DOCUMENT RESUME

ED 278 957 CS 008 689

AUTHOR Balajthy, Ernest
TITLE Evaluation of Models of the Reading Process.
PUB DATE May 84
NOTE 23p.; An earlier version of this paper was presented at the Annual Meeting of the International Reading Association (29th, Atlanta, GA, May 6-10, 1984).
PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Comparative Analysis; Elementary Education; Evaluation Criteria; *Evaluation Methods; *Methods Research; *Models; Qualitative Research; *Reading Processes; *Reading Research; *Research Methodology; Statistical Analysis; Theory Practice Relationship

ABSTRACT

A variety of reading process models have been proposed and evaluated in reading research. Traditional approaches to model evaluation specify the workings of a system in a simplified fashion to enable organized, systematic study of the system's components. Following are several statistical methods of model evaluation: (1) empirical research on components, in which the researcher predicts behavior based on an analysis of the model's components; (2) tally sheet summarization of studies addressing particular issues; (3) meta-analytic techniques, a more sophisticated inter-study summarization than the tally sheet approach; and (4) causal modeling, in which researchers go beyond the constraints of correlation analysis to make statistical inferences concerning causation. Comparisons of reading models are often descriptive in nature rather than analytical. Leading texts on reading are generally comprised of key theoretical articles that provide readers with important and original introductory materials on various reading models. English/language arts education has long emphasized philosophical reading theories based on what reading should be, although several attempts have been made to evaluate reading models by constructing classification systems as well. The ultimate test of any reading model rests on whether it can be effectively applied in the classroom. A four-page reference list is provided. (JD)

EVALUATION OF MODELS OF THE READING PROCESS

Ernest Balajthy

Department of Elementary & Secondary Education & Reading

State University of New York at Geneseo

Geneseo, NY 14454

An earlier version of this paper was presented at the International Reading Association, Atlanta, Georgia, May, 1984 (Co-sponsored presentation with the Alvina Treut Burrows Institute).

EVALUATION OF MODELS OF THE READING PROCESS

Reading is a complex, covert activity. Component processes are not directly observable. As a result research on reading depends critically on design of theoretical models of the internal processes involved. A wide variety of such models have been proposed and researched, and more such proposals appear annually.

What is it that makes a good model, a model that effectively portrays components of reading? This paper surveys attempts to evaluate these theoretical models of the reading process.

THE TRADITIONAL APPROACH TO MODEL EVALUATION

Historically in the sciences, models involve specification of the workings of a system. These workings are usually portrayed--either physically, diagrammatically, or verbally--in simplified fashion. That is, for the sake of precision of study, the variables involved in the model are fewer in number than those in the actual system. The relationships of variables proposed in the model are thought by the modeler to be suitably analogous to the system.

The purpose of such a theoretical model is to enable an organized, systematic study of its components. By manipulating the variables and observing the consequences, inferences can be made about the actual system upon which the model is based. "A model in this sense is merely the indication

of a simpler and more accurately determinable state of affairs, with the intention of facilitating deduction of further consequences which can then be tentatively reapplied to the more complex and elusive real system" (Gellner, 1964). Models are meant to be tested empirically, then to be accepted, refined, or rejected on the basis of experimental data.

Two complementary approaches to this traditional modeling process have been apparent. Newell (1973), for example, criticized the lack of general principles arising from the mass of research carried out in cognitive psychology. His solution, and the solution of others who followed him, was to develop complex models of the human mind using information processing paradigms based on computer simulations. These complex models were designed to synthesize a wide variety of psychological processes, including reading.

Posner and McLeod (1982) contrasted this synthesis approach to modeling with an analytic approach. They surveyed research in which modelers had analyzed psychological processes into elementary operations which appeared to yield specific performances, constructing comparatively simple models of those operations for experimental study. A proposed model of letter and letter string recognition, for example, might be tested with matching or detection tasks (e.g., Healy, 1976; Carr et al, 1979).

