DOCUMENT RESUME

ED 214 110 CS 006 519

AUTHOR Beck, Michael D.
TITLE An Adaptation of the "Standards for Evaluations" for the Interpretation of Reading Program Evaluations. Final Report.
INSTITUTION International Reading Association, Newark, Del.
SPONS AGENCY National Inst. of Education (ED), Washington, D.C.
PUB DATE 30 Jun 81
GRANT NIE-G-80-0075
NOTE 49p.

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Administrator Guides; Elementary Secondary Education; *Evaluation Criteria; *Evaluation Methods; *Program Evaluation; *Reading Programs
IDENTIFIERS *International Reading Association

ABSTRACT

This paper reports on work by an International Reading Association (IRA) committee to develop criteria--for use in judging the superiority of a particular reading program--that may be used by school personnel in the interpretation of evaluation reports. The paper outlines (1) the background of the project, (2) project activities, (3) anticipated uses and usefulness of the guidelines, (4) project funding and staffing, and (5) dissemination and distribution plans. Appendixes include a draft of the guidelines used for the field test, a summary of the comments on the field test guidelines by local district reviewers, a draft of the final IRA guidelines, a memo to IRA committee reviewers of the final draft guidelines, a summary of comments on the final guidelines by IRA committee members, field test reviewers' comments concerning future evaluation activities, and recommendations concerning local uses of the guidelines. (HOD)

Project Title: AN ADAPTATION OF THE "STANDARDS FOR EVALUATIONS" FOR THE INTERPRETATION OF READING PROGRAM EVALUATIONS

Principal Investigator: Michael D. Beck
The Psychological Corporation
757 Third Avenue
New York, NY 10017
(212) 888-3546

Project No.: NIE-G-80-0075
Starting Date: July 1, 1980
Ending Date: June 30, 1981

Final Abstract

The International Reading Association (IRA) revised the Stufflebeam joint committee's "Standards for Evaluations" to assist the reading profession in interpreting reading program evaluations. Under development by IRA for a year prior to field test, the adaptation work resulted in a draft document. This project involved a field test of the IRA adaptation in four local school districts. The purposes of the field test were to assess the usefulness of the adapted standards and to enable IRA further to improve and tailor its work to local school personnel's needs.

School districts currently conducting or interpreting an evaluation of their local reading program participated. Local staff were oriented to the IRA draft "Guidelines" and then applied them in their districts. Project staff gathered feedback concerning the Guidelines from the local districts during and immediately after the field test. Based on the results, the Guidelines were revised, reviewed by various IRA Committee members, and prepared in final draft form. They will soon be published by IRA and distributed to its membership. Other dissemination procedures are underway.

Local school personnel consider reading program evaluation to be an important activity. However, most districts are without access to professional assistance in interpreting evaluations. Based on the many requests IRA receives for assistance of this type, the publication that results from this project has the potential for guiding a large number of local districts toward better use of evaluations. The field study enabled IRA to conduct an empirically based revision cycle for the Guidelines, thus making them more useful to local school districts.

Executive Summary

An Adaptation of the "Standards for Evaluations" for the Interpretation of Reading Program Evaluations (NIE-G-80-0075)

I. PROJECT BACKGROUND

Evaluations of educational programs and materials are an important concern of almost all school administrators. Although many procedures for conducting such evaluations have evolved, most local school personnel are not trained in these techniques. Further, local staff is often even less equipped to evaluate an evaluation -- to analyze the procedures used, data included, and conclusions reached. It is toward this latter problem -- interpreting an evaluation report -- that this proposal is directed.

In 1978, the International Reading Association (IRA) organized a Committee on the Study of Reading Program Evaluations. This Committee had as its primary charge to develop "criteria to be used to judge the adequacy of the evidence to support the superiority of a particular reading program..." The Committee used the Standards for Evaluations of Educational Programs, Projects, and Materials document as a starting point in evolving a written set of "Guidelines for Reviewing Evaluation Reports" to be used by people attempting to assess an evaluation report. This work went through three revisions prior to the field test funded by this grant. The purpose of the field trial was to assess the Guidelines' usefulness and improve upon them before making available a revised set of Guidelines to IRA's general membership.

IRA members are involved in the interpretation and use of literally thousands of reading program evaluations annually. The adaptations we propose to publish, assisted by this funding, will find wide use among our 70,000 members. Further, this work will improve both the understanding of evaluations by school personnel and the interpretation and application of such evaluations on a local level.

