

DETERMINISTIC AND PROBABILISTIC THEORIZING IN STRUCTURAL LEARNING

REINHARD HILKE, WILHELM F. KEMPF & JOSEPH M. SCANDURA

1. Introduction

In the previous chapters a variety of formal models of the teaching-learning process have been presented. With one major exception, however, all of the models have a largely stochastic foundation. The structural learning theory (SCANDURA, 1973a, forthcoming, also this volume) is based (partly) on the assumption that deterministic theorizing about teaching and learning might actually provide a more useful first step. Moreover, it was suggested that there are close relationships between deterministic and probabilistic theories which might increase both the generality and the usefulness of both types of theory. The present contribution is designed to put these insights and conjectures regarding deterministic and probabilistic theories on a firmer theoretical base. In particular, the goal is to identify and illustrate conditions that allow for probabilistic extensions of the deterministic theory of structural learning.

The preparation of this report was supported by IPN and, in part, by N.I.H. Grant HD 09185-01.

2. Criteria for Probabilistic Theories

We begin with the general assumption that any formalization of a teaching and learning theory should:

- 1) make the theoretical concepts and assumptions (of the underlying theory) precise and
- 2) establish unambiguous relationships between the theory and methods of testing it.

In order to accomplish the above, each concept and assumption of the formalized theory should have a psychological equivalent. Otherwise, it would be impossible to determine whether a refutation of the theory is a refutation of arbitrary or of psychological content.

In particular for probabilistic theories it must be possible to

- a) test the structure of the theory,
- b) estimate the parameters, and
- c) make comparisons among parameters (if relevant) in a methodologically satisfactory manner (KEMPF, 1976).

This requirement holds irrespective of whether the parameters of the models (theories) are probabilities, as in models involving binomial distributions or in the probabilistic automaton model by SUPPES & MORNINGSTAR (1972), or whether the probabilities are reflected indirectly in terms of parameters, as in models involving the poisson distribution or in the linear logistic test model (LLTM) (COX, 1968; FISCHER, 1973, 1976, also this volume).

To perform tests of structure in stochastic theories, and especially to make comparisons among parameters, it is essential that the theory satisfy the condition of specific objectivity (RASCH, 1961): If a probabilistic theory is postulated for a given population of events (contacts of persons on items under specified testing conditions), it is postulated to hold for each event from this population. Correspondingly, statistics to be used in testing the theory must similarly be independent of sampling. Sampling otherwise introduces interfering conditions

(KEMPF, 1974, 1977) which make it possible to maintain any theory irrespective of empirical findings (cf. HOLZKAMP, 1968). For instance, consider a simple stochastic model which assumes that all subjects learn with the same probability. Although this assumption can be shown to be inconsistent with data in most experiments (cf. SCANDURA, 1971), it is possible to show this only insofar as such a model (cf. BOWER, 1961) satisfies specific objectivity. On the other hand, if a theory assumes that there are individual differences and there is no way to separate them out (e.g., to estimate learning parameters independently of individual differences parameters), then contradictory results can always be attributed to sampling. One can include general learning effects only if one can separate out the effects of individual differences (FISCHER, 1971; SCHEIBLECHNER, 1972).¹⁾

The above example makes clear one important limitation of classical approaches to the study of learning, namely that they neglect individual differences. Similarly, classical approaches to testing ignore structural relationships among test items. As noted by SCANDURA & DURIN (1971) this leads to statements such as "on the average, he should get eight out of ten items correct." It is essential, we think, for any viable theory to take both individual differences and structural relationships into account. Only in this way will it be possible to say anything about individual behavior in specific situations.

Notice, however, that the restrictions imposed by the above conditions (esp. specific objectivity) place strict limits on (stochastic) model construction (cf. RASCH, 1965; FISCHER, 1974).

Another limitation of current formalized learning theories is

1 It is important to qualify the above. Independence of sampling is necessarily relative, for deterministic as well as stochastic theories. If a theory is postulated to be valid for a population of events and there is a subpopulation of events for which the theory does not hold (while it holds for the rest of the events), the precision of tests of the theory will depend on the percentage of events from this subpopulation included in the sample.

that they often deal with a narrow range of phenomena (e.g., learning of nonsense syllables). Moreover, there is often little relationship among such "miniature theories." The lack of such apparent relationship could make the task of integrating phenomena practically insurmountable.