Beaugrande (1981) has argued that these traditional views of the modeling process have not been substantiated in the realities of the scientific endeavor. Kuhn's (1970) survey and analysis of scientific "revolutions" suggested that the history

of the field of science shows that scientists reject theories for other reasons than that of contrary evidence. Stegmuller, in his 1976 analytical treatment of theory testing, argued that the interrelationships of theory and evidence are so intricately entwined that the two cannot be separated. All data gathering is influenced by theory. "Facts" cannot be completely separable from "theory." A researcher's expectations or bias may well influence results, for example. Research is often aimed at proving a point rather than at determining "truth," whatever truth may be. This may be especially true in the social sciences. Certainly the field of reading is replete with a grand variety of theories and models for which extensive testing has never been carried out.

STATISTICAL METHODS OF MODEL EVALUATION

Empirical Research on Components

In this method, the modeler/researcher makes predictions of behaviors based on components of the model. These predictions are then translated into testable hypotheses, that is, into hypotheses which lend themselves to empirical research and data gathering. Bit by bit, piece by piece, the model is analyzed into its smallest components, each of which is tested. The ultimate goal is to produce a body of research evidence so impressive as to lead to general acceptance of the overall model. Each small, targeted study contributes a "brick" to the

foundation supporting the theory.

Within the reading community, this approach to model evaluation has been widespread. Rothkopf's (1966, 1972) "mathemagenic" model of reading and learning, for example, was the basis for some two to three hundred research studies on adjunct questioning, a good sampling of which were published in Journal of Educational Psychology. Goodman's (1967) description of the reading process as a "psycholinguistic guessing game" resulted in a flurry of doctoral dissertations which analyzed oral reading errors in an uncounted number of variations. More recently, Carver's "rauding theory" has been examined in several Reading Research Quarterly studies (Carver, 1982, 1983, 1984).

Tally Sheet Summarization

What does one do with the overwhelming morass of data and studies resulting from popular acceptance of a model? How can one wade through dozens of published and unpublished studies to evaluate accuracy of the model or to suggest refinements? A common approach is to examine the research studies and construct what amounts to a tally sheet (or "box score" or "taking a vote") on the pertinent issues. How many of the studies found positive results? How many found negative results? How many found "no significant differences"?

For example, in the wake of interest in programmed learning and in Rothkopf's mathemagenic model in the 1970's, an ensuing deluge of research on questioning and other mathemagenic

activities (such as underlining and outlining) concentrated on examining adjunct pre- and postquestions, but the findings at first did not seem to result in readily identifiable patterns. In their comprehensive survey of the literature of adjunct questions, Anderson and Biddle (1975) collated the results of numerous studies to conclude that adjunct questions do indeed influence learning from text. Questions even seemed to obtain positive results when administered in relation to lectures and films, as well as printed text. Anderson and Biddle tallied research results to draw three conclusions:

a) Prequestions facilitate learning of direct, question-relevant information. Ten of fourteen studies demonstrated this result.

b) Prequestions have negative effects on indirect, incidental learning, actually depressing amount of retention. Thirteen of eighteen studies demonstrated this result.

c) Postquestions facilitate the learning of both relevant and incidental information. Thirty-seven of forty studies demonstrated that the direct effects of postquestions are positive, and twenty-six of thirty-nine studies demonstrated that the indirect effects are likewise positive.

Reviews of this type are subject to several serious flaws. For example, they do not deal with the issue of differing effect sizes among the studies. Do eight show very minor (but statistically significant) positive effects, while five or six show major negative effects? A simple tally may not make this apparent. Another problem involves the obtaining of statistically significant results. Larger studies tend to

achieve significance more easily than smaller studies. Still another objection involves the selection of studies to be included in the tally. Ought seriously flawed studies be included, and if not, how shall distinctions be made to weed out such studies? Light and Smith (1971) argue that reviews in which such judgments have been made are particularly susceptible to reviewer bias. Finally, it has been generally recognized that research journals only infrequently publish studies in which no significant differences have been found. Unless tally sheet summaries include unpublished studies such as dissertations and research included in the ERIC system (and even the inclusion of such studies does not deal with those findings which the researchers themselves have shelved as unpublishable), the results may be biased in favor of positive effects.