II. PROJECT ACTIVITIES

The IRA Study of Reading Program Evaluations Committee conducted this project, with assistance from the IRA Research Department. There were essentially eight stages to the project, as outlined below.

1. Finalize the Guidelines draft and Field Test plans. This was completed in two stages. An interim set of Guidelines was developed for review in May, 1980. These Guidelines were then further revised and adapted at an August, 1980 meeting of the Committee. Also at that time, general plans for the field test were discused, site selection procedures and materials were finalized, and plans for the orientation for local field test coordinators were made.

2. Site Selection. During September, 1980, 105 IRA local council presidents were sent information concerning the field test, soliciting the participation of their member school districts. A second mailing was made to a somewhat-expanded list of local councils during October. These two mailings resulted in approximately 20 inquiries or expressions of interest. From this number, a list of five districts

from four states with varied evaluation experiences, materials, and interest was determined. These districts were formally invited to participate. Since one of these districts later was forced to withdraw, the field test was conducted with four sites.

3. Orientation for Local Districts. This meeting was held January 18-20, 1981 at IRA headquarters in Newark, Delaware. At that time, committee members and appropriate IRA staff members met with representatives of the four field test districts. These districts were: Webster (NY) Central School District, Rock Falls Twp. (IL) High School District, Houghton Lake (MI) Community Schools, and Downers Grove (IL) Grade School District.

During this meeting, the project was explained to the field test coordinators, the draft Guidelines were discussed in detail, and local implementation procedures for the field test were discussed. One IRA committee member was assigned to each local district as its "IRA contact person." Based on feedback received at this meeting, additional revisions were made to the Guidelines. The Guidelines revised as a result of the orientation meeting was the set used for the field test.

4. Field Test. Following the meeting in Delaware, each of the field test coordinators conducted informal "orientation sessions" with several persons in their district who had interest in or responsibility for some aspect of a reading program evaluation. Then various local personnel were asked to review independently the Guidelines draft document. The focus of the review was on how the Guidelines could best be improved in content, organization, scope, or wording. The 22 local reviewers were also asked to complete a "Comments & Reactions" page.

The review led to several conclusions concerning the Guidelines draft. First, selected portions of the Guidelines required additional clarification, editing, and/or rewriting. Second, despite these many suggestions, the reviewers generally found the document to be very comprehensive and clearly organized, readable and reasonable for school use, and sufficiently non-technical. The primary negative reaction received from reviewers was that the Guidelines were almost too comprehensive and needed to be shortened somewhat, with jargon eliminated.

5. Site Visitations. Two site visits were scheduled for each field test district. The first was to be made by the Committee member assigned to the district; the second visit was to be made by a Research Specialist from IRA headquarters. The purpose of the two visits was the same: to meet with/interview local people who had participated in the field test review process to obtain their reactions to and suggestions for the Guidelines.

Both Committee members and local coordinators reacted well to the format and scope of the visits and no difficulties were encountered. Further, local reviewers gave willingly of their time for these interviews and voiced positive reactions both to the project or their district's participation.

6. Committee Revision of the Guidelines. During the IRA Annual Meeting, Committee members met to revise the Guidelines. Before this time, Committee members had submitted written summaries of their site visits for review. Both general topics concerning the purpose, audience, overall tone, and scope of the Guidelines, as well as specific wording and rewritting suggestions were explored. Guidelines drafts reviewed by the various local personnel were used to make revisions in the draft. Many changes made in the Guidelines at this stage, based on the field test results.

7. IRA Review of the Revised Guidelines. A revised version of the Guidelines draft was sent to 101 selected IRA volunteers, including chairpersons of each of the 48 IRA Standing and Ad Hoc Committees and all members of four committees, the memberships of which is heavily weighted with people interested in and knowledgeable about evaluation... The 101 people were asked to "critically review" the draft and make either editorial or substantive comments. They were also asked to complete a "Comments & Reactions" form similar to that used in the field test.

8. Final Draft. Although this phase of the project was not covered by NIE funding, nor does it fall within the activities of the grant, it is outlined here for closure purposes. As with the field test responses, the most worthwhile portions of the IRA review were the comments item by item and major section by section. These comments were used to pull together yet another revision of the Guidelines which is currently being reviewed by Committee members. The proposed final draft will be submitted to the IRA Board of Directors in March for approval. This "final draft" version will be referred to the IRA Publications Committee for dissemination.