3. Arguments in Favour of Deterministic Theorizing

Although knowledge is necessarily partial, this is not the only approach to understanding complex human behavior. SCANDURA, for example, has had promising success with deterministic partial theories which allow for systematic enrichment. In constructing such partial theories, it is essential that one keep in mind the requirements of a broader reality. Thus, for example, a partial theory of phenomenon A will be extendable to a partial theory of phenomenon B just to the extent that the partial theory of A is compatible with the requirements for an adequate theory of B. In effect, while the limits of such enrichment cannot be predetermined the success achieved to date suggests that it may be easier to construct comprehensive theories of teaching and learning on a deterministic basis than on a probabilistic basis.

There is, however, an important argument against the use of deterministic theories in psychology. It is commonly felt that they impose too strict restrictions on data (e.g., LORD & NOVICK, 1968) and, hence, are "unrealistic." (According to FISCHER (1968, p. 73) a model is unrealistic if it can be refuted by almost any data.)

Nonetheless, one cannot reject deterministic theorizing a priori because of the empirical assumption that the data available to psychologists are essentially random (i.e., subject to unidentified and unidentifiable influences). To the extent that psychologists act on this presupposition, experience can never be gained with deterministic theories and one is thus in

danger of creating a circle (i.e., there is no way to reject the above assumption that data are random, even if it is false). Such a circle can be avoided only if data can be shown to be independent of a priori theoretical commitment. Since this implies the use of both probabilistic and deterministic approaches, deterministic theorizing can be rejected (if at all) only on an a posteriori basis (cf. HILKE forthcoming).

Reviewing the psychological literature we find that implications of adopting a deterministic approach in constructing theories and in their testing is rarely discussed. This suggests that such experience has rarely been obtained. Nonetheless, all deterministic theories have a major limitation when it comes to empirical confirmation: deterministic prediction and testing is possible only under idealized conditions (SCANDURA, 1971, this volume).

In order to clarify what is implied by such a statement, it is important to consider the nature of scientific laws.

Discussions in the philosophy of science typically, and independently of the point of view, refer to the natural sciences (i.e., physics), partly to demonstrate the diversity of other disciplines and partly to demonstrate their essential communality. In particular, such communalities are often used to qualify a discipline as an empirical one.

In attempting to demonstrate the empirical nature of a discipline, the laws of physics are frequently interpreted in different ways. In psychology, for example, universal statements of the form $(\forall x(Fx \supset Gx))$ can and have been interpreted in two different ways:

- 1) As assumptions about the relation between variables.
- 2) As rules which determine the relation between variables. If the rule concerns the relation between observables and theoretical entities it is called a rule of correspondence (cf. HARNATT, 1975).

This distinction is useful, though philosophers themselves have not yet even succeeded in making a sharp distinction between

such fundamental notions as law-like statements and statements of "accidental" universal cooccurrence. The only criteria for which there is general agreement are:

- a) law-like statements are non-analytic, and
- b) law-like statements are essentially universal

(KUTSCHERA, 1972, p. 329ff).

Such universal empirical statements cannot be verified; their logical structure allows only for falsification. However, assumptions can be regarded as more or less corroborated.

The most common type of interpretation in contemporary psychology, that a statement of the form $(Fa \wedge Ga)$ is sufficient to falsify the statement $(\forall x(Fx \supset Gx))$, is due perhaps in part to superficial and incomplete analogies to physics and specifically to the desire to have theories which can easily be falsified.

The all pervasive nature of stochastic theorizing in behavioral science; the general lack of theories, which are both formal and comprehensive; and the "naive" empiricism which characterizes much of contemporary psychology, have also contributed in this regard.

On the other hand, if statements are interpreted as rules, then they can not be tested independently of the total theoretical systems of which they are a part. (In psychology, of course, many so-called "theories" consist simply of isolated rules expressing relationships between observables.) Indeed, rules of correspondence lead to testable predictions only in combination with other assumptions and rules and, therefore, they cannot be falsified at all; one can only falsify the total system of which they are a part.

In physics, for example, rules of correspondence in combination with other rules and assumptions provide a basis for deriving statements about observables. Such derived statements, however, are only logical consequents. The empirical content of a theory derives from the assumption that the assumed entities, contained in those theoretical assumptions and rules from which the statement is derived, are the only ones necessary to ex-

plain the observed phenomena. Therefore, if a derived statement is not supported empirically, the system can be maintained by assuming that not all relevant entities are included. In this case, the theorist is obligated to enrich the theory by adding new assumptions and/or rules which make it possible to account for the deviations. This is in fact what physicists had to do with respect to the force laws in arriving at the principle of linear superposition. "Dieses Prinzip ist - wie die protophysikalischen Prinzipien - kein Meßergebnis der Physik, es ist auch keine Hypothese wie die Kraftgesetze, es eröffnet vielmehr erst die Möglichkeit, weitere Kraftgesetze hypothetisch aufzustellen, wenn die bisherigen die Beobachtungen nicht erklären" (LORENZEN & SCHWEMMER, 1973, p. 172). As DUHEM has argued for the force laws in 1906, only total systems can be compared with empirical findings.