Tally sheet reviews are not typically limited to the tally, however. Most reviewers include thoughtful synthesis and careful description of the component studies. While such synthesis may well be influenced by reviewer bias, if sufficient care is taken to include descriptions of the studies, the reader is able to make independent judgments as to the conclusions drawn. A simple tally sheet will satisfy few knowledgeable readers.

The next procedure to be discussed has been criticized on this very point. Its statistical complexity and apparent precision may well serve to mislead readers into believing they have been presented with exact findings when in fact the analysis could be seriously flawed.

Meta-Analysis Summarization Techniques

Meta-analytic techniques are more sophisticated inter-study summarization methods than the tally sheet approach. As with tally sheets, research studies on the same issue are gathered together for combined analysis. In most meta-analysis techniques, however, studies are combined for the analysis with differential weighting. A study with 1000 subjects, for example, would be weighted more heavily than a study with 50 subjects. Rather than simply tallying the results for or against, the data are combined and statistically reanalyzed.

The actual original data of the component studies are not used in meta-analysis. The unit of analysis is the study. Glass (1977) suggests, for example, that in experimental-control studies, findings of each study can be standardized by calculating the mean difference in performance between the experimental and control groups and dividing that by the within-group standard deviation of the control group. The resulting standardized figure is called "effect size."

Meta-analytic studies have been criticized on a wide variety of points (Jackson, 1980), but increasing use of the technique is evident in such journals as Review of Educational Research and American Education Research Journal. These reviews are subject to most of the same criticisms as the tally sheet reviews discussed earlier. The statistical re-analysis involved does adjust for differences in sample size between component studies, however. The strongest denunciations of meta-analytic

procedures (e.g., Eysenck, 1978) have been based on the "comparing apples to oranges" argument: Combining the results of diverse studies can yield results that simply make no sense. Proponents of the procedures suggest that between-study differences can be included in the meta-analytic computations and any interactions involving them be emphasized. Even the factor of study quality can be computed: Rate the component studies for quality and check for interactions involving this factor (Glass, McGaw, & Smith, 1981).

Slavin (1984) admits that meta-analysis, properly used, can be a useful tool for researchers in the social sciences. His detailed analysis of eight important publications of this type, though, led him to conclude that meta-analytic studies to date have been seriously flawed and may lead to erroneous conclusions which--due to the apparent precision--stifle future research on the topics involved. "Meta-analysis can pull a veil of numbers over the critical information" (p. 26).

Causal Modeling

Causal modeling (also called linear structure equation, simultaneous equation, and path analysis) is an increasingly popular tool with social scientists, though its applications specifically in reading and educational psychology have been far fewer than in sociological research. The procedures involved attempt to go beyond the generally recognized constraints of correlation analysis to arrive at some statistical inferences concerning causation.

Beginning students in educational research courses are consistently warned not to confuse correlation with causation. If variable A and variable B are correlated, we have no guarantee whatsoever that one causes the other. "One can never infer the causal ordering of two or more variables knowing only the values of the correlations!" (Duncan, 1975). A common pedagogical illustration of this is the correlation of shoe size with position within the educational administrative hierarchy: The bigger one's shoes, the higher one's position is likely to be. Does shoe size therefore cause an increase in position? No, of course not. Shoe size is correlated with sex, and since males remain dominant in higher administrative positions, shoe size is therefore correlated with occupational position.

Causal modeling proceeds in the opposite direction, however. While we cannot reason from correlation to causation, we may be able to reason from our ideas and theories of causation to correlation. That is, assuming we have a model of the reading process, we can make certain predictions as to relationships which should be found among components of that model if the model is valid--predictions which should potentially be subject to refutation if the model is invalid (Bentler, 1980).

Bear in mind that, even if the predicted correlations are found, there is no ground to conclude that one's theories of causation must be true. There may be, after all, alternate models which could fit the data. Failure to reject a tested model does not require acceptance of it.

While causal modeling can be a powerful technique for

statistical substantiation of models, the key lies in the actual prior formulation of the model. "If one's causal analysis goes astray, it will more likely be due to carrying out the earlier steps in the research process poorly rather than to any misuse of the techniques....Probably the best advice that one could offer to someone contemplating the use of causal modeling is to begin with a model in which one has substantial confidence. Presumably this confidence results from some theoretical or substantive reasoning about the linkages between the variables of interest" (Asher, 1976).

