributes sharing an underlying similarity which gives the variable its name and distinguishes it from other variables.[2] It is important for conceptual clarity to recognize that although in a technical sense a variable is a relation between an object domain and a predicate domain, the elementary scientific variable is *not* a relation in the sense that the laws of empirical science are relations. Rather, these variables are the units among which such relations hold. Thus to define, with Maze (1954), the IV as the empirical relation between observables is not merely to lose sight of the original meaning of "intervening variable" but to contaminate the whole concept of "variable."

Empirical laws. When two or more variables have been defined over the same domain of objects, it is possible that empirical relations exist among them. To say that a (nontrivial) empirical relation exists between two elementary scientific variables x and y[3] is merely to say that although y is not logically dependent on x, the value of y for a specific object o can be predicted with greater accuracy when the value of x is known for o than when x is not known. This relation takes the form, $\hat{y} = f(x)$, where $\hat{y}$ is the best estimate of y, given the value of x.[4] Such relations are traditionally referred to as empirical "laws." Unfortunately, the term "law" carries with it connotations of a degree of predictability approaching certainty. In actuality, an empirical law whose accuracy is expressed by a correlation coefficient

---

[1] For simplicity, only the case of a single reference and predicate component has been described. More complex elementary scientific variables may have multiple reference and predicate components.

[2] Formally, ontological similarity of the attribute set defining a variable is superfluous. But even formally the functional similarity remains.

[3] More explicitly, the variables x and y should be written "$P(o,\, x)$" and "$Q(o,\, y)$", where P and Q name propositional functions and x and y merely represent the blanks to be filled when a specific value of o is given.

[4] Empirical laws are usually written "$y = f(x)$" with "y" replacing "$\hat{y}$". This has the effect of implying, somehow, that the equation gives the "true" value of y for a given x, and that departure of the real value of y from this computed value is merely a contamination of the law in the same sense that the law of falling bodies is contaminated by air resistance effects (the latter law, of course, being a theoretical derivation and not an empirical relation).

of $\eta = .01$ has the same systemic and ontological status as a law whose accuracy is $\eta = .99$. It is important to note that an empirical law cannot be discovered unless the related variables have first been explicitly identified. That is, the empirical relations in which a variable participates cannot contribute to the definition of the variable and are thus irrelevant to its ontological content.

I now wish to state three lemmas for future reference. These contain certain cold facts of life which, though hardly controversial, have implications which are usually overlooked by the traditional, highly oversimplified discussion of psychological methodology.

Lemma I. Empirical laws as actually observed by psychologists never attain perfect predictability. We have yet to exhibit an observation variable, psychological or otherwise, which could be predicted with absolute certainty from a set of other known observation variables. This is not a question of whether perfect relations exist in fact, but merely recognition that we, as errant humans, never quite succeed in discovering all the empirical determinants of a given variable. Even for empirical laws which seem to admit of no practical uncertainty, we still have a small range of indeterminancy which we ascribe to "errors of measurement."

Corollary. An empirical relation between two observation variables x and y must always be expressed "$\hat{y} = f(x)$" and never "$y = f(x)$." The latter form claims an exact relationship between x and y and is hence falsified by the first observation not precisely satisfying the function, no matter how minute the divergence.

Lemma II. No two (logically distinct) observation variables can in practice be demonstrated to have exactly the same determination of a given third observation variable. Because of a popular misconception concerning the notion of predictive equivalences, some explanatory remarks are in order here. It not infrequently occurs that two or more different experimental manipulations have similar effects on the behavior of the organism. Thus we might discover, e.g., that (a) a 12-hr, food deprivation, (b) a 24-hr, food deprivation with ad lib sugar solution ingestion, and (c) a 36-hr, food deprivation with intravenous glucose injection one hour before running all have approximately the same effect on a rat's rate of lever-pressing for food reinforcement. Under such circumstances, we are accustomed to stating that these are alternate variables with the same behavioral consequences. Unfortunately, this terminology, in leaping from operations to variables defined by the operations without explicitly constructing the propositional functions describing the variables, carries the implication that it is possible for only one of a set of variables to hold for an organism. Actually, as commented earlier, each object of a scientific system partakes of some value for every elementary variable in the system. When we say that two variables x and y are equivalent in their consequences for a third, z, we in reality conceive of x and y having neutral areas in their distributions (e.g., the value corresponding to the rat's not having received an intravenous injection), and mean that there are nonneutral values of x conjoined with neutral values of y on the one hand, and nonneutral values of y coupled with neutral values of x on the other hand that have the same predictive consequences for z.

Thus the notion of predictive equivalence among variables is a somewhat misleading abbreviation for the notion that more than one point in a joint multivariate distribution can have the same predictive implication for some dependent variable. That approximations to such equivalence points exist has been repeatedly demonstrated. The claim of this lemma is that we are never able to demonstrate the absolute equivalence of two distinct points. We have many quasi-empirical laws which predict the same value of the dependent variable from multiple points in the joint distribution of the independent variables-for example, all objects having the same ratio of volume to weight have the same prediction as to whether they will float in water. Such laws, however, are never strictly empirical with the prediction from every point independently checked experimentally, but involve the fitting of a simplified equation, usually continuous and frequently involving additional rational assumptions, so that the estimates computed by the equation are seldom strictly the best *empirical* estimates available. This again is not to take a stand on the factual existence of equivalence points, but merely to point out that we, as ruthless empiricists, are never able to demonstrate them exactly, and for behavioral variables, seldom even approximately.

