DOCUMENT RESUME

ED 355 470 CS 011 217

AUTHOR	Dagostino, Lorraine; Carifio, James M.
TITLE	Establishing the Logical Validity of Instructional Activities for Teaching Reading Evaluatively.
PUB DATE	9 Mar 93
NOTE	27p.; Revised version of a paper presented at the Annual Meeting of the Eastern Educational Research Association (Clearwater Beach, FL, February 17-22, 1993).
PUB TYPE	Speeches/Conference Papers (150) -- Reports - Research/Technical (143)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Analysis of Variance; Critical Reading; *Literacy; Models; Reader Text Relationship; *Reading Comprehension; *Reading Instruction; Reading Research; Secondary Education; *Teaching Methods; *Validity
IDENTIFIERS	Interactive Reading; Reading Uses

ABSTRACT

A study established the logical validity of instructional activities for teaching reading evaluatively. These activities operationalized a theory of literacy and a model of reading comprehension developed by L. Dagostino and J. Carifio that specify 20 general characteristics of literate students who read evaluatively. Nineteen secondary classroom teachers were randomly assigned three characteristics for which they were to construct two instructional activities each. The 114 activities constructed were then randomly assigned to the 19 teachers to assess the degree to which the activities logically reflected the characteristic for which they were constructed. In terms of the instructional activities, all 3 judges agreed that 66 (58%) of the activities reflected the characteristics and 2 of the 3 judges agreed that 108 (95%) of the activities reflected the characteristics. These results supported the view that these 20 characteristics can be translated into valid instructional activities by experienced classroom teachers. (Three tables of data are included.) (Author/RS)

Establishing The Logical Validity of Instructional Activities For Teaching Reading Evaluatively

Lorraine Dagostino

James M. Carifio

University Of Massachusetts Lowell

March 9, 1993

Establishing The Logical Validity of Instructional Activities For Teaching Reading Evaluatively

Abstract

This study established the logical validity of instructional activities for teaching reading evaluatively. These activities operationalized a theory of literacy and model of reading comprehension developed by Dagostino and Carifio (1993) that specify 20 general characteristics of literate students who read evaluatively.

Nineteen classroom teachers were randomly assigned 3 characteristics for which they were to construct 2 instructional activities each. The 114 activities constructed were then randomly assigned to the 19 teachers to assess the degree to which the activities logically reflected the characteristic for which they were constructed.

In terms of the instructional activities, all 3 judges agreed that 66 of 114 (58%) of the activities reflected the characteristics and 2 of 3 judges agreed that 108 of 114 (95%) of the activities reflected the characteristics.

These results supported the view that these 20 characteristics can be translated into valid instructional activities by experienced classroom teachers

The present study sought to establish the logical validity of instructional activities for teaching reading evaluatively. The construct of reading evaluatively used in this study was derived from a theory of literacy and a model of reading comprehension based on a cognitive point of view and its logical implications for instruction of higher-order cognitive abilities (Dagostino and Carifio, in press). Aspects of this theory and model are presented below.

This paper presents the background for the present study, descriptions of the categories used to classify the instructional activities, a description of the procedures for collecting data, and a summary of the results and implications of the findings. The results of the present study not only have direct instructional implications, but also provide a general model of how theory may be translated into direct instructional practices in a validated and systematic fashion.

Background

Reading evaluatively is a major component of a model of reading comprehension developed by Dagostino and Carifio (in press). Including such a component in a model of reading comprehension necessitated delineating the construct of reading evaluatively and testing its validity. Of particular

interest to the researchers was an application of the construct to instruction. This interest and focus led to the development of twenty general characteristics that we hypothesized constituted reading evaluatively. Each of these general characteristics we hypothesized could be translated into instructional activities by classroom teachers.