In view of the above, one might think that interpreting universal empirical statements as assumptions would have the basic advantage of allowing for unambiguous falsification. Unfortunately, however, even this is not true. One can never be sure, whether a discrepancy between prediction and observation is due to inadequacies in the theory or in the theory on which observation is based. Moreover, statements about observables cannot either be verified or falsified irrefutably on the basis of a finite set of observations (cf. KUTSCHERA, 1972, p. 501).

In effect, a theory and its corresponding observation theory must necessarily be viewed as a total system, and hence it is not possible to falsify part of it only. In physics this objection is practically irrelevant because physics allows relatively precise measurement, although even here there is still room for rational doubt. In psychology, particularly in those areas which deal with complex human behavior, observations are much less precise. This relative lack of precision leaves correspondingly greater room for rational doubt, and hence a strict application of the principle of falsification to universal empirical statements obviously cannot be allowed.

As a consequence of this argumentation it follows that one has not only to specify the domain to which a theory applies, but also to specify the conditions under which it can be tested empirically (i.e., the conditions under which the assumed entities contained in those theoretical assumptions and rules, from which the statements about the observed phenomena are derived, are the only ones necessary to explain the data). These conditions are what we refer to as "idealized conditions" (SCANDURA, 1971, p. 26). Although the real world of observation (e.g., the classroom) is not "idealized," idealized conditions can be approximated in varying degrees in laboratory situations. One should note, however, that this implies accepting certain statements about observables as true.

Once having accepted that the empirical test of a deterministic partial theory can be performed in the laboratory only, it is obvious that such a test does not require the use of any statistical methods. If a scientist fails to establish idealized conditions in the laboratory, no statements about whether or not the theory holds are possible. In particular, it is impossible to "prove" a deterministic theory by rejecting the null hypothesis that the data are purely random. Similarly, one cannot "prove" that the linear logistic test model (LLTM) (FISCHER, 1973, this volume) holds, for example, by a statistically significant multiple correlation between the item- and basic-parameters.

4. The Weakened Form of the Structural Learning Theory

In view of the above restrictions on the applicability of any deterministic theory, even laboratory support is not sufficient. To have didactic relevance, for example, a deterministic theory of learning must be shown to be applicable in the everyday world of the classroom. Since pedagogic reality rarely satisfies idealized conditions, deterministic theories of teaching

and learning accordingly must be weakened in order to apply. Specifically, it would be desirable to have some explicit way to account for the effects of deviations from idealized conditions while, at the same time, retaining as much of the theoretical and empirical content as possible. Trivially, in weakening the theory of structural learning it is necessary to retain the essentials like, for example, the possibility of assessing behavior.

Before deciding how to weaken the structural learning theory it is necessary to first ask what one would like to accomplish. In view of the nature of the theory, any reasonable weakening should still provide: information about which rules are not known by a subject and for which there exists a substantial probability of failure, corresponding information about needed instruction, and specific information about deviations from the ideal.

Internal detail aside, the following rule of correspondence between theoretical and observable statements is essential in the SCANDURA theory (e.g., 1973a, Chapter 9):

$$(1) \quad \left\{ \begin{array}{l} \underline{S} \text{ knows and has} \\ \text{a rule available} \end{array} \right\} \iff \left\{ \begin{array}{l} \underline{S} \text{ uses the rule successfully} \\ \text{when needed to solve problems} \end{array} \right\}$$

In any probabilistic extension of the theory, this relation must be weakened. One possibility would be to drop the dichotomy "know"/"not know" and to introduce a quantitative dimension "a rule is known more or less":

$$(2) \quad \left\{ \begin{array}{l} \underline{S} \text{ 1 "knows" a rule} \\ \text{better than } \underline{S} \text{ 2} \end{array} \right\} \iff \left\{ \begin{array}{l} \underline{S} \text{ 1 has a higher probability}^1 \\ \text{of using a rule successful-} \\ \text{ly when needed than } \underline{S} \text{ 2} \end{array} \right\}$$

¹ Notice that this rule is between theoretical entities only (i.e., "knows" and "probability"). Hence, rules of correspondence are needed to connect probabilities and observables. Such connections are provided by stochastic measurement models.