Lemma III. The empirical laws discovered and formulated by psychologists always presuppose specific fixed values of additional variables. The laws that prescribe the behavior of an organism presume that it is alive and that the neuromuscular structures and environmental supports necessary to mediate the behavior are intact. Again, except for certain special studies, it is further assumed that the animal is not ill, has not received various drug injections, is not deficient in the relevant sensory modalities, and so forth for many other items. That empirical behavioral laws as actually formulated presuppose specific constants in no way detracts from their validity, but it is important to realize that such laws in no way specifies the predictive consequences of their independent variables when the relevant constants are altered. For a proposit tion "P implies Q" in no way specifies the implications of not-P.

Mediation Variables

Before differentiating between intervening variables and hypothetical constructs, it is expedient to examine the more generic concept of "mediation variable." Let us suppose that : $P(o,\, x)$ and $Q(o,\, y)$ are propositional functions whose meanings have already been established, and furthermore, that an empirical relationship has been discovered between x and y.[5] This means that given a value for x, we can estimate y as a function of x, a situation which may be symbolized ' "x" $\rightarrow$ "$y = f(x)$)", ' the quotation marks indicating that $y = f(x)$ is not a law of nature, but rather that $f(x)$ is the value of y which is anticipated s by the scientist under belief that x has the given value. To represent this relation without reference to the extrasystemic scientist, we symbolize the relationship, as discussed earlier, by "$\hat{y} = f(x)$." (The latter is able to represent an approximation to a fact of nature because y is a property not of a given object, but of the class of objects having a common value of x.)

It is seldom, however, that the scientist—psychologist included—is content merely to infer y from x. Whatever his motives and methodological orientation, he usually avails himself of a mediation variable, $K(o, m)$, the value of which is stated as a function of x and which, in turn, is used to predict y. The scientist's behavior in this respect may be symbolized ' "x" $\rightarrow$ "$m(= g(x))$" $\rightarrow$ "$y(= \phi(m)$",' recognizing by quotation marks, as before, that the elements related are beliefs of the scientist and not principles of nature. When we attempt to justify the scientist's behavior by translation of this sequence of beliefs into a statement about the properties of the system, however, we find not one possible notation but three: (a) "$m = g(x),\ \hat{y} = \phi(m)$"; ($b$) "$\hat{m} = g(x),\ y = \phi(m)$"; and (c) "$\hat{m} = g(x),\ \hat{y} = \phi(m)$." In each case, elimination of m can yield no more predictive accuracy than the original empirical relation $\hat{y} = f(x)$, but in case (a), m is linked to its antecedent by a rigorous, nonempirical relation, while in (b) and (c), passage from x to m is empirical and hence subject to error.

In other words, a mediation variable may be introduced as a rigorous function or as an empirical function of its antecedent variable(s),[6] but not as both. If the function is rigorous, the dependency of m on x cannot be empirical (Lemma I) and m must derive analytically from x, with a corresponding analytic relation between the ontological contents of x and m. Conversely, if the dependency of m on x is an empirical relation, the definition and hence the ontological content of m must be independent of that of x, and of the relationship between x and m. Thus investigation of the nature and functioning of mediation variables must proceed along two distinct (though not orthogonal) dimensions: (a) the ontological content of the variable—i.e., what it is that the propositional function defining the variable predicates of things; and (b) the grounds upon which the scientist is able to justify his introduction of m into the estimation of y from x—i.e., the type of relation between the mediation variable and its antecedents. Much of the confusion which has arisen concerning the differences between HC's and IV's seems to be

---

[5] For simplicity, we discuss only the case of a single independent variable. The analysis can be extended to the multivariate case by conceiving x as a vector.

[6] Here, as henceforth, "antecedent variable(s)" refers to the variable(s) by means of which the mediation variable is systemically introduced.

due to interpretation of the MacCorquodale and Meehl distinction as primarily a distinction of ontological content. Actually, as will become apparent, the ontologies of both HC and IV cover essentially the same range, and it is primarily along the second dimension that the distinction is to be made.

Intervening variables. By "analytic mediation variable," or "analytic variable" for short, let us understand a mediation variable which is derived from its antecedent variables by an analytic relation. Analytic variables, in turn, are of two types: (a) primitives, which are logically prior to their antecedent variables, and (b) transformation variables, for which logical priority goes to their antecedents. The first type, primitive analytic variables, are of little interest to us here and need be mentioned only for technical accuracy. They comprise the class of mediation variables whose antecedents are logical compounds containing m as a constituent—e.g., the derivation of a person's height given the vector describing his vital statistics, the latter including height as a component. The second type, transformation variables, is our primary focus, for these comprise the class of mediation variables which MacCorquodale and Meehl have named "intervening variables."