III. ANTICIPATED USES AND USEFULNESS OF THE GUIDELINES

In general, we would make the following observations concerning the anticipated usefulness of the final Guidelines:

1. The Guidelines as a whole are too comprehensive for most local evaluations. Maximum usefulness will probably result if a small group of local personnel select the relevant portions of the Guidelines and the entire set is reduced to some limited subset of items especially pertinent locally. This is now suggested in the Introduction.

2. The two primary uses of the Guidelines will probably be: (A) in reviewing plans for an internal evaluation before the evaluation begins -- looking for additional components to be considered; and (B) in judging the quality of an external evaluation after it has been completed. We believe that, for these two purposes, the Guidelines will be helpful as a stand-alone document for most districts.

3. Use of the Guidelines for either the planning or evaluating function would probably be improved if an orientation to the topic and document was available. We will continue to explore this impression through the sessions to be conducted over the next several months using the Guidelines.

IV. PROJECT FUNDING AND STAFFING

The project was funded for the budgeted total of $11,082. Despite greatly increased travel costs during the contract period, the total cost to NIE of this project was $9551.64. All professional time for this project, except that spent by our IRA Headquarters Assistant, was donated to IRA. Records of this time were kept by each Committee member and local field test coordinator during the project. These items totaled to 137 professional days. It is clear that the project was completed at a minimal cost due largely to the donation of time by these concerned professionals, all of whom we sincerely thank.

Figuring donated staff time at an estimated $150 per day, including benefits, and including no additional overhead expenses, it is estimated that the total "cost" of this project was over $31,000, only 30% of which was charged back to NIE. Projects such as this, seem to be an ideal way for governmental agencies to derive maximum impact for their shrinking contract/grant funds.

V. DISSEMINATION AND DISTRIBUTION PLAN

Clearly, the ultimate impact of our efforts will depend on the extent to which the Guidelines are disseminated to the IRA membership and interested others. We have been keenly aware of this throughout the project and have already finalized the following plans:

A. Presentations at Conferences. Sessions have been proposed and accepted for presentation as of this date at three conferences. Presenters will be either Committee members or participants in the field test. In addition to these currently scheduled sessions, two proposals for additional presentations are planned. Feedback from these sessions will assist IRA in additional refinement of the Guidelines for publication.

B. Publications. The final draft of the Guidelines will be submitted by the Committee for approval by the IRA Board of Directors. Following approval, the Publications Committee will decide on the means of distribution for the Guidelines via the IRA publications system. Finally, an article describing the development of the Guidelines will be submitted to *The Reading Teacher* or *Journal of Reading* for publication. All of these efforts should ensure the wide distribution of the Guidelines to interested professionals.

FINAL REPORT

An Adaptation of the "Standards for Evaluations" for the Interpretation of Reading Program Evaluations (NIE-G-80-0075)

I. PROJECT BACKGROUND

Evaluations of educational programs and materials are an important concern of almost all school administrators. Although many procedures for conducting such evaluations have evolved over the past 30 years or so, most local school personnel are not trained in these techniques. Further, local staff is often even less equipped to evaluate an evaluation -- to analyze the procedures used, data included, and conclusions reached. It is toward this latter problem -- interpreting an evaluation report -- that this proposal is directed.

The recently completed Standards for Evaluations of Educational Programs, Projects, and Materials (McGraw-Hill, 1981) were developed to fill the widely perceived need for standards for program evaluation similar to the APA-AERA-NCME standards for published tests. These Standards for Evaluations, partially funded by the National Institute of Education, may eventually find wide use in guiding the conduct and evaluation of educational evaluations. However, the Standards in their complete form are very comprehensive and broadly based.

Because of this, and because of frequent requests from members for such assistance, in 1978 the International Reading Association (IRA) organized a Committee on the Study of Reading Program Evaluations. This Committee has had as its primary charge to develop "criteria to be used to judge the adequacy of the evidence to support the superiority of a particular reading program..." The Committee used the Standards for Evaluations document and various other evaluation guidelines as starting points in evolving a written set of questions, termed "Guidelines for Reviewing Evaluation Reports," to be used by people attempting to assess an evaluation report. This work went through three revisions by the Committee prior to the field test funded by this grant. The purpose of the field trial was to assess the Guidelines' usefulness and improve upon them before making available a revised set of Guidelines to IRA's general membership.

IRA members are involved in the interpretation and use of literally thousands of reading program evaluations annually. It is our belief that the adaptations we propose to publish, assisted by this funding, will find wide use among our 70,000 members. Further, this work will serve to improve both the understanding of evaluations by school personnel and the interpretation and application of such evaluations on a local level.