It is this possibility, for example, which provides the basis for most applications of the linear logistic test model (LLTM) to the analysis of thinking and learning (SCHEIBLECHNER, 1972; FISCHER, 1973; SPADA, 1976).

A second possibility is to weaken the rule of correspondence (1) so that the arrow points in one direction only:

- 1) Let knowing a rule imply successful use of the rule or
- 2) let successful use of a rule imply knowledge of a rule.

In the former case the assumption of knowing a rule is falsified if the subject fails to use the rule successfully on any one problem. Conversely the S would be assumed to know a rule only where he succeeds on all items. These requirements may be reasonable in the case of simple tasks, where the probability of guessing correctly is high. However, they seem unrealistically stringent when the probability of guessing is low, as in most school learning tasks, and where the possibility of "careless errors" is considerable.

For our purposes, the second weakening seems more reasonable. Specifically:

- (3) $\left\{ \begin{array}{l} \underline{S} \text{ knows and has} \\ \text{a rule available} \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \underline{S} \text{ uses the rule successfully} \\ \text{when needed to solve problems} \end{array} \right\}$
- (4) $\left\{ \begin{array}{l} \text{It is not the case} \\ \text{that } \underline{S} \text{ knows and has} \\ \text{a rule available} \end{array} \right\} \rightarrow \left\{ \begin{array}{l} \text{It is not the case that } \underline{S} \\ \text{uses the rule successfully} \\ \text{when needed to solve problems} \end{array} \right\}$

Hence, the probability of success is 0 when S does not know or does not have available a rule. Nothing is implied if S does know a rule. In this case we assume only that S has a positive probability of success $0 < p \leq 1$, with the equality $p = 1$ holding under idealized conditions.

Assumptions (3) and (4) differ from assumption (2) in that the former build on assumption (1) (i.e., that the Scandura theory holds in the laboratory under "idealized conditions") whereas assumption (2) does not. As a consequence, assumptions (3) and

(4) allow for deterministic analysis and, in this regard, are consistent with the structural learning theory. In particular, the deterministic theory is a special case.

Adopting the generalization based on assumptions (3) and (4) makes it possible to take two aspects of individual differences into account:

- 1) That part of the data which is consistent with the stronger assumption (1) of the deterministic theory can be used to specify which rules are and are not known by particular Ss in the sample - namely, those Ss who either succeed or fail uniformly on all of the problems associated with given rules.
- 2) The remainder of the data can be used to evaluate individual differences with respect to other combinations of rules and Ss. These individual differences will reflect individual deviations from idealized conditions rather than just "latent abilities" in the traditional sense.

Similarly, the data can be used to provide information concerning the difficulties of individual rules, and to make overall judgments about the degree of deviation of the data from the ideal.

Such information can be used in a variety of ways:

- 1) It can be used to make specific decisions about instruction needed (e.g., about which subjects need which information).
- 2) It can be used to identify possible overall weaknesses in instruction (e.g., about which rules require more explicit instruction).
- 3) It can be used to decide whether deviations from the ideal are due: to incompleteness of the theory, to improper implementation of technologies based on the theory or to miscellaneous factors pertaining to implementation (i. e., professional "know-how" - cf. SCANDURA, 1973b, p. 9).

The diagram in Figure 1 summarizes the flow of information that is envisioned here among the teacher, the curriculum con-

structor and the theorist/experimenter. For more detail see HILKE, forthcoming).

TEACHER

- 7 ... Teacher's unit plan.
- 8 ... Actual instruction and testing.
- 9 ... Has criterion been achieved?
- 10 ... Individualized instruction.
- 11 ... Are there overall inadequacies in teaching?
- 12 ... Do data deviate substantially from ideal?