Actually, there is little by way of exposition that can be added here to that already presented by MacCorquodale and Meehl. An intervening variable (transformation variable) is a variable which has been systematically defined in terms of its antecedents and is hence dependent upon the latter for its meaning. To recapitulate the analysis by means of a formal example, let us suppose that $z = (x+y^2)/w$ is an empirical relation expressing the best estimate of a known variable z as a joint function of other known variables w, x, and y. We are now able to express z as a function of analytic mediation variables by defining one or more of these as mathematical functions of w, x and y. There are infinitely many different ways in which this can be done.[7] To list but a small sample :

$$a =_{\mathrm{def}} (x + y^2),\ b =_{\mathrm{def}} w^{-1},\ \hat{z} = ab\ ; \tag{1}$$

$$c =_{\mathrm{def}} x/w,\ d =_{\mathrm{def}} y/\sqrt{w},\ \hat{z} = c + d^2\ ; \tag{2}$$

$$e =_{\mathrm{def}} 1/2(x + y^2)/w,\ \hat{z} = 2e\ . \tag{3}$$

Although an intervening variable is usually introduced so as to simplify as much as possible the written expression of the empirical relation into which it is introduced, any transformation of the antecedent variables into an arbitrary domain of

---

[7] MacCorquodale and Meehl slightly understated the case when they observed that Hull could have introduced intervening variables into his law of response strength by 15 ways other than the one he in fact chose. The actual number of possibilities is one of the higher transfinities.

symbols is a perfectly legitimate definition of an IV, and can (with the perhaps necessary assistance of other similarly defined IV's) replace its defining variables in the empirical equation under consideration. Intervening variables are of frequent and important application in statistical analyses when it is desired to transform a distribution of raw scores into a distribution, such as the unit normal distribution, with special mathematical properties.

In description, then, of the nature of the IV along the second dimension mentioned above, we see that the IV is derived nonempirically from its antecedent variables. That is, passage from the latter to the former does not admit of uncertainty, for when an object's values of the antecedent variables are known, its value of the IV is also rigorously defined. On the other hand, the empirical law relating the IV to the dependent variable can at best be no more accurate than the original unmediated empirical relation and actually represents a loss of predictive accuracy unless the definition of the IV in terms of its antecedents is a one-one function. This is because in general, as pointed out in Lemma II, no two points in a multivariate distribution ever have quite the same strictly empirical prediction to the dependent variable, and whenever two or more such points define the same value for the IV (as obtains, for example, when the IV is defined as a ratio of two antecedent variables), prediction from that value of the IV never can have quite the same predictive accuracy as prediction from any one of the corresponding values of the antecedent variables. As for the ontological content of the IV, it is precisely the same as that of the antecedent variables which define it. There are various technical devices by which a definition can be accomplished in a formal system-explicit definition, recursive definition, reduction sentences, etc.-but all have the common characteristic that any sentence containing a defined term can be replaced by a sentence not containing that term. If $K(o, m)$ is an IV defined in terms of an antecedent variable $P(o, x)$, and $a_1, a_2, \ldots, a_n$ are specific values of x generating the same specific value, b, of m, then the assertion "$K(o, b)$" about an object o has the same meaning as "$P(o, a_1)$ or $P(o, a_2)$ or $\cdots P(o, an)$." To say "o has a density of 2 gm./cc." is to say, "Either o has a volume of 2 cc. and weighs one gm., or has a volume of 3 cc. and weighs 1.5 gm., or $\cdots$, or $\cdots$" etc. The latter sentence completely exhausts the meaning of "density."

Hypothetical constructs. In the last section we considered those mediation variables which derive analytically from their antecedents. For the remainder, passage from antecedents to mediators must be made with less than logical certainty, and such variables we shall designate as "inferred variables." Precisely because inferred variables are not analytically contained in their antecedents, it follows immediately that (a) the meaning, or ontological content, of the inferred variable is not reducible to that of its antecedents, and (b) a certain degree of error must be expected in passage to the inferred variable from its antecedents.

With some reluctance, we now distinguish between two main types of inferred variables: (a) inferences where the specific value of an object is estimated on a variable already accepted as existent for that object, and (b) inferences where not only specific values are estimated, but where for that object the very existence of the generic property defining the variable is subject to critical doubt.[8] As a practical example of the first type, we might conceive of a situation in which a construction foreman, needing a heavy worker to ballast an unriveted beam, selects the tallest and widest man he can find. Although the foreman has probably never done an empirical study relating height and width to successful ballasting, he knows that on the one hand, ballasting ability is directly determined by weight, and that on the other hand, weight in humans can be closely inferred from height and width. To illustrate the second type, we may take the traditional example of the kinetic theory of gases and the empirical gas laws. Given two of the three observation variables—pressure, temperature, and volume—for an enclosed volume of gas, we can infer the distribution and kinetic energy of the gas particles, and deduce the third observation variable, in turn, from this inference. In addition, however, we must hypothesize that the property of being composed of a large number of small elastic particles does, in fact, exist and hold for gases.