Obviously, there is a multitude of types of evaluations available to school personnel. Such variables as whether the evaluation is internally or externally conducted; its duration and scope; whether it contains qualitative, quantitative, or both types of data; whether its purpose is to assess ongoing programs or propose revisions; etc. all distinguish various types of such evaluations. The Guidelines developed clearly cannot be equally valuable in interpreting all types of evaluations. However, they are generally applicable to internal and external, short and long-term, and quantitative and qualitative formal, written evaluations. We believe that they will assist local personnel in assessing the content of most of the wide variety of evaluations with which they come in frequent contact.

II. PROJECT ACTIVITIES

The International Reading Association (IRA) Study of Reading Program Evaluations Committee conducted this project, with assistance from the IRA Research Department. There were essentially eight stages to the project, as outlined below.

1. Finalize the Guidelines draft and Field Test plans. This was completed in two stages. An interim set of Guidelines was developed for review in May, 1980 by Committee members. These Guidelines were then further revised and adapted at an August, 1980 meeting of the Committee in Indianapolis. Also at that time, general plans for the field test were discussed, site selection procedures and materials were finalized, and plans for the orientation for local field test coordinators were made. This activity was completed essentially as originally planned and scheduled. NIE funding financed the Committee's August meeting. Committee members and their current professional affiliations are listed below:

Michael Beck, Chair
Senior Project Director
The Psychological Corp.

Mary Seifert, IRA Headquarters
Research Specialist
International Reading Association

Roger Farr
Director, Smith Research Center
Indiana University

Frank Ferris
Director of Reading
Spencerport (NY) Central Schools

Edward Smith
Basic Skills Improvement Office-DOE
Washington, D.C.

Alden Moe, IRA Board of Directors Liaison
Professor of Education
Purdue University

Rebecca Barr
Department of Education
University of Chicago

Jerome Harste
Assoc. Professor of Reading Education
Indiana University

Susan Stuber
Title I Director
Oscoda (MI) Area Schools

2. Site Selection. During September, 1980, 105 IRA local council presidents were sent information concerning the field test, soliciting the participation of their member school districts. We had anticipated a faster and larger return from this mailing than we actually experienced. For both of these reasons, a second follow-up mailing was made to a somewhat-expanded list of local councils during October. These two mailings resulted in approximately 20 inquiries or expressions of interest. From this number, seven apparent good candidates were selected and contacted for more information concerning their evaluations work. Finally, a list of five districts from four states with varied evaluation experiences, materials, and interest was determined. These districts were formally invited to participate. Unfortunately, one of these districts later was forced to withdraw due to a new local board of education ruling prohibiting travel outside the district by any staff member, regardless of cost reimbursement. Despite discussions with the superintendent on this matter, we were unable to obtain exemption for the project. Due to the lateness of the date and the lack of an appropriate substitute district, the field test was conducted with four sites.

3. Orientation for Local Districts. The schedule for this meeting was finalized when the site selection activities were completed. The meeting was held January 18-20, 1981 at IRA headquarters in Newark, Delaware. At that time, committee members and appropriate IRA staff members met with representatives of the four field test districts. These districts and the local field test coordinators were:

Webster (NY) Central School District -	Janice Strege Reading Coordinator
Rock Falls Township (IL) High School -	Edward Mulvaney Reading Specialist
Houghton Lake (MI) Community Schools -	Ken Roberts Curriculum/Federal Projects Dir.
Downers Grove (IL) Grade School Dist -	Ronald Hale Asst. Supt. for Instruction

During this meeting, the project was explained to the field test coordinators, the draft Guidelines were discussed in detail, a sample evaluation report was "rated" using the Guidelines, and local implementation procedures for the field test were discussed. One IRA committee member was assigned to each local district as its "IRA contact person." This person was in charge of maintaining contact with the local field test coordinator concerning the project and with making a site visit to the district. Based on feedback received at this meeting, additional revisions were made to the Guidelines. The Guidelines revised as a result of the orientation meeting was the set used for the field test. Although we had not originally planned to make a revision at this point, the concensus of opinion was that the revisions were important and that the usefulness of the field test would be increased if the revised version was used. A copy of the field test Guidelines is presented in Appendix A.