GENERAL METHODS OF MODEL EVALUATION

While researchers in the field of reading, with their foundations in educational psychology, have traditionally valued empirical research, statistical analysis is certainly not the only method of evaluation of reading theories and models.

Comparative Summarization: A Historical Approach

Comparisons of reading models have often been descriptive in nature, rather than analytical. Leading texts in the field have consisted of collections of key theoretical articles designed to provide readers with important original introductory materials on the various models discussed. The important International Reading Association publication, Theoretical Models and Processes of Reading (Singer and Ruddell, 1976), for example, provided basic resource articles on models proposed by Holmes, LaBerge and

Samuels, Goodman, Gough and many others.

The final report of the U.S. Office of Education Targeted Research and Development Program in Reading, The Literature of Research in Reading With Emphasis on Models, edited by Frederick Davis (1971) attempted to provide some structure to the increasing variety of relevant model proposals from the many fields related to reading, including psychology, psycholinguistics, information processing, sociolinguistics, and the other behavioral sciences. Papers included were largely descriptive surveys of extant models, however, summarizing and drawing together information garnered during the vast overview of reading research carried out during the course of the project's literature search.

Kling (1984) has proposed an eclectic examination system centered on some twelve analytical questions dealing with a variety of approaches to model evaluation. Used primarily as a pedagogical instrument, his interrogatory instrument deals with the following issues:

1) What aspects of the reading process are covered, and what are the basic components of the theory or model?

That is, summarize the major components of the theory, including a step-by-step description of the processes and explanations of the key concepts.

2) What is the historical background of the theory or model?

That is, demonstrate how the model fits into the historical development of models of reading from related disciplines. Pay particular attention to prior models which served as foundations

for the model under consideration.

3) How does the theory extend previous theory?

Discuss the theoretician's improvements upon prior related models.

4) What is the definition of reading?

Concisely state the theoretician's definition, with emphasis upon quotes from his or her writings.

5) What are the assumptions, and what evidence is given to substantiate these assumptions?

Examine foundational (and often implicit) assumptions. Pay special attention to those assumptions which are unique to the model under discussion, rather than to those which are characteristic of models of reading in general.

6) What are the theoretical paradigms or research tools used to explicate the theory?

What methodologies are typically used by researchers involved in testing the model?

7) What are the key hypotheses stemming from the theory and what studies support them?

List the fundamental hypotheses upon which the model is constructed. Cite major studies dealing with those hypotheses.

8) What is the relationship of the theory to other models or theories?

Use a classification system. Pay special attention to the model's unique contributions, contrasting it to other models which may be similar.

9) What are the strengths and weaknesses of the theory or model?

10) What further research or modeling needs to be done?

11) What are the educational or curricular implications?

Apply the theory to the classroom. Offer practical ideas suggested by the model. Provide a link between theory and classroom practice.

12) What obstacles are likely to be encountered in implementing this model in the classroom?

Discuss issues of a practical nature. Offer advice to the classroom teacher and to the model's proponents as to applied problems which may arise during implementation.

Beaugrande (1981) has offered another summarization approach to evaluation of models of reading. He presented a checklist of sixteen criteria designed to compare and contrast processes usually included in reading models and illustrated its use by analyzing some ten models, including his own.

Philosophical Approaches

While somewhat foreign to the general empirical emphases traditionally found in the field of reading, evaluation of reading theory based on what reading should be has long been the emphasis in the field of English/language arts education. Such writings provide an important basis for thought and underscore the close interrelationships between language, thinking, reading and the more affective, metaphysical and spiritual concerns so often ignored by empiricists. There is more to reading than

reading (as defined empirically or operationally) itself--that is, reading affects the soul, not merely the neurons. And, of course, the effects occur in the opposite direction as well--one's spiritual and philosophical relationship with print is interactive in nature.

Classification Approaches

Several attempts have been made to evaluate models of reading by construction of classification systems. Singer and Ruddell (1976), for example, divide their attention between looking at isolated processes involved in reading and the more comprehensive approaches to modeling reading. Under "Processes of Reading," they include:

1) Language
2) Visual processing
3) Perception
4) Word recognition
5) Cognition
6) Cultural interaction.

