A DETERMINISTIC APPROACH TO RESEARCH IN INSTRUCTIONAL SCIENCE

JOSEPH M. SCANDURA
University of Pennsylvania

ABSTRACT

A deterministic alternative to probabilistic theory and research is proposed and justified on both rational and empirical grounds. Direct tests of deterministic behavior theories are demanding but they are not only possible but are shown to be quite feasible. The main requisite is that of realizing suitable idealized test conditions. Some deterministic theories have been tested under idealized behavior conditions with a level of empirical support which goes far beyond what is normally found in behavioral research. Moreover, deterministic support in the laboratory is shown to necessarily imply probabilistic applicability in the real world—without field testing.

With the possible exception of testing and psychometrics, which drew at least part of their inspiration from educational needs, educational research traditionally has relied on other disciplines for its methodologies. Experimental research in education, for example, has drawn from a Fisherian experimental design and statistics that was developed initially to evaluate crop yields in agriculture. This methodology has led over the years to many-fold increases in farm production. (In education, and I will propose also in psychology, the results have been equivocal.)

The research methodology and illustrative studies that I want to describe in this paper have a theoretical/philosophical basis that is fundamentally different from this tradition. Although dated, perhaps the simplest general introductions to the theoretical/research and instructional design aspects of the general structural approach are given, respectively, in Scandura (1971) and (1973b). Scandura (1973a) provides perhaps the most complete formal treatment of the theory; refinements, important theoretical extensions and research are given in Scandura, et al (in press).

The present paper deals with only one part of this work, with the way in which the cognitive aspect of the structural theory is tested empirically, and with an important relationship between such tests and traditional methods of testing behavioral theories. The first treatment of similar concerns is included in my initial paper on structural learning (Scandura, 1971). Later and more extensive treatments have been given by Hilke, Kempf and Scandura (1976) and Scandura (1976), with the most complete treatment to appear in a forthcoming book (Scandura, et al, in press).

Theoretical and Methodological Preliminaries

Until recently, most theories and hypotheses in behavioral science have been inherently probabilistic. Thus, for example, we have theories which indicate that

---

1Preparation of this paper was supported in part by N.I.H. Grant HD-09185 to the author and in part by the MERGE Research Institute. The author's address is Graduate School of Education, University of Pennsylvania, Philadelphia, PA 19174.
DETERMINISTIC APPROACH TO RESEARCH

certain behavior will occur with some probability in given situations and theories which refer to average behavior in given classes of situations. In the literature, for example, we frequently read, on the one hand, such things as "the probability of behavior A is a function of the number of reinforcements or each subject learns on each trial with probability c" and, on the other hand, such things as "discovery learning (on the average) leads to more transfer than does expository learning."

In testing probabilistic theories and hypotheses, the essential data have been means and other group statistics. This works fine, and is appropriate, where the emphasis is on external, manipulable variables, or even where the reference is to diffuse underlying cognitive constructs. Modern information processing psychology, however, has a different theoretical base, one that is inherently deterministic. Theories and hypotheses of this type provide a potential basis for stating what individual subjects will do in particular experimental situations.

To date, however, the methods and experimental paradigms that have been used to test such theories, have been essentially the same as those used to test probabilistic theories. Suppose, for example, we want to test alternative information processing theories, each of which consists of some hypothesized rule (process/procedure). In this case, experiments are set up to determine which rule best accounts for average group behavior. In experimental psychology, we often use response latencies for this purpose. Related work in education generally deals with more complex tasks but uses less refined response measures.

There is a serious problem with this approach to theory verification. Just because a rule or process provides a viable account of average, group behavior, this doesn't say anything necessarily about how individuals perform — in particular, it doesn't say anything definitive about individual processes. Perhaps the clearest example of this goes back to the old controversy over whether learning is incremental or all-or-none. As it turned out, as you may recall, the same group predictions could be made using either assumption about individual processes. For the most part, the data were simply inadequate to answer the question.

Consider the philosophical implications of the above approach (to theory verification). It is a basic fact in the philosophy of science that deterministic theories cannot be tested in the same way as probabilistic theories. Whereas probabilistic theories can be tested under actual behavior conditions, because the presence of random error is always assumed, this is not the case with deterministic theories. Deterministic theories must be tested under idealized conditions. Thus, for example, the deterministic laws of mechanics in physics — involving say the inclined plane — hold only where friction, and other peripheral factors are not involved. One could, of course, devise alternative stochastic theories of such phenomena under nonidealized conditions, but the resulting theories would be far more complex and ad hoc than the relatively unified, simple, and aesthetic mechanics developed by Newton.