The meeting was considered very successful by Committee members and field test coordinators, and we believe it improved significantly the implementation of the field test and the results obtained therefrom. This meeting was funded by the NIE grant. Following the meeting, the IRA public information office sent a press release describing the field test to the hometown newspaper of each Committee member and local field test coordinator. These stories, which mentioned NIE's support for the project, highlighted the cooperation of the local districts in the field test.

4. Field Test. Following the meeting in Delaware, each of the four field test coordinators conducted informal "orientation sessions" concerning the field test with several persons in their district who had interest in or responsibility for some aspect of a reading program evaluation. The purpose for the development of the Guidelines was discussed, as well as the intent that the Guidelines be "generic," rather than being tailored to any particular type of evaluation. Following the orientation, all local personnel were asked to review independently the Guidelines draft document. The focus of the review was on how the Guidelines could best be improved in content, organization, scope, or wording. The reviewers were made aware that both fine, "editorial" suggestions as well as more global, "philosophical" comments were being sought. Reviewers were asked to make their comments with their own local evaluation concerns in mind, while remembering that the final Guidelines were to be used with various types of evaluations. After working through the Guidelines item by item, local reviewers were asked to complete a "Comments & Reactions" page.

A total of 22 local personnel (4-7 per field test district) completed a review of the Guidelines; of these, 19 people completed the "Comments and Reactions" form. According to primary job responsibility, the 22 local reviewers included 6 classroom teachers, 4 principals, 3 reading teachers, 2 reading coordinators, 2 curriculum directors, 2 counselors, and 1 each superintendent, assistant superintendent, and test director. Following the review, the local field test coordinators collected the Guidelines and reactions forms and returned them to their IRA Committee contact person. A summary of the "Comments and Reactions" form is presented in Appendix B.

The review resulted in the Committee reaching several conclusions concerning the Guidelines draft. First, selected portions of the Guidelines required additional clarification, editing, and/or rewriting. Second, despite these many editorial-type suggestions, the reviewers generally found the document to be very comprehensive and clearly organized, at least somewhat readable and reasonable for school use, and sufficiently non-technical. Many reviewers surprisingly indicated that the document would be more (or at least as) useful in planning an evaluation as in interpreting an evaluation already completed. This led to much discussion in later committee work. The primary negative reaction received from reviewers was that the Guidelines were almost too comprehensive and needed to be shortened somewhat, with jargon eliminated.

5. Site Visitations. Two site visits were scheduled for each field test district. The first was to be made by the Committee member assigned to the district; the second visit was to be made by Mary Seifert, the Research Specialist from IRA headquarters. The purpose of the two visits was the same: to meet with/interview local people who had participated in the field test review process to obtain their reactions to and suggestions for the Guidelines. Two visits were used since we considered it likely that somewhat different qualitative evaluations would be obtained if two different people met with district staff at two different times.

The "formats" of the site visits varied between, and often within, the districts. That is, for some site visits, the IRA Committee representative met individually with the local reviewers. In other instances, a group meeting was used in which the local field test coordinator participated. Still other visits were group meetings excluding the local coordinator. Arranging the formats of the site visits was the responsibility of the local coordinator, although they were urged to vary the format for the two visits. Both Committee members and local coordinators reacted well to the format and scope of the visits and no difficulties were encountered. Further, local reviewers gave willingly of their time for these meetings or interviews and in most instances voiced positive reactions both to the project or their district's participation.

Due to various schedule conflicts and the variability of local progress on the review of the Guidelines, one of the four field test sites was visited only once. The Principal Investigator interviewed the local field test coordinator via phone in this instance as the "second visit." Although the failure to make a second visit at this site was unfortunate, it is not likely to have affected materially the results of the field test.

6. Committee Revision of the Guidelines. During the IRA Annual Meeting, Committee members and their Headquarters Research Specialist met on two occasions to revise the Guidelines. Before this time, Committee members and Mary Seifert had submitted written summaries of their site visits for review. Both general topics concerning the purpose, audience, overall tone, and scope of the Guidelines, as well as specific wording and rewritting suggestions were explored. Guidelines drafts reviewed by the various local personnel were used to make revisions in the draft. Major changes made in the Guidelines at this stage, based on the field test results were:

A. The Guidelines were reduced in length by almost a quarter. The number of specific items included in the Guidelines was reduced from 156 to 119. Most of these eliminations were made because of specific recommendations of the field test reviewers, as well as the frequent field test comment that the Guidelines were currently too long and complex.