Because the very existence of inferred mediation variables of the second type is inferential, and because, as seen below, such properties can be creatively imagined, the term "hypothetical construct" is singularly appropriate for variables of this sort. However, this term calls undue attention to the unverified nature of the variables, and not infrequently leaves the impression that such variables are nights of metaphysical fancy. Actually, all properly formulated HC's, even the more abstract ones, maintain close contact with testable reality and many involve only the most pedestrian kind of speculation. In fact, it can be argued (and argued convincingly, in my opinion) that *all* scientific variables are hypothetical constructs or transformations thereof, even such primitive observables as length, weight, lever-pressing frequency, etc., for even these, as part of their implicit definitions, contain notions of properties such as constancy under spatial translation, existence in a "real" world, etc. which far transcend the immediately given experiences of the scientist. Since MacCorquodale and Meehl have stressed the existential, rather than the hypothetical, aspects of HC's, I suggest that by "hypothetical constructs" we understand all cases of inferred mediation variables, even though in a few instances, the generic existence of the variable and its applicability to the objects concerned may admit of little practical doubt.

At this point, it becomes mandatory to say a few words about the ontological content of hypothetical constructs, even though philosophers have yet to produce

---

[8] To claim, prior to knowledge of the specific value of x for o, that a variable $P(o,x)$ holds for an object o is to assert the proposition "$(\exists x)P(o,x)$." The present distinction is in reference to whether or not this proposition is considered doubtable.

any substantial agreement on this (cf. Feigl, 1950 for a comprehensive summary of the major interpretations). There would appear to be three distinct levels of uncertainty for the existence of a HC. On the lowest level are variables whose generic defining property is already accepted as existent, but whose legitimate predication of the objects being dealt with is uncertain.[9] Thus in regard to the variable, "The particles comprising the gas in container o have a total kinetic energy of x," we already have a positivistic understanding of an aggregate of discrete particles—as, for example, experientially gleaned from observing a table of pool balls—and the properties abstracted from their motions, but do not know whether, in fact, the gas in container o is comprised of such particles. Inferences on this level involve no more methodological subtlety than does inferring from a man's stated preferences in female beauty the color of his wife's hair, even though it may not be known whether the man is married.

On the second level of uncertainty are those hypothetical constructs consisting of a compound predicate whose constituent properties are accepted as existing separately, but whose conjoint existence is uncertain. Most of the hypotheses of everyday life are on this level. Thus as Bergmann (1951) points out, to speculate on the existence of a unicorn is simply to hypothesize that the properties describing the appearance of a horse and the property of having a long horn on the forehead exist conjoined in a single object. Or when we infer from an empty rat cage and scurrying sounds under the cage rack that an escaped rat is in the corner of the laboratory, we hypothesize that there is an object which possesses both the properties of being small, white, furry, etc., and being located at the spatial-temporal coordinates of that corner.

Third-level constructs, a more sophisticated version of level two, and the level comprising the bulk of scientific constructs, are also compounded of separately defined predicates. On this level, however, at least one member of the compound is a relational predicate entailing a functional association, usually a rigorous or semi-rigorous[10] one between the HC and other variables of the system. Thus part of the hypotheses underlying the kinetic theory of gases is that the particles of which a gas is presumably comprised obey certain laws of elasticity, conservation of momentum and energy, etc. which permit the derivation of the observable properties of gases as a rigorous analytic function of the underlying HC's (subject to certain secondary disturbances). We shall return to the relational assumptions of the HC in slightly greater detail later.

---

[9] Formally, this level comprises variables for which "$(\exists o)(\exists x)P(o,x)$" is accepted, but where for a specific object o, "$(\exists x)P(o,x)$" is questionable. For second and third level constructs, even "$(\exists o)(\exists x)P(o,x)$" is open to practical doubt.

[10] By saying that the determination of y by m is "semi-rigorous," we mean that m exactly determines y by the specified relation so long as certain supplementary variables remain constant or under certain theoretical constraints.

It is important to note that the HC is no more unverifiable than the predicates from which it is constructed. For constructs of the first and second level, whatever test procedures are considered adequate to verify the constituent predicates taken separately are in principle jointly applicable as a test for the HC. Thus the particulate nature of gases is in principle directly verifiable (or refutable), given a sufficiently powerful magnification device. The relational predicates of the third level are not so easily tested directly, but since such predicates need be built into the construct only when some testable covariation among observables is entailed, they are subject to techniques for inductive corroboration.