In order to get a clearer picture of the way in which each type of theory must be tested, consider Galileo's famous experiments at the leaning Tower of Pisa. In particular, suppose instead of dropping two iron balls of differing weights, that Galileo had dropped an iron ball and a feather. How different his results would have been. Moreover, if counselled by a Twentieth century behavioral scientist, Galileo might have been duped into comparing the average rates of fall of 100 iron balls and 100 feathers. In another age, indeed, such a study might well have precipitated an experimental rush to uncover the scientific laws governing the rates of fall of various types of droppings.

Although such things are not as common as they once were, there are still many who believe that they are not as common as they once were. There is no more direct way to test a theory of learning than by dropping two objects with differing weights. In this case, the result is clear: the iron ball falls faster than the feather.

Tests of De Novo Learning

It is my contention that this approach is possible, but that it is not without limitations. Indeed, much of the recent research in the area of learning and memory has focused on the role of experience in determining the rate of learning. The problem is how to combine the information obtained from these different methods in order to make a more accurate prediction of future performance.

Consider the case of a student who fails an exam. This student is then given a course on how to improve their study habits. The student is then retested on the exam and their performance improves. This suggests that the student has learned something new from the course. However, it could also be the case that the student simply practiced more for the exam or that they were simply more motivated to study. Thus, it is important to consider the role of experience in determining the rate of learning.

The importance of experience in learning is illustrated by the fact that the amount of information that a student can retain and recall varies depending on the amount of time they spend studying. This suggests that the role of experience in learning is not limited to the acquisition of new information, but also includes the retention and recall of previously learned information.

Let me descri...
Although such findings could conceivably have immediate interest, it is unlikely that they would have inspired Newton in his attempts at theoretical synthesis. Of much more direct relevance were Galileo’s actual experiments, which correspond more directly to what would happen under idealized conditions — say, where a feather and an iron ball are dropped in a perfect vacuum. The point to emphasize here is that the much simpler results associated with Galileo’s actual experiments, where gravitational force was essentially uninfluenced by other forces, had far greater and longer lived generality — far greater and longer lived than would alternative scientific laws concerning droppings.

One wonders in contemporary cognitive psychology and instructional science whether specific laws (rules, information processing theories) introduced to account for unconstrained average behavior could in the long run have far less generality, less generality that is than simpler deterministic laws determined under idealized behavior conditions.

The problem, of course, is to determine just what these idealized conditions might be. Ebbinghaus, for example, thought that nonsense syllables might provide a promising way to study human learning processes. After years of debate and research, including the decline of early structuralism and the rise of behaviorism, it is increasingly recognized that this was largely a false hope. The years of prior learning and experience, which each subject brings into any learning situation, affect learning in fundamental ways, even with respect to unfamiliar material.

Tests of Deterministic Behavior Theories Under Idealized Conditions

It is my contention that deterministic theorizing in behavioral science is not only possible, but that it is also feasible to realize idealized conditions in many behavioral situations.

Indeed, much of the Gagne inspired work pertaining to learning hierarchies makes this assumption implicitly. Consider, for example, the generalization that students who fail on subordinate tasks will also fail on relatively superordinate tasks. This generalization is a deterministic statement, one which moreover frequently holds in a near deterministic sense. The generalization, moving in the opposite direction, is more fallible. Thus, students who learn all of the prerequisites of a given task, as determined through a standard hierarchical task analysis, are not always able to solve the superordinate task. More accurately, standard hierarchical task analyses do not identify all important prerequisites. The missing prerequisites relate to higher order rules/ processes by which relatively lower order prerequisites may be combined and, or modified to provide a basis for solving the superordinate task.

The important thing for present purposes is to clarify and illustrate the methodological underpinnings of deterministic research. I should caution you, however, that what I shall have to say deals with only a few aspects of the problem. An extension and further development of these ideas will appear elsewhere (Scandura, et al, in press).