B. The field test Guidelines contained essentially two sections for each included item -- one indicating the topic was addressed in planning the report, the second indicating the topic was included in the actual report itself. These dual categories were used because of suggestions made at the field test orientation meeting. However, the two sets of ratings were considered by many field test reviewers to be confusing and/or unnecessary. Therefore, in the revised draft, we returned to a single set of categories.

C. The field test Guidelines topics were worded as questions rather than statements. This issue had been discussed at every jucture in the development of the Guidelines. Our best judgement at the time of field test was that questions were preferable. However, the field test results indicated otherwise, and we returned to declarative statements.

D. Various changes were made in the wording of well over half of the included items. Many of these changes had the effect of reducing the length and the "jargon load" of the items, as recommended during the field test.

E. One of the six sections, "Local Conditions," was reduced substantially in length and eliminated as a separate section. Remaining topics were included in the "Focus" section. The "Local Conditions" section was the most heavily criticized portion of the field test draft.

F. The "Information Sources" section was reduced substantially in length (by about 1/3). Many field test reviewers had commented that the draft version was too technical and was unrealistically beyond the scope of typical school personnel.

G. A section was added to the Introduction to more clearly indicate that many of the topics included would not be appropriate for every particular local evaluation. This is, the Guidelines are intended to be more comprehensive than any given report. Therefore, the best use of the Guidelines in a local situation may be to narrow down the entire list of topics to those on which there is some consensus as to importance to the local evaluation.

Along the same lines, a sentence was added stating that responses of "No" did not necessarily imply weaknesses in the evaluation. Some field test reviewers feared people would merely tally the "Yes" and "No" responses and attempt to derive some quantitative "score" for the report. This was to be discouraged.

7. IRA Review of the Revised Guidelines. Appendix C contains a copy of the Guidelines as revised based on the field test. This version was sent to 101 selected IRA volunteers, including chairpersons of each of the 48 IRA Standing and Ad Hoc Committees and all members of the following committees: Evaluation of Tests, Teacher Education, Studies and Research, and Interpretation and Analysis of NAEP Reading Assessment Data. These four committees were selected because their memberships are heavily weighted with people interested in and knowledgeable about evaluation.

The 101 people were asked to "critically review" the draft and make either editorial or substantive comments (see Appendix D). They were also asked to complete a "Comments and Reactions" form similar to that used in the field test. By the deadline for responding, 43 replies had been received, including 41 who completed the evaluation form. These data are summarized in Appendix E.

Concurrently with this review, members of the Committee for the Study of Reading Program Evaluations made additional comments on this most current draft of the Guidelines. These comments were pooled with those of the other reviewers.

Interestingly and probably not surprisingly, there were several parallels between the ratings during and after the field test. Both sets of respondents saw the Guidelines drafts as being well-organized, comprehensive, and useful as an "ideal." There appears to be significant improvement in the post-field test Guidelines in terms of the readability of the questions and in the Information Sources section, two areas that received much comment during the field test. (The alternate explanation is, of course, that the IRA respondents were a more sophisticated group, for whom these areas were less likely to be confusing than they were to typical school personnel. Which, if either, of these explanations is closest to the truth is difficult to determine.)

8. Final Draft. Although this phase of the project is not covered by NIE funding, nor does it fall within the activities of the grant, it is outlined here for closure purposes. As with the field test responses, the most worthwhile portions of the IRA review were the comments item by item and major section by section. These comments were recorded on large summary sheets and used to pull together yet another revision of the Guidelines. This (hopefully) next-to-final draft is currently being reviewed and polished by Committee members and the proposed final draft will be submitted to the IRA Board of Directors in March for approval. This "final draft" version is the one that will be referred to the IRA Publications Committee for dissemination as indicated in the Dissemination and Distribution Plan below.

III. ANTICIPATED USES AND USEFULNESS OF THE GUIDELINES

It may be instructive to convey some qualitative perceptions gained during the field study concerning the anticipated usefulness of the Guidelines in the schools. These perceptions were obtained from the Comments and Reactions forms or, more often, the interviews conducted on site by Committee members or Mary Seifert. Appendix F contains a summary of the comments made by field test respondents to the Comments and Reactions question, "As a result of using these Guidelines, what specific changes, if any, would you make in any future evaluation activities in which you might participate?" This provides some indirect indications of the reactions of our field reviewers as to the Guidelines' usefulness.

Additional impressionistic views concerning the ultimate utility of the Guidelines were obtained during the field test site visits. Selected reactions relevant to this question are listed in Appendix G.