The feature of the HC which gives it its particular fruitfulness as a mediation variable and its convincingness as an explanatory device is precisely that feature which has caused some of our more methodologically chaste psychologists to protest in moral indignation—the "excess meaning" of the HC. Since the HC is defined independently of the variables from which it is inferred, the HC may have analytic implications or known empirical relations far beyond anything known or surmised about its antecedents. In regard to first- or second-level elements of the HC, we may have much independent empirical knowledge concerning the behavior of constituent properties. Suppose, for example, that, in the analysis of visual discrimination, we hypothesize that discrimination behavior is mediated by electromagnetic excitation of retinal cells. By this inference, we have immediately brought into consideration a mass of empirical data concerning electromagnetic waves and retinal cells arrived at independently of discrimination research (e.g., the observation that electromagnetic waves will not pass through certain types of obstructions or that retinal cells have such-and-such limits of response) which would otherwise be irrelevant to discrimination problems but now permits prediction of the dependent discrimination variable (e.g., prediction that the organism cannot discriminate wearing opaque goggles, or that discrimination cannot exceed such-and-such acuity). With third-level constructs, the relation of the dependent observation variable to the HC is either rigorously or semi-rigorously entailed by the constitution of the HC. When we have an empirical relation $\hat{y} = f(x)$, and insert a HC as a mediation variable so that $\hat{m} = g(x)$, $\hat{y} = \phi(m)$, the uncertainty of passage from m to y may be as little as we please, depending upon our construction of m. Passage from x to m must be a calculated risk, not merely in inference to a specific value of m but to the very existence of m. Once the risk is accepted, however, we have not only a "rational" explanation for $\hat{y} = f(x)$, but also (a) prediction as to the relation of x and y beyond the actual limits tested empirically at the time, and (b) predictions of empirical relations between x or y and other variables designated through their relations to m but bearing no intuitive relevance to the $\hat{y} = f(x)$ association. That is to say, inference to the existence of an HC inevitably entails many more predictions to empirical relationships than merely the one for which the construct was initially invoked.

Let us now consider some attempts to clarify several misconceptions that seem to have arisen concerning intervening variables and hypothetical constructs.

Intervening variables do not lack existential reference. MacCorquodale and Meehl have been accused by Ginsberg (1954) of defining the IV as something entirely arbitrary and without existential reference. We have just seen, however, that while the algorithmic device by which we formally define and compute the IV is essentially arbitrary (as are, indeed, all our symbolic techniques), the ontological referent of the IV is the same as that of the variables from which it is derived. Hence if the antecedent variables designate "real" properties, so must the IV, and similarly, the latter must be empirically related to any other variables that are empirical functions of its antecedents.

The hypothetical construct is not a "convenient fiction." There seems to be a widespread belief, not merely among psychologists, that HC's are "convenient fictions," or artificial formal devices. If the properties which are hypothesized to co-exist in the HC have "real" referents individually, they do not somehow become fictional just because their conjoint existence has not as yet been verified. An HC cannot simultaneously be both convenient and fictional, because if the hypothesis expressing the HC is in fact false, the entire system incorporating the HC becomes basically useless. For a conjunction of propositions is falsified by the falsity of any one of its constituents.

The hypothetical construct is not unobservable in principle. Unless we adopt an extreme solipsism and assert that, e.g., the objective world exists only when we see it, there is nothing about an HC which in principle bars it from direct confirmation. This has already been discussed above, and the point is simply illustrated through the example of a mechanic diagnosing a malfunctioning automobile. His inference of dirt in the carburetor is to a hypothetical construct, but to verify or reject the hypothesis, all he need do is to look in the appropriate part of the engine.

The excess meaning of the hypothetical construct does not consist of vague, undesirable prescientific connotations. I suspect that much of the distrust shown toward HC's by some of our more tough-minded psychologists has stemmed from fear that the HC is a methodological Trojan Horse by which mentalistic notions such as "idea," "desire," "self," etc. are surreptitiously to be re-introduced into behavioral psychology. Any vagueness inhering in an HC, however, can only be due to vagueness in the definitions of its component properties. It is true that the ramifications of an HC may not at first be fully comprehended, but these additional commitments to testable expectancies are precisely wherein the potency of a good HC lies. Vagueness in a concept, on the other hand, has just the opposite systemic effect. The reason that older mentalistic notions or more modern refinements such as "libido," "r_g", etc., have proven objectionable is precisely because their lack of explicit description prevents derivation of testable, unambiguous consequences.

The intervening variable is not a manifestation of, nor is the hypothetical construct the antithesis of, methodological rigor. Since an IV is merely another way of expressing its antecedent variables, its introduction in no way adds to or subtracts from whatever formulational rigor the system may possess. On the other hand, while it is not difficult to use HC's poorly, their proper employment necessitates explicit delineation of essential properties and careful derivation of both empirical and analytic implications to testable propositions—demanding, in short, a high degree of methodological rigor.

Intervening variables and hypothetical constructs do not represent different degrees of operational validity. Marx (1951) has attempted to redefine IV's and HC's as distinguishable only as different degrees of operational validity, high for the former and low for the latter. I am no longer certain what "operationally valid" means, since the term has all but lost its operational definition and has, rather, become a synonym for "good concept" or an antonym for "sinful." The notion, however, scarcely seems applicable to the IV-HC distinction. Since the IV is analytically derived from its antecedents, the former has precisely the same operational validity as the latter, which may be of any degree. Similarly, the range of operational validities for HC's must cover almost the entire spectrum of possibilities, according to the uncertainty of the inference by which the HC is introduced.

Neither the intervening variable nor the hypothetical construct can be "anchored on both sides." The requirement that a mediation variable be anchored on both sides, first introduced by Hull (1943, p. 22), stipulates that a mediation variable must be introduced from observation variables by an equation or chain of equations, and must, in turn, lead by an equation or chain thereof to another observation variable. The fact that all empirical laws actually known to psychologists are but imperfect estimates of the dependent variable precludes the linkage of dependent and independent variables by any chain of exact equations. This means that in any chain of mediation variables linking observables, at least one link must be nonanalytic and hence irrelevant to the definition of the mediation variables. Thus a mediation variable bracketed by observation variables can be anchored rigorously to at most one end of the empirical relation mediated, while the remaining linkage remains strictly empirical and hence logically unnecessary.