In their section entitled "Models," they include four categories:

1) Psycholinguistic models (Ruddell, Goodman)
2) Information processing models (Gough, LaBerge & Samuels, Anderson)
3) Developmental models (Holmes, Singer)
4) Affective models (Mathewson)

As noted above, Kling (1971) organized models according to an

"interdisciplinary matrix" drawn from (but not limited to) five major disciplines:

1) Psychology

2) Psycholinguistics

3) Information processing

4) Sociolinguistics

5) Biobehavioral sciences (e.g., neurobehavioral and psychophysiological).

In Stogdill's The Process of Model-Building in the Behavioral Sciences, Luce (1970) categorizes models according to their complexity:

1) Models of variables

2) Simple models of phenomena

3) Complex models of phenomena

4) Models of experiments

5) Models of interaction among individuals

6) Models of social institutions and mechanisms.

A three-part classification system which has gained considerable popularity in recent years has been the division of reading theories into serial-stage ("top-down" and "bottom-up") and "interactive" categories. Top-down models stress the influence of the mind in reading. Bottom-up models place primary emphasis upon the printed page and the decoding process. Models of both types are large serial in nature, with analysis proceeding through a series of stages (Levy, 1981).

Interactive models, drawing heavily from artificial intelligence research, form a middle ground, asserting that reading is a combination of processes which can operate with both top-down and bottom-up directionality. Primary direction of processing at any given time depends upon the momentary purpose and circumstances of reading. Recent interest in interactive processes has been stimulated largely by Rumelhart's (1977) "Toward an Interactive Model of Reading."

Educational Approaches

In the first edition of their International Reading Association book on reading theories and models, Singer and Ruddell (1970) noted that "the more we understand the [reading] process, the more likely we are to develop and devise instructional programs and materials that will enhance the reading abilities of our students, the ultimate goal of research in reading" (xiii). Ultimately, especially to those of us who consider ourselves to be educators first--and only secondarily researchers--the test of a model rests on whether its application "works" in the classroom, and under what particular circumstances that application succeeds.

A final word of warning to those interested in what has been called "theory into practice." Nice psychological theories do not always translate into successful educational practice. Spiro, Bruce and Brewer (1980) warn that, "Even if we had a complete understanding of the reading process (and we certainly do not), we would not automatically know how reading instruction

should be carried out" (p. 3).

Too often we have jumped upon the latest theories of what goes on in the mind during reading, too hurriedly drawing conclusions as to application. Not only must theory be tested, but the relationship of theory to practice must be examined as well. While some writers often bemoan the lag time between findings of educational researchers and application of those findings to the classroom, the fact of the matter is that this lag time allows for closer critical examination of research before massive numbers of our children are affected.

REFERENCES

ANDERSON, R. C. & BIDDLE, W. B. (1975). On asking people questions about what they are reading. In G. H. Bower (Ed.), The psychology of learning and motivation (Vol. 9, pp. 89-132). New York, NY: Academic Press.

ASHER, H. B. (1976). Causal modeling. Beverly Hills, CA: Sage Publications.

BEAUGRANDE, R. DE. (1981). Design criteria for process models of reading. Reading Research Quarterly, 16, 261-315.

BENTLER, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. In M. R. Rosenzweig, & L. W. Porter (Eds.), Annual Review of Psychology (Vol. 31, pp. 419-456). Palo Alto, CA: Annual Reviews.

CARR, T. H., POLLATSEK, A., & POSNER, M. I. (1981). What does the visual system know about words? Perception and Psychophysics, 29, 183-190.

CARVER, R. (1982). Optimal rate of reading prose. Reading Research Quarterly, 18, 56-88.

CARVER, R. (1983). Is reading rate constant or flexible? Reading Research Quarterly, 18, 190-215.

CARVER, R. (1984). Rauding theory predictions of amount comprehended under different purposes and speed reading conditions. Reading Research Quarterly, 19, 205-218.