In general, we would make the following observations concerning the anticipated usefulness of the final Guidelines:

1. The Guidelines as a whole are too comprehensive for most local evaluations. Maximum usefulness will probably result if a small group of local personnel select the relevant portions of the Guidelines and the entire set is reduced to some limited subset of items especially pertinent locally. This is now suggested in the Introduction. Most local, non-evaluation people would probably find the complete set overwhelming and, therefore, not especially helpful.

2. The two primary uses of the Guidelines will probably be:

   A. in reviewing plans for an internal evaluation before the evaluation begins -- looking for additional components to be considered.

   B. in judging the quality of an external evaluation after it has been completed

We believe that, for these two purposes, the Guidelines will be helpful as a stand-alone document for most districts.

3. Use of the Guidelines for either the planning or evaluating function would probably be improved if an orientation to the topic and document was available. We will continue to explore this impression through the sessions we will be conducting over the next several months using the Guidelines (see Section V).

IV. PROJECT FUNDING AND STAFFING

The project was funded for the budgeted total of $11,082. Despite greatly increased travel costs during the contract period, the total cost to NIE of this project was $9551.64. Charges to NIE, compared with budget, were as follows:

	Budget	Actual Expenditures
Personnel salaries/benefits, incl. indirect costs	$2,235	$1,640.02
Travel	8,347	7,682.99
Communications	500	228.63
	$11,082	$9,551.64

All professional time for this project, with the exception of that spent by our IRA Headquarters Assistant, was donated to IRA. Records of this time were kept by each Committee member and local field test coordinator during the project. These times totaled to 137 professional days. It is clear that the project was completed at a minimal cost to NIE due largely to the donation of time by these concerned professionals, all of whom we sincerely thank.

In addition to the donated time, an estimated $1600 of actual costs were paid by IRA and employers of the Committee members and field test districts. These expenses included multilith and other copying charges, telephone and postage, clerical and typing support, and miscellaneous supplies.

Figuring donated staff time at an estimated $150 per day, including benefits, and including no additional overhead expenses; it is estimated that the total "cost" of this project was over $31,000, only 30% of which was charged back to NIE. Projects such as this seem to be an ideal way for governmental agencies to derive maximum impact for their shrinking contract/grant funds.

V. DISSEMINATION AND DISTRIBUTION PLAN

Clearly, the ultimate impact of our efforts will be determined by the extent to which the Guidelines are disseminated to the IRA membership and interested others. We have been keenly aware of this throughout our work on the project and have already finalized the following plans:

A. Presentations at Conferences. Sessions have been proposed and accepted for presentation as of this date at three conferences. Presenters will be either Committee members or participants in the field test. The currently scheduled presentations are:

MEETING	DATE	PRESENTER(S)
Illinois Reading Council	March 18-20, 1982	Ed Mulvaney
Michigan Compensatory Education Directors Workshop	August 19-20, 1982	Roger Farr, Susan Stuber
New York State Reading Conference	November 3-6, 1982	Mike Beck, Jan Strege, Mary Seifert

All of these presentations will be workshop-type sessions in which the development of the Guidelines will be discussed, followed by a hands-on discussion of the Guidelines with participants. Feedback from these sessions will assist IRA in additional refinement of the Guidelines for publication.

In addition to these currently scheduled sessions, two proposals for additional presentations are planned. These presentations will be at the IRA national convention in Anaheim in May, 1983 and at the IRA Great Lakes regional convention in October, 1983. Attendance at these two conventions will exceed 20,000, thus providing an excellent forum for the dissemination of the Guidelines. These 1983 presentations will be made using the final version of the Guidelines which, hopefully, will have been published by IRA by that time (see below). NIE's support of the field test activities will be acknowledged in each professional presentation made.

B. Publications. The final draft of the Guidelines will be submitted by the Committee for approval at the April Board of Directors meeting. We will request at that time through the Committee's Board Liaison, Alden Moe, that the Guidelines be referred to the Publications Committee for printing and distribution. Our intent is that the Guidelines be distributed in booklet form as a regular publication of IRA. For the several booklets of a similar nature already published by IRA, over 350,000 copies have been distributed. These booklets are regularly promoted in various IRA publications received by the Association's 50,000 individual and 20,000 institutional members.

Following approval, the Publications Committee will decide on the means of distribution for the Guidelines via the IRA publications system. IRA will send a notice of the availability of the Guidelines to each of IRA's State Coordinators, asking that the notice be published or distributed through whatever channels are appropriate to the several hundred Local Councils. Finally, an article describing the development of the Guidelines will be submitted to The Reading Teacher or Journal of Reading, IRA journals, for publication. This article will also contribute to the dissemination of the work to interested professionals. All of these efforts should ensure the wide distribution of the Guidelines to IRA members.