Consider some of our recent problem solving experiments (Scandura, 1973; 1974). These experiments were run under idealized, memory-free conditions. That is, the experiments were run under conditions designed to eliminate the effects of memory, processing capacity, and response latency. The latter were effectively factored out of the experiments.

Let me describe one of these experiments briefly (Scandura, 1974). In this
experiment, children were taught how to trade objects of one kind for objects of another kind.

The card at the top of Figure 1, for example, denotes a rule (cognitive process) for exchanging paper clips for blue chips. In particular, if 2 paper clips are presented to a child, he is to trade for 3 blue chips; similarly, if presented with 3 paper clips, he is to trade for 4 blue chips. The composite card at the bottom denotes a composite rule in which trades are made in two steps. In the particular example shown, the child first trades paper clips for pencils by adding 2 extra paper clips and then white chips for the paper clips by adding 1 extra white chip. For discussion purposes, compatible simple rules are denoted \( A \rightarrow B \) and \( B \rightarrow C \) whereas corresponding composite rules are denoted \( A \rightarrow B \rightarrow C \).  

Given tasks of this type, suppose we teach a child an \( A \rightarrow B \) and a \( B \rightarrow C \) rule, and then present him with a certain number of A objects, and ask him to trade for the appropriate number of C objects. Under these conditions: (1) How many children will succeed on the A-C task? (2) If some of the children fail, what else do they need to know in order to succeed? (3) Is there something we can take for granted about the capabilities of all children—something which is basic, which is innate to all children, say, of the ages of 7 and above.

In an experiment conducted with children ranging from ages 7-9, we found that only 6 of 30 children were able to solve a new A-C task of this type without explicit training on an unc

order rule by which the other 12 children with a new pair of rules may serve as the 12 children who had seen before.

The results show that a-C task where

Several things of a solution (rule components together applying higher order rule problems where order rules, altho sufficient. Presumably universal capability attacking a problem.

With regard to mechanism which This mechanism be expressed in terms of a hypothesis: Given the ability of one or subject will use (does not have automatically sh (learning and rule derived rule is ad to the original g

To see how t into C objects, A objects into C objects, B operates on pair.

Is our postulate these subjects? A simple performance is involved. The subject shifts to the higher order rule (C rules) to gain higher level goal base and control, so the new problem is solved.

Could the an is much more
training on an underlying \( A \rightarrow C \) rule. Of the 24 who failed, half were taught a higher order rule by which arbitrary pairs of compatible rules (in which the output of one may serve as the input of the other) could be combined to form composite rules. The other 12 children were not given this training. Then, all of the children were presented with a new pair of compatible rules, \( A' \rightarrow B' \) and \( B' \rightarrow C' \), rules none of the children had seen before. Finally, they were presented with the corresponding \( A' \rightarrow C' \) task.

The results showed that all of the children who received training succeeded on the \( A' \rightarrow C' \) task whereas not one of the others did so.

Several things should be observed about this study. (1) Knowing the components of a solution (rule) for a task is not sufficient. The subject must know how to put the components together appropriately. (2) Integrating components is accomplished by applying higher order rules to lower order ones (e.g., component rules). Indeed, higher order rules seem to be both a necessary and a sufficient condition for solving problems where needed components are available. (3) Component rules and higher order rules, although behaviorally sufficient for problem solving, are not logically sufficient. Presumably, a human being must have some prior and presumably universal capability which tells him how and when the various rules are to be used in attacking a problem.

With regard to the third point, let us consider a basic, assumedly universal mechanism which is fundamental to my structural learning theory (Scandura, 1973). This mechanism tells how known rules interact in learning and performance and may be expressed in terms of three simple hypotheses. Hypothesis 1 (simple performance hypothesis): Given a goal (which the subject is attempting to achieve) and the availability of one or more rules, each of which generates the desired response, then the subject will use one of them. Hypothesis 2 (control shift hypothesis): If the subject does not have a rule immediately available for achieving his goal, control automatically shifts to the higher level goal of deriving such a rule. Hypothesis 3 (learning and reversion hypothesis): Once a higher level goal is satisfied, the newly derived rule is added to available knowledge (i.e., is learned) and control reverts back to the original goal.

To see how this mechanism operates, consider the task of converting \( A \) objects into \( C \) objects, and two students, one who enters a task knowing a rule for converting \( A \) objects into \( C \) objects and another who only knows rules for converting \( A \) objects into \( B \) objects, \( B \) objects into \( C \) objects, and the higher order composition rule which operates on pairs of compatible rules and forms composite ones.