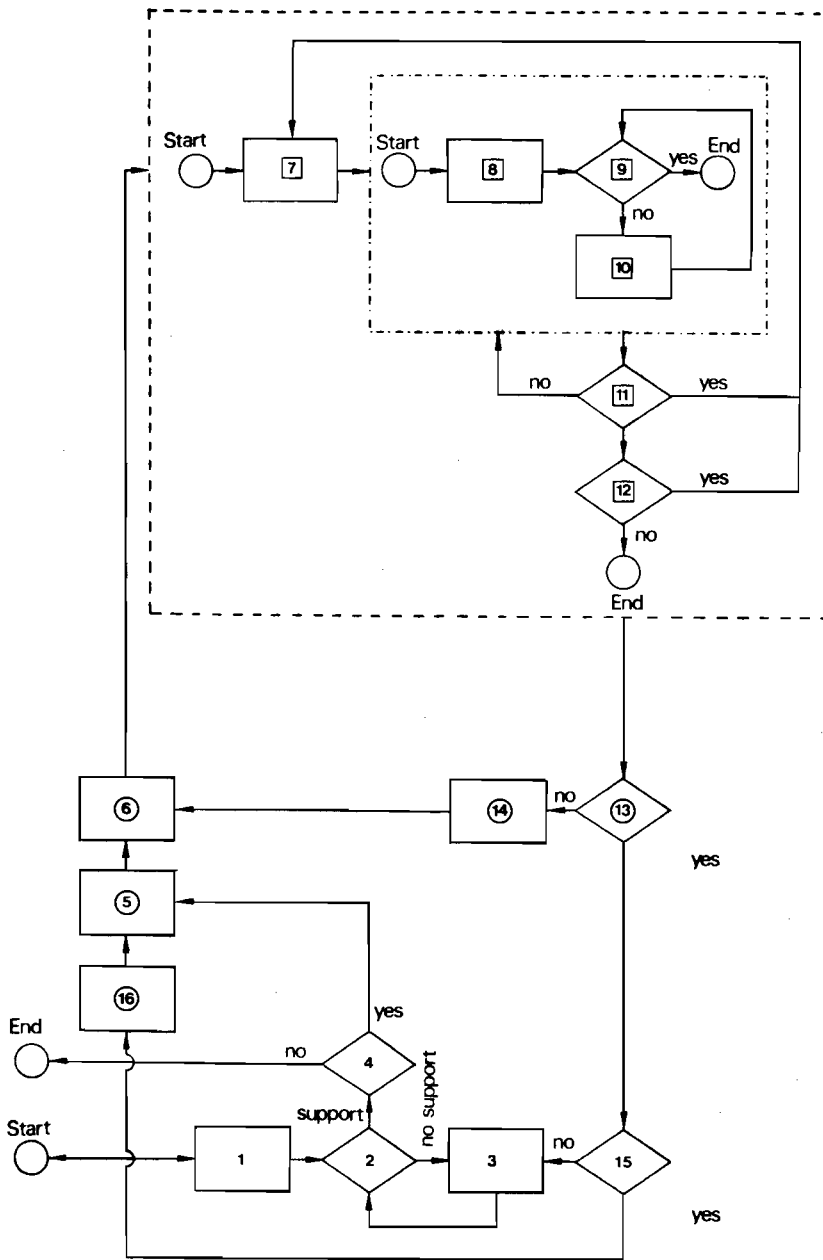
CURRICULUM CONSTRUCTOR

- 5 ... Curriculum and test construction.
- 6 ... Overall plan for implementation and implementation.
- 13 ... Has the technology/theory been properly applied in curriculum construction?
- 14 ... Revise curriculum.
- 16 ... Improve overall plan for implementation
(including relevant educational research) and implement.

THEORIST/EXPERIMENTER

- 1 ... Development of a subject matter specific structural learning theory by applying a partial structural learning theory to some specific subject matter.
- 2 ... Is test of the theory in the laboratory successful?
- 3 ... Enrichment of refutation and development of a new theory.
- 4 ... Can needs of schools be met by applying theory and can adequate resources be obtained for same?
- 15 ... Is theory supported by further tests in laboratory?

Figure 1



In order to avoid inconsistencies in this flow of information and in the decisions based on it, it is essential that the teacher, the curriculum constructor, and the theorist work closely together and have a common methodological base. Thereby, the teacher plays a central role. Information provided by the weakened form of the deterministic theory is sufficient for the teacher. Thus, for example, the teacher can identify which students need what information, and thereby meet the needs of individual students by counting the frequencies b_{voj} of success of individual students (v) on items associated with particular rules (j). Similarly, information (not necessarily just simple counting) about overall performance on rules which have been taught, provides a basis for improving the teacher's overall lesson plan. On the other hand, in order to gain information concerning the strongest form of the deterministic theory, the theorist must obtain data under idealized conditions. Effectively, the curriculum constructor must make use of both kinds of information (i.e., information pertaining to both the weakened and the strong theory), and specifically the relations between them provided by the above probability statements.