Laws containing only intervening variables have no inductive potency. The mathematical expression for a finite number of points may be extrapolated by an infinite number of alternate equations. Also, as seen earlier (Lemma III), an empirical law always presupposes specific fixed values for other implicit variables. Hence an empirical law may be considered reliable only for the actual values of the variables tested, and provides no justifiable predictions for the relation of the empirical variables outside the range tested, much less to predictions of the relation when the values of the implicit variables have been changed. Since a law containing only IV's is merely shorthand for an empirical relation, we must conclude that such laws cannot rationally be extrapolated beyond the data from which they were derived without additional assumptions. On the other hand, if passage from the HC to the observation variables is in part analytic (as is almost always the case), or if the excess meaning of the HC permits an evaluation of its relation to the implicit variables held constant in the original empirical law. then introduction of an HC as a mediation variable into an empirical relation may extend the relation (subject to experimental disconfirmation) far beyond the limits of the original observations. The intervening variable is NOT a causal mediation variable. Our earlier analysis of IVs as a type of mediation variable employed the term "mediation variable" strictly to designate a syntactic property-i.e., to indicate a formal expression written by the scientist between his formal symbols for the dependent and independent variables of an empirical relation. We frequently have reason to believe, however, that given an empirical relation in which a variable y covaries with a variable x, there exist one or more "real" variables v whose identities may be unknown to us but which causally mediate between x and y. That is, we suspect x tends to determine y only because x tends to determine v and v, in turn, causally influences y, so that if v were to be prevented from varying, the relation between x and y would vanish-e.g., as in the case of the relation between the first and last gear in a train of gears. Now, since an HC has an ontological content independent of that of its bracketing variables, it can and frequently does occur that the HC, if it in fact exists, behaves as a causal mediator for the empirical relation which it is invoked to explain. Thus, neurophysiological constructs such as nerve impulses, etc., are frequently invoked as causal mediators between stimulus and response variables. Intervening variables, on the other hand, can in no way represent causal mediators, since the IV is merely a syntactic transformation of its antecedents and has the same ontological referent as the latter. Similarly, the breaking up of a complex empirical relation into several simpler equations through the utilization of several intervening variables in no way need reflect the dependency structure of causal mediators, since the resolution into IV's is entirely arbitrary.

The intervening variable does NOT tie together sets of independent and dependent observation variables and permit a reduction in the number of laws describing their interrelations. As originally observed by Feigl (1945) and recently spelled out in detail by (Seward, 1955), a mediation variable may permit a sharp reduction in the number of laws required to express the relationship between k independent and n dependent variables. Taken pairwise, k independent and n dependent observation variables determine kn empirical relations. If these relations follow a certain pattern, however, a common mediation variable m can be written as separate functions of the independent variables—$m = f_1(x_1), \ldots, m = f_k(x_k)$—and the dependent variables, in turn, written as individual functions of the mediator—

$y_1 = \phi_1(x_1), \ldots, y_n = \phi_n(m)$—so that the total number of equations becomes reduced to $k + n$. Seward has explicitly claimed to find no distinction between HC's and IV's in regard to such a mediator.

The discerning reader will observe immediately that the above system of interrelations, where k independent variables appear to be equivalent in the determination of each of n dependent variables, is a more complex form of the situation previously discussed in Lemma II, concerning the equivalence of more than one variable in the determination of a single dependent variable. It now becomes apparent that a mediation variable which functions in the manner just described cannot be an IV but must, rather, consist of a HC. Reasons:

1. An IV cannot be defined simultaneously by two or more independent functions. To define, e.g., an IV simultaneously as $m =_{\text{def}} f_1(x_1)$ and $m =_{\text{def}} f_2(x_2)$ implies that $x_2 = f_2^{-1} f_1(x_1)$ thus asserting a rigorous covariation between x_1 and $x_2$which in practice never occurs when x_1 and x_2 are not analytically equivalent. Hence defining an IV in terms of more than one observation variable can only be accomplished by: $m =_{\text{def}} [f_1(x_1) \text{ or } f_2(x_2)]$. But since observation variables x_1 and x_2 are never demonstratively related in the same way to the dependent observation variable y, prediction of y from m is less accurate than from either x_1 or x_2, and the value of y predicted from m differs from that predicted from either x_1 or x_2. In other words, since no exact equivalences can be found among k independent observation variables in their determination of n dependent variables, it is not mathematically possible exactly to reproduce the kn empirical relations by $k + n$ equations with an IV as the mediation variable. For psychological laws in particular, where we have yet to demonstrate even close equivalences, the loss of predictive accuracy through attempting to introduce a single IV as a common mediator would be tremendous. On the other hand, when k independent variables seem reasonably similar in their determination of n dependent variables, it seems very reasonable to suspect the existence of some additional variable, a HC prior to its actual detection, whose value is substantially determined alternatively by x_1 or x_2 or $\ldots x_k$ and which, in turn, substantially determine $y_1, y_2, \ldots,$ and y_n. Thus for the example in Lemma II, where different nutritional schedules were presumed to have similar effects in determining lever-pressing rate, it would seem not illegitimate to conclude that while these operations are not identical in their influence on lever pressing and might, for example, be quite different in their cue value (S^D) for that behavior, they may nonetheless similarly influence an underlying behavioral variable called "degree of hunger" whose meaning is independent of the operations by which it is aroused.