Is our postulated control mechanism sufficient for predicting the performance of these subjects? The success of the first student (on \( A \rightarrow C \)) follows directly from the simple performance hypothesis. An available rule applies in the task situation, so he, therefore, uses it. The situation with the second student is only slightly more involved. The student first checks his available rules, and since none applies, control shifts to the higher level goal of generating one. In this higher level goal situation, the higher order composition rule applies, so it is used (i.e., applied to the \( A \rightarrow B \) and \( B \rightarrow C \) rules) to generate the composite rule \( A \rightarrow B \rightarrow C \). This composite rule satisfies the higher level goal; therefore, the \( A \rightarrow B \rightarrow C \) rules is added to the available knowledge base and control reverts to the original goal. At this point, the newly derived rule applies, so according to the simple performance hypothesis, it is used and the problem is solved.

Could the answer to how learning takes place be all that simple? Not quite. There is much more that might be said about such things as breaking problems into
DETERMINISTIC APPROACH TO RESEARCH

subproblems, processing time, how much information a person can keep in mind at the same time, and so on (for details see Scandura, 1973; Scandura, et al, in press). But for present purposes we can ignore these factors.

The central issue here pertains to internal external conditions under which the experiment was run. In running this experiment we took every feasible precaution to ensure that the subjects were indeed attempting to achieve the stated goals, that the requisite to-be-learned rules really were learned, and, perhaps most important, that the learned rules were available to each subject at the time of testing. The subjects were also given physical mnemonics to help them remember the critical rules, and all the time they needed. Under these “memory-free” conditions, the deterministic theory was shown to account almost perfectly for the performance of individual subjects on specific problems.

Now, of course, not all data are idealized. Most real world data come from nonidealized situations. This is certainly true, for example, in the classroom.

Probabilistic Generalization of Deterministic Findings

Nonetheless, and perhaps surprisingly, deterministic theories, and parallel data collected under idealized conditions, have an important advantage over probabilistic theories and corresponding group data. This advantage has to do with the generalizability of laboratory results to the real world of education.

Consider, first, the situation with regard to group data. In this case, results obtained in the laboratory may be generalized to the population from which the experimental samples have been drawn. Usually, this population is poorly defined and one can never be sure whether or not an obtained finding will hold up in a new situation or not. It is this fact, above all, which has made so many educators skeptical of basic behavioral research and its applicability to education. Hence, the well known clarion call for lots of expensive “field testing.”

Although representativeness of experiments is necessary in testing both probabilistic and deterministic theories, to insure generalizability of results, the situation in the latter case is quite different. In testing deterministic theories, results do not refer to average effects. The results pertain to what individual subjects will do in specific situations. We are essentially replicating an experiment every time a new subject is tested in a given situation, or, for that matter, every time the same subject is tested in a new situation. Accordingly, it is far easier to obtain representative data in the case of deterministic theories than in the case of probabilistic ones. (On the other hand, of course, deterministic theories only apply deterministically under idealized boundary conditions.)

By way of analogy, for example, consider the inclined plane law mentioned earlier. Suppose we pull a given cart up a given inclined plane and measure the force needed. Each time we do this, we are obtaining a critical data point which may be compared directly with predictions from the inclined plane law. If the inclined plane were perfectly smooth, and the cart were frictionless, this law would allow us to determine the needed force (almost) perfectly. Of course, there is no perfectly smooth inclined plane, or frictionless cart. But, under laboratory conditions, one can approximate the ideal situation so closely that the error of measurement may be ignored for all practical purposes.

For present purposes, the important point is that while deviations from the ideal may cloud experimental comparisons, they cannot eliminate differences. Thus, given any inclined plane, bumpy or otherwise, and given any cart, frictionless or otherwise,
the force required to move the cart up the plane will depend directly on the angle of inclination. This dependence, however, is no longer deterministic but rather is probabilistic because of measurement error.

An analogous thing happens with respect to human behavior. As we have seen, subjects will uniformly succeed on simple problems if they know both the components of a solution and a higher order rule by which these components may appropriately be combined. If any of these essentials is missing the subjects will fail. This basic result has been replicated in the laboratory with different tasks and different kinds of rules, with individuals ranging in age level from preschool through the Ph.D. (This is like checking the inclined plane law in the laboratory at different angles of inclination.)