5. Probabilistic Models

In order to accomplish this, there is a need for probabilistic measurement models which are both consistent with the weakened form of the structural learning theory, applicable in realistic educational settings, and compatible with the type of scores on which the teacher may base his decisions on the classroom. One class of stochastic measurement models follows directly by generalization of a theorem by RASCH (1965) from the type of data just mentioned. The models in this class are defined by the following item characteristic.¹

¹ To the knowledge of the authors this class of models has so far been mentioned only in unpublished papers by Kempf and by Stegelmann. Further statistical details of the models have been worked out by KEMPF (1976). The results will be published in STEGELMANN (forthcoming).

$$(5) \quad p\{a_{vi} | (f_{ij})\} = \frac{\exp\{a_{vi} \sum_j f_{ij} (\xi_{vj} - \sigma_{ij})\}}{1 + \exp\{\sum_j f_{ij} (\xi_{vj} - \sigma_{ij})\}}$$

where $a_{vi} = \begin{cases} 1 & \text{if person } v \text{ solves problem } i \\ 0 & \text{if not} \end{cases}$

$f_{ij} = \begin{cases} 1 & \text{if rule } j \text{ is necessary for the solution of} \\ & \text{problem } i \\ 0 & \text{if not} \end{cases}$

$\xi_{vj} =$ person v 's "ability" with respect to rule j

$\sigma_{ij} =$ the difficulty of rule j in problem i .

(Letting the difficulty σ_{ij} of rule j depend on problem i allows for adding more internal structure to rule j in accordance with the structural learning theory, cf. SCANDURA 1973a, Chapters 7 and 9. Moreover, allowing more than one rule j to enter into the solution of a given problem corresponds roughly to allowing rules to operate on other rules in the solution of problems (same reference), though the correspondence is not exact.)

The model in Eq. 5 is a direct generalization of the Rasch model. Specifically, it is a generalization which takes into account individual abilities with respect to specific rules and where item difficulties are represented in terms of difficulties of specific rules.

The generalized model has the same desirable characteristics as the Rasch model. In particular, there exist minimally sufficient statistics for the parameters, which can be used as a basis for specifically objective comparisons. Thus, if we assume stochastic independence of items and individuals, the likelihood of

the data matrix $((a_{vi}))$ ($v = 1, 2, \dots ; i = 1, 2, \dots$) becomes

$$(6) \quad p \{((a_{vi})) | ((f_{ij}))\} = \frac{\prod_v \prod_i \exp \{a_{vi} \sum_j f_{ij} (\xi_{vj} - \sigma_{ij})\}}{\prod_v \prod_i (1 + \exp \{\sum_j f_{ij} (\xi_{vj} - \sigma_{ij})\})} \\ = \frac{\exp \{\sum_v \sum_j b_{voj} \xi_{vj} - \sum_i \sum_j b_{oij} \sigma_{ij}\}}{\prod_v \prod_i (1 + \exp \{\sum_j f_{ij} (\xi_{vj} - \sigma_{ij})\})}$$

where $b_{voj} = \sum_i a_{vi} f_{ij}$ and $b_{oij} = \sum_v a_{vi} f_{ij}$.

Since $p \{((a_{vi})) | ((f_{ij}))\}$ is the same for all matrices with the same two marginal matrices $((b_{voj}))$ and $((b_{oij}))$, the joint distribution of the marginal matrices is

$$(7) \quad p \{((b_{voj}), (b_{oij}))\} = \left[\begin{array}{c} ((b_{voj})) \\ \\ ((b_{oij})) \end{array} \right] p \{((a_{vi})) | ((f_{ij}))\}$$

where the coefficient on the right denotes the number of possible data matrices compatible with the marginals. Summing (7) over all $((b_{voj}^*))$ compatible with $((b_{oij}))$ we obtain the likelihood function

$$(8) \quad p \{((b_{oij}))\} = \sum_{((b_{voj}^*))} \left[\begin{array}{c} ((b_{voj}^*)) \\ \\ ((b_{oij})) \end{array} \right] \\ \frac{\exp \{\sum_v \sum_j b_{voj}^* \xi_{vj} - \sum_i \sum_j b_{oij} \sigma_{ij}\}}{\prod_v \prod_i (1 + \exp \{\sum_j f_{ij} (\xi_{vj} - \sigma_{ij})\})}$$

Now, by dividing (6) by (8) we obtain the conditional likelihood of the data matrix, given the marginal matrix $((b_{oij}))$

$$(9) \quad p \{((a_{vi})) | ((f_{ij})), ((b_{oij}))\} = \frac{\exp \{ \sum_v \sum_j b_{voj} \xi_{vj} \}}{\sum_{((b_{voj}^*))} \left[\begin{array}{c} ((b_{voj}^*)) \\ ((b_{oij})) \end{array} \right] \exp \{ \sum_v \sum_j b_{voj}^* \xi_{vj} \}}$$