2. A mediation variable is frequently inserted into a matrix of k independent and n dependent variables before all kn empirical relations have been experimentally determined. Thus Seward (1955) speaks of testing the validity of a mediator by

observing empirical relations: $\hat{y}_1 = g_1(x_1)$, $\hat{y}_2 = g_2(x_2)$; inserting a mediation variable: $m = f_1(x_1), \hat{y}_1 \phi_1 2(m_1)$; $m = f_2(x_2)$, $\hat{y}_2 = \phi_2(m)$; and *subsequently* testing the validity of the inferred relations: $\hat{y}_2 = \phi_2(f_1(x_1))$ and $\hat{y}_1 = \phi_1(f_2(x_2))$. When m is an HC, inferences to previously untested empirical relations may easily be justified by the excess meaning of the construct. On the other hand, if such a mediator is an IV, there is no reason to infer that the complete $k \times n$ covariance structure is of the sort reducible to $k + n$ laws. For no IV can provide a basis for any inference not already inferrable directly from the empirical relations into which the IV has been introduced.

The Content of Hypothetical Constructs

By now, I hope, the reader will have come to regard with some favor the writer's view that the methodological properties of the intervening variable and hypothetical construct are drastically distinct; that the IV, while aseptically respectable, achieves this through inductive, explanatory, and heuristic sterility; that the HC, precisely because of its fertility, involves a calculated inferential risk and a certain virtuosity in construction; and finally, that either to deny the legitimate usage of the latter or to introduce the latter with the complacency entitled by the former is to demonstrate a less than adequate comprehension of basic scientific methodology. From this, it follows immediately that one of the most urgent problems facing psychological theory is an explication of the more detailed nature of hypothetical constructs and of their optimal construction and employment. While such an undertaking is considerably more formidable than any possible of attempt here, I wish, nonetheless, to offer some suggestions as to the necessary and sufficient ontological content of HC's suitable for behavioral theory.

It has been almost unanimously presumed that any HC invoked to account for molar behavioral principles must constitute an inference specifically about neurophysiological processes underlying that behavior. This belief, coupled with the faith that behavioral laws are capable of formulation on a molar level alone, has doubtless been a major factor in the aversion for HC's shown by some behaviorists, and why others have attempted, through redefinition, to infuse the term "intervening variable" with the formal properties of the hypothetical construct (cf. Bergmann, 1953 for an especially flagrant example). Personally I have the fullest sympathy for those who patiently seek an explanation of organic behavior in terms of neurophysiological mechanisms. That molar behavioristics is in theory completely reducible to underlying neurophysiological principles is not a hypothesis but a logical derivation from the fact that every instance of a response is analytically concomitant with a specific set of neurophysiological events, so that a completely deterministic neurophysiology must of necessity permit derivation of all molar behavioral laws. It is seriously to be doubted, however, whether the present state of physiological knowledge is sufficient for an HC constructed of neurophysiological predicates to have many logical or empirical implications for behavioral relations. Most, if not all, studies permitting an integration of behavioral and neurophysiological principles have resulted in extension of our knowledge not so much about behavior as about neurophysiology. For while existent neurophysiological data impose little or no restrictions on possible behavioral relations, existent knowledge of the principles of behavior prescribes an important set of minimal properties that neurophysiological processes must possess.

Neurophysiological content is not only at present inadequate for behavioral HC's, but also basically irrelevant. Presuming the switchboard theory of neural action to be true, we could in principle replace each neurone with a mechanical or electronic device with the same connections and input-output characteristics as the replaced cell without affecting the validity of any behavioral law. Since neurophysiological attributes are thus not essential for a physical behavior system, a set of HC's need possess no neurophysiological content in order to describe the necessary properties of the causal mediators underlying behavior. Nor does the number of HC's required for an adequate theory of behavior need be on the order of the number of neurones in the nervous system any more than the kinetic theory of gases requires a separate variable for each gas molecule. The functional units of behavior—i.e., the essential properties of the system which actually make a difference for behavioral outcomes—are undoubtedly properties of neural aggregates or mass actions, just as in a Hebbian brain the functional unit is the phase sequence, the behavior of which can be formulated without reference to any constituent neurones.