DAVIS, F. B. (Ed.). (1971). The literature of research in reading with emphasis on models. New Brunswick,

NJ: Rutgers University Graduate School of Education.

DUNCAN, O. D. (1975). Introduction to structural equation models. New York, NY: Academic Press.

EYSENCK, H. J. (1978). An exercise in mega-silliness. American Psychologist, 33, 517.

GELLNER, E. A. (1964). Model (Theoretical Model). In J. Gould, & W. L. Kolb (Eds.), A Dictionary of the Social Sciences (p. 435). New York: Free Press.

GLASS, G. V. (1977). Integrating Findings: The meta-analysis of research. In L. S. Shulman (Ed.), Review of research in education (Vol. 5, pp. 351-379). Itasca, IL: F. E. Peacock.

GLASS, G., McGAW, B., & SMITH, M. L. (1981). Meta-analysis in social research. Beverly Hills, CA: Sage.

GOODMAN, K. (1967). Reading: A psycholinguistic guessing game. Journal of the Reading Specialist, 6, 126-135.

HEALY, A. F. (1976). Detection errors in the word the: evidence for reading units larger than letters. Journal of Experimental Psychology: Human Perception and Performance, 2, 235-242.

JACKSON, G. B. (1980). Methods for integrative reviews. Review of Educational Research, 50, 438-460.

KLING, M. (1971). Quest for synthesis. In F. B. Davis (Ed.), Literature of research in reading with emphasis on models (pp. 2-1--2-18). New Brunswick, NJ: Rutgers University Graduate School of Education.

KLING, M. (1984). Models of reading. Paper presented at the annual meeting of the International Reading Association,

Atlanta, GA.

KUHN, T. (1962). The structure of scientific revolutions. Chicago: University of Chicago Press.

LEVY, B. A. (1981). Interactive processing during reading. In A. M. Lesgold and C. A. Perfetti (Eds.), Interactive Processes in Reading (pp. 1-35). Hillsdale, NJ: Lawrence Erlbaum.

LIGHT, R. J., & SMITH, P. B. (1971). Accumulating evidence: Procedures for resolving contradictions among different research studies. Harvard Educational Review, 41, 429-471.

LUCE, R. D. (1970). What are mathematical models of behavior models of? In R. M. Stogdill (Ed.), The process of model-building in the behavioral sciences (pp. 115-132). Columbus, Ohio: Ohio State University Press.

NEWELL, A. (1973). You can't play 20 questions with nature and win. In W. G. Chase (Ed.), Visual Information Processing. New York: Academic Press, 1973.

POSNER, M. I., & MCLEOD, P. (1982). Information processing models--In search of elementary operations. In M. R. Rosenzweig, & L. W. Porter (Eds.), Annual Review of Psychology (Vol. 33, pp. 477-514) Palo Alto, CA.: Annual Reviews.

ROTHKOPF, E. Z. (1966). Learning from written instruction materials: An exploration of the control of inspection behavior by testlike events. American Educational Research Journal, 3, 241-249.

ROTHKOPF, E. Z. (1972). Structural text features and the control

of processes in learning from written materials. In R. O. Freedle and J. B. Carroll (Eds.), Language comprehension and the acquisition of knowledge (pp. 315-336). Washington, D.C.: V. H. Winston.

RUMELHART, D. E. (1977). Toward an interactive model of reading. In S. Dornic (Ed.), Attention and Performance (Vol. 6, pp. 215-245). Hillsdale, NJ: Lawrence Erlbaum.

SINGER, H., & RUDDELL, R. B. (Eds.). (1970). Theoretical models and processes of reading. Newark, DE: International Reading Association.

SINGER, H., & RUDDELL, R. B. (Eds.). (1976). Theoretical models and processes of reading, (2nd ed.). Newark, DE: International Reading Association.

SLAVIN, R. E. (1984). Meta-analysis in education: How has it been used? Educational Researcher, 13, 6-15, 24-27.

SPIRO, R. J., BRUCE, B. C., & BREWER, W. F. (1980). Theoretical Issues in Reading Comprehension. Hillsdale, N.J.: Lawrence Erlbaum, 1980.

STEGMULLER, W. (1976). The structure and dynamics of theories. NY: Springer.