Aspects of a Theory of Literacy and A Model of Reading Comprehension

The theory of literacy underlying the model of reading comprehension in this study emphasizes the idea of multiple proficiencies, and represents those proficiencies through four spheres of literacy. The four spheres of literacy are (1) functional literacy, (2) specialized literacy, (3) cultural and multicultural literacy, and (4) critical literacy. This view of literacy models the changing needs and expectations of different environments and kinds of literacy. It recognizes the multiple kinds of literacy represented in the literature and the potential relationships between and among the spheres. Most importantly, this theory of literacy emphasizes the ability to evaluate critically what is read. Regardless of the sphere or spheres of literacy, a reader finds that simply attaining basic comprehension is not enough for living in today's world.

Developing a theory of literacy with evaluation as a driving force led us to develop a comprehensive model of reading comprehension that also focused on processes of evaluation. This model of reading comprehension is organized around the principle that the process of evaluation is used continuously throughout the reader-text interaction. This view means that evaluation of a text is a part of the process, as well as the reader's response to what is comprehended and represented in an interpretation of the text. Also important to this model is the idea that a reader's evaluation of a text may vary with (1) the degree of objectivity and emotional distance maintained by the reader, (2) the stringency of any application of criteria, and (3) the latitude of selected criteria. These three decisions, in turn, influence the nature of the continuous interaction.

This idea of continuous interaction throughout the reader-text interaction contrasts with the strict sequential, and hierarchical view of reading comprehension that suggests a direct progression from literal to inferential understanding, and then to evaluation in step-like levels. The inclusion of evaluation as a continuous interaction in reading comprehension also differs from other interactive models of reading comprehension that focus on literal and

inferential comprehension. The view that the evaluation of text occurs continuously may be derived from a cognitive view of information-processing and learning (see Dagostino and Carifio for details, in press).

The model we have proposed establishes an evaluation component in reading comprehension that needs to be more clearly delineated so that we can know what constitutes the evaluation process and how well synthesized it is. We have proposed general characteristics that may constitute the process of evaluation. The question addressed by the present study is How well can these proposed characteristics be translated into instruction?

Description of the Characteristics of Reading Evaluatively

The twenty general characteristics of reading evaluatively which were being tested for their logical validity are listed and described below.

1. Reader's General Approach To Text: The reader's approach to text includes activating prior knowledge, ensuring openness to the author's message, stating the reader's knowledge and beliefs, and testing the message against experiences, logic and other sources.

2. Reading For Intention and Assumptions: This characteristic of reading evaluatively means comprehending

explicit and implicit messages and testing the plausibility of the reader's conclusions. Part of determining explicit and implicit messages means ascertaining the writer's assumptions about the topic and generating other possible assumptions either complementary or contradictory.

3. Reading Different Kinds of Material Evaluatively: This characteristic of reading evaluatively entails understanding how the reading process is transformed for poetry, exposition, narrative and argumentation by using specific criteria for formal evaluation.

4. Establishing Purpose In Reading Evaluatively: This characteristic entails establishing purposes beyond getting literal and basic inferential meaning. Instead the reader focuses on getting at underlying intentions, assumptions and theme.

5. Teaching How To Read For A Point Of View: This characteristic entails understanding that a point of view represents and interprets events differently. This is true in fiction as well as non-fiction although it takes slightly different forms.

6. Determining Attitude Toward Reading: A reader must recognize how she feels about a topic, types of material, styles of writing, specific points of view because attitude

determines how you interact with the ideas and the message and your level of disagreement.

7. Sorting Personal Feelings From Application of External Criteria: A reader needs to know how emotional responses differ from intellectual responses by identifying the "criteria" of emotional response and the "criteria" of objectivity.

8. Constructing A Point Of View: A reader needs to understand how having a point of view logically leads you to certain actions and reactions.

9. Drawing Conclusions, Making Predictions and Suspending Judgment Through Maintaining Tentativeness and Hypothesis-Testing: This characteristic of reading evaluatively is an information-gathering process that requires the reader to bring details to closure periodically, to withold personal judgment and set the direction for thinking that predicts what will occur. Logic and information must direct the predictions and the final conclusions for plausible endings to be legitimized.

10. The Effects of Interpretation As A Preliminary Step To Evaluation: This characteristic requires understanding how multiple interpretation help create the openness to message necessary for evaluative reading.

Seeing that interpretation influences how external criteria are applied and accepted.