In the behavioral situation I have described, then, what corresponds to friction and bumpiness of the inclined plane? A large part of the answer essentially concerns the availability in working memory of the requisite rules. To the extent that the needed rules have been learned and are available in a given test situation, to that extent, and only to that extent, will a subject succeed. If any one of these rules is missing the subject should fail.

In the real world of education, we cannot be sure, without prior testing, which rules are and are not known. Hence, any advantages of instruction will accrue just to the extent that the instruction results in the student learning a critical rule that he or she did not know before.

To summarize, whereas valid deterministic theories only apply deterministically under idealized boundary conditions, they necessarily also apply probabilistically in nonidealized situations. They lend themselves to generalization from the laboratory to the real world in a way which probabilistic theories cannot.

Test of an Analogous Probabilistic Theory Under “Real World” Conditions

To keep you from thinking that I am just talking about possibility and not reality, I would like to briefly outline a study by Ehrenpreis and myself (1974). This study directly parallels the one I described earlier but was conducted under "real world" conditions. Basically, what we did was to identify the rules and higher order rules underlying a given mathematics curriculum.

For example, the following tasks were identified in Mathematics: Concrete Behavioral Foundations (Scandura, 1971) on pages 182 and 191, respectively.

Task A. Given a whole number \(m\), determine whether or not \(x\) is an additive identity for \(m\).

Task B. Given a whole number \(m\), determine whether or not \(y\) is a multiplicative identity for \(m\).

The corresponding rules were:

Rule A. Find the sum \(m + x\) and then the sum \(x + m\). If \(m + x = x + m = m\), then \(x\) is an additive identity for \(m\); if \(m + x \neq m\) or \(x + m \neq m\), then \(x\) is not an additive identity for \(m\).

Rule B. Find the product \(m \times y\) and then the product \(y \times m\). If \(m \times y = y \times m = m\), then \(y\) is a multiplicative identity for \(m\); if \(m \times y \neq m\) or \(y \times m \neq m\), then \(y\) is not a multiplicative identity for \(m\).

The obvious relationship between rules A and B can be represented by the following task and its underlying higher order rule.

Task H. Give a rule for demonstrating that a given set of numbers provides an
Deterministic Approach to Research

instance of a property (e.g., commutativity) under some operation, generate a corresponding rule involving another operation.

Rule H. In the given rule, replace the original operation by the new operation and any "special" element (e.g., the identity) by its counterpart.

Higher order rules of this type made it possible to eliminate those tasks and corresponding rules that are derivable by application of the higher order rules to other rules. Thus, for example, Rule B can be eliminated inasmuch as it can be generated by applying Rule H to Rule A.

In all, 129 of the 303 lower order rules that were originally identified could be derived by applying the five higher order rules identified to the remaining 174 rules. To determine the instructional value of the higher order rules, the derivable rules were eliminated in the training of one group of students. This higher order rules (H) group was also taught the five higher order rules. Another group (D) was taught all 303 lower order rules as discrete units. After training all students were tested on: (a) lower order tasks common to both groups, (b) higher order tasks on which only the H group was trained, (c) transfer tasks on which only the D group was trained, and (d) transfer tasks on which neither group was trained.

The results showed that: (a) both groups performed at a high level on the (a) tasks (about 95% correct), (b) the H subjects had 71.5% success on the H tasks after training and the D subjects had 53% success without training (difference, p < .01), (c) the H and D subjects performed at roughly the same level (88% vs. 93%, a nonsignificant difference) on those transfer tasks on which only the D subjects had been trained, and (d) the H and D subjects, respectively, had 68% and 51% success (p < .01) on the transfer problems, on which neither group had been trained.

These results can be summarized without serious distortion by saying that the H subjects were taught less but learned more.

Notice, however, that the higher order rules training was considerably less than perfect (71.5% success after training), especially considering the fact that many (53%) of the trainees apparently could use these relatively simple higher order rules before training. This suggests that many simple higher order rules have already been acquired by the time people reach adulthood, and that, one can normally expect less than perfect acquisition unless special precautions and extended and carefully detailed training are undertaken. Put differently, the transfer tests were conducted under nonidealized "real world" conditions.