I suggest, therefore, (a) that despite the complexity of the central nervous system, those of its properties which are essentially involved in the causal determination of the dependent variables of behavioral psychology may possibly be reflected by a relatively small number of HC's without neuro-physiological reference; and (b) that the necessary ontological content of such HC's need consist only of existential quantification (the logical operator "$\exists x$") and relational predicates. In conjunction with this latter suggestion, I wish to point out that an HC may be introduced with no more meat on its bones than "There exists a variable m such that for observation variables x and y, $\hat{m} = g(x)$and $\hat{y} = \phi(m)$." For example, given the existence of an empirical relation $\hat{z} = wxy$, we might hypothesize, "There exist variables m and n such that $\hat{m} = wx$, $\hat{n} = y$, and $\hat{z} = mn$." It is very important to distinguish such a hypothesis from a definition of intervening variables: "$m =_{\text{def}} wx$, $n =_{\text{def}} y$, and $z = mn$." The existential hypothesis asserts an *empirical* relationship between the mediation variables and their antecedents, thereby leaving room for (a) additional hypotheses concerning the nature of m and n, (b) possible independent substantiation of the existence of m and n, and (c) improvement or correction of the inferred relations among the mediation and bracketing variables in the light of advancing knowledge.

But most important of all are the *structural* inferences contained in the . existential hypotheses defining HC's.[11] Decomposition of a complex empirical relation by means of IV's has no necessary relation to the covariance structure of the causal mediators underlying the empirical relation. An intervening variable breakdown, "$m =_{\text{def}} wx,\ n =_{\text{def}} y,\ \hat{z} = mn$," has precisely the same ontological meaning and predictive implications as "$m =_{\text{def}} w,\ n =_{\text{def}} xy,\ \hat{z} = mn$," To hypothesize "$\hat{m} = wx,\ \hat{n} = y,\ \hat{z} = mn$" on the other hand, involves expectancies, e.g., about the manner in which the variables hang together under altered supporting conditions, which differ from the expectancies involved in the hypothesis "$\hat{m} = w,\ \hat{n} = xy,\ \hat{z} = mn$" and which are capable of experimental substantiation. Thus if we were to discover that under the influence of a certain drug, the empirical relation in the present example became changed to $\hat{z} = y\sqrt{wx}$, we can readily interpret the result by inferring that $\hat{m} = wx,\ \hat{n} = y$ and that the drug changes the manner by which m enters into the determination of z to $\hat{z} = n\sqrt{m}$. This is a more parsimonious interpretation of the drug result than any which can be made under the hypothesis that $\hat{m} = w,\ \hat{n} = xy$.

Once we realize, in fact, that the structure of a system of HC's is in itself a predicate sufficient to carry a set of existential operators, we become aware that it is not even necessary for the hypothesis to specify in quantitative detail the functional relations by which the variables are connected. It is sufficient that the hypothesis merely designates for each construct the other variables of which it is an immediate function and those which, in turn, the construct immediately helps determine. Thus if we treat the mediation variables of Hull's system not as IV's but as HC's (which, of course, the excess meaning introjected by Hull reveals them to be), we are, if we wish, able to conceive D, ${}_SH_R$, etc. as real attributes of the organism, not necessarily burdened with specific neurophysiological content, whose exact quantitative relations to each other and to extra-organismic variables still await experimental clarification, and yet variables whose structural assumptions generate testable predictions, especially when determination of the more detailed nature of these interrelations is regarded as a major experimental objective. Recognition that the structure of causal linkages among a set of mediation variables is independent of the specific quantitative nature of the relations frees us from preoccupation with excessive quantification of empirical relations during the early phases of a science, enabling us to identify the major immanent

---

[11] This point seems to have been anticipated by MacCorquodale and Meehl. "A hypothetical construct which seems inherently metaphorical may involve a set of properties to which hitherto undiscovered characteristics of the nervous system correspond. So long as the propositions about the construct are not stated in the terms of the next lower discipline, it is always a possibility that the purely formal or relational content of the construct will find an isomorphism in such characteristics. For scientific theories this is enough, since here, as in physics, the associated mechanical imagery of the theorist is irrelevant" (MacCorquodale & Meehl, 1948, p. 106; cf. also "The Nomological Net," Cronbach & Meehl, 1955, p. 290)

determinants of behavior prior to discovery of their exact relations to their causal antecedents and observable consequences.

References

Bergmann, G. (1951). The logic of psychological concepts. *Philosophy of Science*, *18*, 93–110.

Bergmann, G. (1953). Theoretical psychology. *Annual Review of Psychology*, *4*, 435–458.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.

Feigl, H. (1945). Operationism and scientific method. *Psychological Review*, *52*, 250–9.

Feigl, H. (1950). Existential hypotheses. *Philosophy of Science*, *17*, 35–62.

Ginsberg, A. (1954). Hypotheticca constructs and intervening variables. *Psychological Review*, *61*, 119–131.

Hull, C. L. (1943). *Principles of behavior.* New York: Appleton-Century.

Kendler, H. H. (1952). "What is learned?"—a theoretical blind alley. *Psychological Review*, *59*, 269–277.

MacCorquodale, K., & Meehl, P. E. (1948). On a distinction between hypothetical constructs and intervening variables. *Psychological Review*, *55*, 95–107.

Marx, M. H. (1951). Intervening variable of hypothetical construct? *Psychological Review*, *58*, 235–247.

Maze, M. H. (1954). Do intervening variables intervene? *Psychological Review*, *61*, 226–234.

Seward, J. P. (1955). The constancy of the I-V: A critique of intervening variables. *Psychological Review*, *62*, 155–68.