11. Integrating The Reader's Personal Views With The Views of The Larger World: A reader's personal views may be limited or misconceived in relationship to the views of the world at large. This characteristic focuses on the need to match schemas and synthesize them based upon multiple views.

12. Understanding That Reading Comprehension is a process of gathering, assimilating, thinking and restructuring: A reader views reading comprehension as an interactive process where the reader uses skills and strategies to find explicit and implicit meaning. The reader is thinking about meaning and how he arrived at it so that an appropriate response can be made. This means using 1) determining importance of information 2)summarizing text, 3) drawing inferences, 4) generating questions and 5) monitoring comprehension for achieving a mature response.

13. Establishing Readers' Goals: Several questions help the reader determine the author's perspective. These questions are: 1) What is the author's point of view? 2) How does she establish it? and 3) What is her purpose? A reader always aims to determine what the writer has to say. Understanding

the author's point of view and purpose leads to making a reasonable interpretation of text. Recognizing how the writer accomplishes his goal helps the reader evaluate the message. Once purpose and point of view are established a response can be made.

14. Detecting Deception: This facet of reading evaluatively focuses on recognizing propaganda and whether it sells something or someone of value. When the purpose is to persuade the reader must recognize the writer's slant of information and ideas. The reader must recognize the situational purpose of a text and weigh the message and the presentation.

15. Seeking Credibility and Validity: This dimension of reading evaluatively focuses on distinguishing Fantasy from Reality and distinguishing Fact from Opinion on the bases of authenticity, adequacy and relevance as well as on following and evaluating logic and argument. It also entails recognizing the influence of beliefs and attitudes which are part of a person's moral structures and knowing what to believe or to do is based upon the reader's ability to find the logic and truth in the text. Doing this means determining the truth unbiased by moral views of an issue. Readers are looking for plausibility and possibility of occurrence of events as well as

internal validity.

16. Developing Sensitivity to Language: This aspect of reading evaluatively is focused on Diction, Denotation and Connotation, Tone, Figurative Language, and Use of Syntax. Readers must become sensitive to the framing of a message and what words are chosen to create a particular tone. Additional features are use of formal or informal register, use of symbolic or abstract language, and overall effect of phrasing and sentence structure.

17. Building Knowledge Bases: This aspect of reading evaluatively focuses on building general knowledge as well as specific schema for different established bodies of knowledge on different topics. It also entails building points of comparison to test ideas in text and building bases for assimilation and accomodation of new information and ideas.

18. Developing Schema For Reflective Questioning: Readers must learn to monitor comprehension by questioning themselves and the text in a reflective manner.

19. Achieving Automaticity Of Analysis: Readers must automate all of the skills and strategies they have developed so that analysis and evaluation are assimilated in the reading habits and process they take to text.

20. Generating Response To Texts: This dimension of reading

evaluatively focuses on a reader's response to text beyond basic literal and inferential comprehension.

As previously stated, reading evaluatively is a major component of reading comprehension. The twenty general characteristics of reading evaluatively described above were derived from a theory of literacy and a model of reading comprehension based on a cognitive point of view. Our efforts to delineate these characteristics and test them for logical validity should help determine if reading evaluatively is a clearly distinguishable set of characteristics that can be translated into individual instructional activities or a synthesized ability. In either case, our findings are informative on several points, and should prove useful for instruction of reading evaluatively as an integral part of reading comprehension.

The present study does not address instructional sequencing issues or whether our construct is appropriate for a specific age group or for all learners. The instructional activities that were developed in the present study were created with secondary school students in mind as a "baseline". They may be appropriate, with or without modification, for learners in other age groups.