It is of interest in the latter regard, that the degree of deviation from the ideal, in terms of higher order rule availability (H = 71.5%, D = 53%), is closely paralleled by the obtained degrees of transfer (H = 68%, D = 51%). The argument becomes even more convincing when one takes into account the fact that the availability of both higher and lower order rules was necessary for success on the transfer problems. Thus, for example, ignoring rounding error, 95% (the average lower order rules availability) of 71.5% (H group, higher order rules success) is 68% (H group, transfer success). Similarly for the D group. 95% X 53% closely approximates 51%.

Conclusions

By way of conclusion, I would like to reiterate three main points. (1) There are alternatives in instructional research to the usual probabilistic paradigm. Direct tests of deterministic theories are demanding but they are not only possible but, in many cases, they are quite feasible. The main requisite is that of realizing suitable idealized test conditions. (2) Although I have barely touched on one aspect of such
theories, deterministic idealized conditions beyond what is not present in Scandura, et al., in p results in real world. Reaching. They could focus on the relationship by-pass a good deal effectively demonstrated. My argument p theory is sufficient for the theory be tested in idealized conditions.
perations, generate a new operation and note those tasks and higher order rules to the extent as it can be clearly identified could be the remaining 174 rules. The derivable rules higher order rules (H) up (D) was taught all means were tested on: (a) tasks on which only the H up was trained, and (d) the level on the (a) tests on the H tasks after (difference, p < .01), (c) higher level (88% vs. 93%), a only the D subjects had 68% and 51% success (p had been trained. ion by saying that the H is considerably less than the fact that many (53%) higher order rules before rules have already been can normally expect less extended and carefully for tests were conducted deviation from the ideal, in (a), is closely paralleled by argument becomes even at the availability of both on the transfer problems. Average lower order rules was 68% (H group, transfer sely approximates 51%.

main points. (1) There are balistic paradigm. Direct e not only possible but, in s that of realizing suitable hed on one aspect of such theories, deterministic behavior theories do exist. Some have been tested under idealized conditions with strong empirical support; a level of support which goes far beyond what is normally found in behavioral research (Scandura, 1971; 1973; 1976; Scandura, et al, in press). (3) These empirical supports imply the applicability of the results in real-world situations. The implications of this last point are potentially far reaching. They could affect the very way in which we as instructional scientists look at the relationship between research and practice. Specifically, it may be possible to bypass a good deal of expensive “field testing.” (Among other things, this paper effectively demonstrates this possibility.)

My argument purports to show that direct empirical support for a deterministic theory is sufficient alone to insure applicability. The only requirement is that the theory be tested in a representative number of specific situations under suitable idealized conditions.

REFERENCES


SCANDURA, J. M. Structural approach to instructional design. Educational Technology, August 1973, 7:13 (b)


EPILOGUE

In view of my major conclusions, one might wonder whether the converse also holds. That is, since deterministic support under idealized conditions necessarily implies probabilistic support under real world conditions, it is natural to ask whether probabilistic support for a theory or hypothesis necessarily implies deterministic support under idealized conditions.

The answer, I think, is both yes and no. Although this is not the context in which to dwell on what is basically a complex issue, consider probabilistic laws of the form P(R) = f(x1, x2, ..., x). In particular, consider the following simple law: The probability of getting an “A” in a course is a function of I.Q. and grades in prerequisite courses. In this and similar cases, it is assumed (implicitly or explicitly) that we do not yet know enough to make perfect predictions but that, in principle, if we added enough (of the right) predictor variables, we could predict things perfectly. This assumption is equivalent to assuming that the extended list of predictors would effectively define suitable idealized conditions.

In fact, of course, no one really believes that we would succeed in accomplishing

126
DETERMINISTIC APPROACH TO RESEARCH

this. Even if we did succeed, the number of predictors would almost certainly be so large (even countably infinite) that the result would hardly be worthy of the name "theory."

The essential point is that if one wants a deterministic theory that lends itself to empirical testing, it is essential that one start with the right variables. It is not possible to start with variables such as I.Q. and prior grades, just because they happen to be convenient, and fairly expect that they might be extended to a deterministic theory.

PLEASE ASK YOUR UNIVERSITY OR INSTITUTION LIBRARY TO SUBSCRIBE TO EP