A symmetrical argument yields the conditional likelihood

$$(10) \quad p \{((a_{vi})) | ((f_{ij})), ((b_{voj}))\} = \frac{\exp \{ - \sum_i \sum_j b_{oij} \sigma_{ij} \}}{\sum_{((b_{oij}^*))} \left[\begin{array}{c} ((b_{voj})) \\ ((b_{oij}^*)) \end{array} \right] \exp \{ - \sum_i \sum_j b_{oij}^* \sigma_{ij} \}}$$

Finally, we may divide (6) by (7) and obtain the conditional likelihood of the data matrix, given both marginal matrices

$$(11) \quad p \{((a_{vi})) | ((f_{ij})), ((b_{voj}), ((b_{oij}))\} = \frac{1}{\left[\begin{array}{c} ((b_{voj})) \\ ((b_{oij})) \end{array} \right]}$$

On the basis of (10) we may estimate the rule-difficulty parameters independently of the individual parameters which have been replaced by something observable, namely by the statistics b_{voj} . Furthermore, (9) can be used to compare the individual parameters independently of the rule-difficulty parameters which are replaced by the frequencies b_{oij} . Finally, (11) allows for checks of the model which are independent of all of the parameters.

For the practical use of the model (5) however, there is one major limitation: the general, and abstract formulation of rule difficulties relative to items σ_{ij} does not allow for a direct estimation of the rule-difficulty parameters, but only for the estimation of item difficulties $\sigma_{io} = \sum_j f_{ij} \sigma_{ij}$. If difficulty

parameters pertaining to component rules are to be estimated, additional restrictions must be imposed on the model.

For example, if we assume that the difficulty of a given rule is the same on all problems ($\sigma_{ij} = \sigma_{oj}$ for all $i=1,2,\dots$), then the σ_{oj} parameters play the role of parametrization constants only. With the definition $\xi_{vj}^* = \xi_{vj} - \sigma_{oj}$ we may replace the expression $(\xi_{vj} - \sigma_{ij})$ in (5) by ξ_{vj}^* , and hence cannot speak any longer about overall difficulties of the rules. Since rule difficulty is thereby confounded with individual ability, all statistical variation of the data is explainable solely in terms of individual differences with respect to the various rules.

Another situation arises, if we take into account the internal structure of the rules corresponding to the σ_{ij} in accordance with the structural learning theory. Specifically, provision may be made for different paths through an algorithm or rule and for different computations associated with particular paths (SCANDURA, 1973a, Chapter 9). In this case the rule difficulty parameters σ_{ij} may be expressed in terms of difficulties of component (atomic) rules. Thus, specific difficulty parameters η_h are introduced pertaining to the component (atomic) rules (h) and weights q_{ijh} are introduced describing how many times the component rules (h) are needed in application to rule j in solving problem i. We then may replace the σ_{ij} parameters by $\sum_h q_{ijh} \eta_h$.

An interesting special case arises, if we consider a domain of problems which are all associated with the same rules so that $f_{ij} = f_{oj}$, for all $i=1,2,\dots$. The model then reduces to the linear logistic test model

$$(12) \quad p\{a_{vi} | (d_{ih})\} = \frac{\exp\{a_{vi} (\xi_v - \sum_h d_{ih} \eta_h)\}}{1 + \exp\{\xi_v - \sum_h d_{ih} \eta_h\}}$$

with definitions $\xi_v = \sum_j f_{oj} \xi_{vj}$ and $d_{ih} = \sum_j f_{oj} q_{ijh}$. The "cog-

nitive operations" in the study by SPADA, FISCHER & HEYNER (1974) can be understood as such component rules with a domain of problems, the solution algorithm of which consists of just one rule.

By way of conclusion, we make explicit the relation between this model and the weakened deterministic structural learning theory. In general, given data of the type indicated above, some will be compatible with the weakened theory and some will not be. In particular, the behavior of those subjects who fail uniformly on all items, associated with a rule or path, will be consistent with the weakened theory.

The above stochastic theory fills the gap by dealing with the remainder of the data (i.e., where subjects succeed on some (or all) items associated with a rule or path). This "imperfect" data is used to determine parameter values.

Moreover, the overall proportion of the data which deviates from deterministic predictions provides a useful measure of average deviation of the empirical situation from the ideal. As indicated in SCANDURA (this volume), such information could be used to compare alternative empirical situations (as to their relationships to the ideal), and to extrapolate to new situations on the basis of the degree to which the new and old situations differ.