Research Design

The general model and methodology used in this study is an adaptation of Campbell and Fiske's (1959) convergent and discriminant validity paradigm. This paradigm is something called the method of "congruence" (Kerlinger, 1986), the method of "triangulation" (Borg and Gall, 1989), or a "panel" design (Lanza and Carifio, 1992). In general, the focus on this method is to have someone generate an item (instructional activity) that is suppose to operationalize and reflect some objective specification (i.e, characteristic of reading evaluatively) and then have some other independent judge rate whether the item corresponds and logically reflects the objective specification. This basic design or model is strengthened through replication which is considered by most researchers to be the second basic cannon of experimentation after randomization. Consequently, one would have at least two judges rate the item (instructional activity) produced according to the objective specification. If the two judges agreed that the item (instructional activity) reflected the specification then their judgments would "converge" or be "congruent" and this (intersubjective agreement) would be evidence for the item's logical validity; namely, that the item reflected what it claimed

to reflect. This basic model and logic not only supports the logical validity of the instructional activities written, but also is strong evidence for the specifications being relatively clear and precise enough for independent writers to create equivalent activities using them. This latter point is a point that is often missed in the interpretation of outcomes of this design.

Multiple writers devising "logically equivalent" instructional activities, therefore, is strong empirical evidence for both the objective quality of the criterion specification and the discriminability of one criterion (characteristic of reading evaluatively) from another. Such an outcome would also mean that valid instructional activities could be readily and routinely created for each characteristic of reading evaluatively, provided that practicing classroom teachers were used to write the instructional activities. Using the same classroom teachers to rate the instructional activities written, moreover, would be strong evidence for the quality and the discriminability of the specifications of each characteristic of reading evaluatively. It was for this later reason that practicing classroom teachers were used in the present study.

Methodology

The instructional activities that were to be evaluated were generated by 19 classroom teachers in a graduate course in Developmental Reading for the Secondary School. The activities were developed after the teachers discussed each characteristic and after they had reviewed sample activities developed by the researchers. The discussion was the focus of a two hour class session.

The activities that were developed were general guidelines for how to teach one of the twenty characteristics described earlier in this paper. For the most part, the guidelines helped teachers lead students through a step by step process of instruction. Each activity was designed to be used with a number of texts rather than one specific book or passage, thus making them useful for more than one lesson. The descriptions of the instructional activities were approximately one-half to one typewritten page in length. They were not lengthy lesson plans.

To develop the instructional activities, each classroom teacher (n=19) was randomly assigned 3 characteristics, and then was asked to develop and write out 2 instructional activities for each characteristic assigned for a total of 6 instructional activities per teacher. These classroom teachers

were given one week to create the activities.

The 19 classroom teachers in this study generated a total of 114 instructional activities. Once the 114 instructional activities were developed, they were packaged for a blind review by the same 19 classroom teachers. The reviews were done four weeks after the activities were developed. Each reviewer's packet contained 18 instructional activities to rate that were randomly assigned to each reviewer, excluding a given reviewer's own activities.

Each instructional activity was evaluated by at least 3 reviewers. Each packet of instructional activities had (1) a cover sheet with directions, (2) a rating sheet with the rater's code number and spaces for the reviewer's responses to each of 18 activities, (3) 18 instructional activities to be rated and (4) a list of the characteristics of reading evaluatively.

The reviewers were instructed orally and in writing to read each activity to determine if the activity represented the characteristic it was said to represent. In terms of making this congruence judgment, the reviewer could respond "yes, not completely, or no". If the reviewer thought the activity did not represent the stated characteristic (i.e., responded no or not completely), s/he was asked to explain

why not.

The reviewers also were asked to identify other characteristics the activity might represent as an indirect measure of each characteristic's discriminability. Clarification of instructions was given to entire group where needed. The time allotted to performing these ratings was 2 hours. The reviewers worked independently in the same room. They could take breaks, but they could not discuss their judgments. This rating task was done as an activity in a graduate class, and it represented an application of concepts related to reading evaluatively. When the reviewers completed their work, graduate assistants keyed the reviewer's responses into a computer file so that the data could be analyzed.

Results

The first question that was addressed was, "What is the inter-rater agreement when all judges and all activities are analyzed as a group?" Table 1 presents the frequencies and percentages of inter-rater agreements by agreement type across all activities. The square roots of the agreement percentages approximate the inter-rater correlation (agreement) coefficients (See Kerlinger, 1986).

As can be seen from Table 1, all three raters agreed that 66

of the 114 instructional activities (57.9%) operationalized the characteristic it was constructed to represent. Two of three judges agreed that 108 of 114 instructional activities (94.7%) operationalized the characteristic they were constructed to represent. These percentages translate into an extremely conservative lower limit estimate of inter-rater agreement of r = +.76 and an upper case estimate of r = +.97. The percentages of inter-rater agreements show that the raters understood the construct of reading evaluatively represented by the twenty categories and that they were able to develop and judge the activities well. The percentages of inter-rater agreements also showed that the activities fit the categories well.

Table 1: Percentages of Inter-Rater Agreements by Type of Agreement Across All Activities.

Type of Agreement	Frequency	Percent	Cum. Percent	R
1. All 3 raters agreed	66	57.9	57.9	.76
2. 2 of 3 raters agreed; 1 partially agreed	24	21.1	78.9	.89
3. 2 of 3 raters agreed; 1 disagreed	18	15.8	94.7	.97
4. 2 of 3 judges disagreed, or only partially agreed	6	5.2	100	
	114	100%		

As each rater rated 18 activities of the 114 activities, the ratings were correlated. Therefore, to take into account the correlation between the rating for the 19 raters, a 19 x 18 repeated measure ANOVA was used to assess the degree of intra and inter-rater agreement. As raters could all agree with each other but have the "wrong" answer, a deviation score between each rating and the "correct" response (i.e., yes) was computed so that this relativity was removed from the data making the results directly and easily interpretable.

Table 2: One-way Repeated Measures ANOVA for the 18 Ratings of All 19 Judges.

Source	df	MS	F	p
Judges	18	.96	2.34*	<.01
Ratings	17	1.07	2.61*	<.001
ERROR	306	.41		

Table 2 presents a one-way repeated measures ANOVA on the 18 ratings done by all 19 judges in terms of deviations from the "desired or correct" response scores. As can be seen from Table 2, significant differences were found between

judges ($F = 2.34$; $df = 17, 306$; $p < .001$). Post-hoc analyses revealed that these differences were due to three judges who had a disproportionate number of "incorrect" responses which made these 3 judges significantly different from the other 16 judges in terms of their judgment correctness and profiles. These three judges simply may have been poor, or "error prone" judges for a number of reasons. First, they may not have understood the material presented to them or the rating task. Another reason may be that these three judges may have had prior views of literacy and the reading comprehension process that interferred with the views and model inherent in the materials being rated which produced the consistent judgment errors observed in their ratings. Whatever the reason, these three judges were eliminated from the analyses as half or more of their ratings were incorrect and the repeated measures ANOVA run again. Since only 5 or 6 judges are needed in validation studies of this kind, we still had a sufficient number of judges in the reduced analysis to have confidence in the generalizability of our findings and conclusions.

Table 3 presents the one-way repeated measures ANOVA on the 18 ratings done for the remaining 16 judges. As can be seen from Table 3, no significant differences were found

between judges or rating patterns for the remaining 16 judges on reanalysis.The F-ratio for judges fiven in Table 3 converts into an inter-rater agreement coefficient of r = +.94 (See Kerlinger, 1986). These 16 judges agreed with each other and rated the 18 instructional activities they rated as valid operalizations of the characteristics they were linked to logically. Our 20 characteristics were operationally well-defined enough for classroom teachers to construct instructional activities to be valid exemplars of the characteristic by three other classroom teachers.

Table 3: One-way Repeated Measures ANOVA for the 18 Ratings by The Remaining 16 Judges.

Source	df	MS	F	p
Judges	15	.25	0.7	>.05
Ratings	17	.38	1.5	>.05
ERROR	255	.38		

A comprehensive table was constructed that summarized the number of activities constructed for each characteristic, the number of judges that rated the activities constructed,

and the rater agreements for each of the 20 characteristics. This table is not included here because of its size and unwieldiness. However, the pertinent information from this table will be summarized here. As would be seen from the table constructed, agreement of all 3 raters ranged from a low of 55.6% to a high of 100% across the 20 characteristics. Characteristics 11 (Integrating a Personal View), 13 (Establishing Goals), 14 (Detecting Deception), 18 (Developing Schema and 19 (Automaticity of Analysis) were the areas of least inter-rater agreement with values ranging from 55.6% to 66.7%. Characteristics 5 (Reading for Pt. of View), 9 (Conclusions, Predictions), 15 (Credibility and Validity) and 20 (Generating Responses to Text) were the areas of highest inter-rater agreement with values ranging from 91.7% to 100%. Agreement of 2 out of 3 judges ranged from 83.3% to 100% for all 20 characteristics.