References:

- ANDERSEN, E.B.: Conditional inference for multiple-choice questionnaires. *British Journal of Mathematical and Statistical Psychology*, 1973, 26, 31.
- BOWER, G.: Application of a model to paired-associate learning. *Psychometrika*, 1961, 26, 255.
- COX, D.R.: *The analysis of binary data*. London: Methuen, 1969.
- DUHEM, P.: *La theorie physique, son object et sa structure*. Paris: 1906.
- FISCHER, G.H.: Neue Entwicklungen in der psychologischen Testtheorie. In FISCHER, G.H. (Ed.): *Psychologische Testtheorie*. Bern: Huber, 1968.
- FISCHER, G.H.: Einige Gedanken über formalisierte psychologische Theorien. *Psychologische Beiträge*, 1971, 13, 376.
- FISCHER, G.H.: The linear logistic test model as an instrument in educational research. *Acta Psychologica*, 1973, 37, 359.
- FISCHER, G.H.: *Einführung in die Theorie psychologischer Tests*. Bern: Huber, 1974.
- FISCHER, G.H.: Linear logistic models for the description of attitudinal and behavioural changes under the influence of mass communication. In KEMPF, W.F., & REPP, B.H. (eds.): *Some mathematical models for social psychology*. Bern: Huber, 1976 (in print).
- HARNATT, J.: *Psychologie als Wissenschaft*. Unpublished paper, 1975.
- HILKE, R.: Zwei neuere Ansätze aus dem Bereich Denken und Problemlösen in ihrer Bedeutung für die Entwicklung lehrzielorientierter Tests (forthcoming).
- HOLZKAMP, K.: *Wissenschaft als Handlung*. Berlin: De Gruyter, 1968.
- KEMPF, W.F.: Basisprobleme der Diagnostik der Aggressivität. In: ECKENSBERGER, L. (Ed.): *Bericht über den 28. Kongreß der Deutschen Gesellschaft für Psychologie in Saarbrücken 1972*. Göttingen: Hogrefe, 1974.
- KEMPF, W.F.: Parameterschätzung in einem mehrdimensionalen strukturellen logistischen Testmodell. Unpublished report to the Max-Planck-Institut für Bildungsforschung, Berlin, 1976.

- KEMPF, W.F.: Dynamic models for the measurement of "traits" in social behavior. In KEMPF, W.F., & REPP, B.H. (eds.): Some mathematical models for social psychology. Bern: Huber, 1977.
- KUTSCHERA, v.F.: Wissenschaftstheorie II. München: Fink, 1972.
- LORD, F.N., NOVICK, M.R.: Statistical theories of mental test scores. Reading, Mass.: Addison Wesley, 1968.
- LORENZEN, P., SCHWEMMER, O.: Konstruktive Logik, Ethik und Wissenschaftstheorie. Mannheim: Bibliographisches Institut, 1973.
- RASCH, G.: On general laws and the meaning of measurement in psychology. Proceedings of the Fourth Berkely Symposium on Mathematical Statistics and Probability, 1961, 5, 321.
- RASCH, G.: Kolloquium über Meßmodelle. 1965 (mimeo).
- SCANDURA, J.M.: Deterministic theorizing in structural learning. Three levels of empiricism. Journal of Structural Learning, 1971, 3, 21.
- SCANDURA, J.M.: Structural learning I. Theory and research. New York: Gordon and Breach, 1973a.
- SCANDURA, J.M.: Structural learning and the design of educational materials. Educational Technology, August, 1973b, 7-13.
- SCANDURA, J.M.: Human problem solving (forthcoming).
- SCANDURA, J.M., DURNIN, J.H.: Assessing behavior potential: Adequacy of basic theoretical assumptions. Report No. 62, Structural learning series, Graduate School of Education, University of Pennsylvania, 1971.
- SCHEIBLECHNER, H.: Das Lernen und Lösen komplexer Denkaufgaben. Zeitschrift für experimentelle und angewandte Psychologie, 1972, 19, 476.
- SPADA, H.: Modelle des Denkens und Lernens. Bern: Huber, 1976 (in print).
- SPADA, H., FISCHER, G.H., HEYNER, W.: Denkopoperationen und Lernprozesse bei der Lösung von Problemstellungen aus der Mechanik. In ECKENSBERGER, L. (Ed.): Bericht über den 28. Kongreß der Deutschen Gesellschaft für Psychologie in Saarbrücken 1972. Göttingen: Hogrefe, 1974.
- STEGELMANN, W.: Eine Theorie über Lösungen von Testaufgaben. Phil. Diss., Berlin (forthcoming).

SUPPES, P., MORNINGSTAR, M.: Computer Assisted Instruction at
Stanford 1966 till 1968. New York: Academic Press, 1972.