When the percentages of agreement for 2 out of 3 judges are compared with correct responses, the results were still good. Only in one category, 19 (Automaticity of Analysis), does the percentage of agreement fall below 70%. All of the ratings are based upon the judgments of 6 activities for each category with the exception of two categories. They had 4 actitivies each. The number of different judges rating

the activities included in each category ranged from 9 to 14 suggesting further how well the activities fit the specified criterion.

In general, the results support the conclusion that the instructional activities created represent the characteristics (constructs) well, and that those characteristics can be translated,with minimal training, from theoretical constructs (generic characteristics) to instructional activities by classroom teachers. However, we do need to explore those characteristics where there were difficulties classifying some of the activities to determine whether the problem is with the construct (characteristic definition) or with the teacher who develops and rates the activity. Examining how often and why judges said that an activity did not fit a particular category and what other categories it also fit may provide some insights concerning these issues when this more qualitative data is examined later.

Conclusions

The central questions addressed by the present study were How well-defined are the 20 proposed characteristics of reading evaluatively? and Can each characteristic be translated into instruction that logically and validly reflects the characteristic?

Through the use of a convergent and discriminant validity paradigm, the evidence collected warranted a yes answer to both questions. The results of the inter-rater agreements and the one-way repeated measures ANOVA's strongly supported the contention that the 20 characteristics can be translated into valid instructional activities.

The question of how mutually exclusive the characterisitcs are needs to be tested further through other forms of analysis. Further examination of the ratings of the judges eliminated from the study also may give insight into the ability to operationalize the construct.

The findings of this study are an important step in the logical validation of the model of reading comprehension and the theory of literacy proposed earlier in this paper.

The results of this study strongly support the view that the 20 characteristics of the evaluative reader outlined at the beginning of this paper are sufficiently detailed and clear enough for classroom teachers to construct valid instructional activities which reflect and promote these characteristics in students. This finding is important as it means that (1) the dissemination of the definitions of the 20 generic characteristics given at the beginning of this paper to classroom teachers will stimulate the production of

instructional activities that develop these characteristics and (2) classroom teachers may "trade" the instructional activities they develop with each other with a high degree of confidence that what they "give and take" will be valid. This latter knowledge, we believe, is knowledge that is important to fostering our different view of literacy and reading evaluatively.

Lastly, it should be noted that many theories and models are proposed in the areas of instruction and education. However, these theories and models are rarely assessed in terms of how easily, reliably, and validly they may be transformed into actual instructional activities and practices by classroom teachers, which is an important characteristic of any theory or model proposed. We have experimentally assessed how easily, reliably and validly our 20 characteristics of the evaluative reader which are derived from our theory of literacy and model of reading comprehension can be translated into actual instructional activities by practicing classroom teachers. Our results were not only excellent but we also have provided and field-tested a generic model for assessing the "translatability" of theories and models into instructional activities and classroom practices. This later point is an additional strength and benefit of this study.

References

Borg, W.R. and Gall, M.D. (1989), Educational Research, Longman, New York

Campbell, D.T. and Fiske, D.W. (1959), Convergent and discriminant validation by multi-trait-multimethod matrix, Psychological Bulletin, 56, 81-105.

Dagostino, L. and Carifio, J. (in press), Reading Evaluatively and Literacy: A Cognitive View, Allyn and Bacon, Inc., Boston.

Lanza, M. and Carifio, J. (1992), Using a panel of experts to establish validity for patient assault vignettes, Evaluation Review, vol 17, no.1, February, 82-92.

Kerlinger, F. (1986), Foundations of Behavioral Research, Holt, Rinehart, and Winston, New York.