Copies of the Guidelines in their final form will be submitted to ERIC and to NIE for possible distribution through the National Diffusion Network. IRA will also announce the availability of the materials to a wide variety of professional educational associations with which it maintains regular contacts. These groups will include curriculum organizations such as the National Council of Teachers of English and the Association for Supervision and Curriculum Development; research organizations such as the American Educational Research Association, National Council on Measurement in Education, the Evaluation Network; and other education agencies such as the Council for Exceptional Children, American Federation of Teachers, National Education Association, and the American Association of School Administrators. Finally the Guidelines will be shared by IRA with fellow members of the National Consortium on Testing, a national group of over 40 professional organizations concerned with test-related issues. All of these activities will broaden greatly the potential impact of the Guidelines on the general education and research community.

ira INTERNATIONAL READING ASSOCIATION, 800 Barksdale Road PO Box 8139 Newark Delaware 19711 USA
ASSOCIATION INTERNATIONALE POUR LA LECTURE Telephone 302 731 1600
ASOCIACION INTERNACIONAL DE LECTURA Cable Reading Newark Delaware

COMMITTEE CORRESPONDENCE

STUDY OF READING PROGRAM EVALUATIONS COMMITTEE

Guidelines Field Test

The purpose of the attached Guidelines document is to assist local school district personnel in planning, conducting, and interpreting an evaluation of a reading program. The Guidelines are designed primarily to provide a frame of reference for assessing the content and scope of such an evaluation.

The format of the Guidelines is a set of questions, followed by boxes in which you would respond "Yes" or "No" or "Not Applicable." The boxes are further divided into Planning and Report sections. A given procedure might have been considered or documented in the planning of an evaluation, but not included in the final report. In such a case, you would check "Yes" for Planning and "No" under Report. Obviously, other combinations of responses are possible. Questions that do not seem applicable to your local evaluation efforts should be so checked. Finally, you need not necessarily feel constrained by the Yes-No format we provide. Other responses or annotations that may be more appropriate for you should be used if desired.

It should be emphasized that we are not trying to evaluate any of your local reports or plans. Thus your actual answers to the Guidelines questions are not our interest, although we hope they will be helpful to you in either planning or reviewing an evaluation. We are simply looking for feedback from you concerning how useful these Guidelines are in your own local situation -- and how we might make them better.

As you work through the Guidelines in terms of your local evaluation, please feel free to write any notes or comments directly on any questions or sections that are unclear, ambiguous, incomplete, etc. Such comments will be of great value to us when we revise and finalize the draft. After you have completed work on the Guidelines, please complete the "Comments & Reactions" pages at the end of this document. Then return your comments and the annotated Guidelines to

Please return them by

Many thanks for your assistance with this project.

1/81 Draft

I.R.A. GUIDELINES FOR REVIEWING EVALUATION REPORTS[1]

The purpose of these guidelines is to assist individuals in the analysis of reading program evaluations. It is important to realize that not all sections of this document will necessarily apply to a given evaluation. Conversely, certain evaluations might include content not addressed by these guidelines. It is important, however, that the reading program being evaluated is described in terms of the program scope and sequence, and also in terms of its key elements. Inherent in this description would be a statement of the:

- purpose(s) of the evaluation
- focus of the evaluation
- local conditions or context in which the evaluation occurred
- procedures followed in conducting the evaluation
- sources of information used in the evaluation
- conclusions/recommendations based on the evaluation results

It should be noted that responses of "No" to questions in the guidelines do not imply weaknesses in the evaluation being reviewed. Even in excellent and comprehensive evaluations, many questions might be answered "No", depending on the purpose and scope of the given evaluation.

Meaningful evaluation always results in the improvement of instruction for students. It is essentially a plan of/for action that gives perspective to the curriculum.

---

[1] International Reading Association - Study of Reading Program Evaluations Committee

Michael Beck (Chair), Rebecca Barr, Roger Farr, Frank Ferris, Jerome Harste, Edward Smith, Susan Stuber

I. PURPOSE

The report should provide the reasons for the conduct of the evaluation. Included should be a statement of why the evaluation was being conducted and who decided on the advisability and the purposes of the study.

	Planning		Report		Not Applic
	YES	NO	YES	NO	
-Is (Are) the reason(s) for conducting the evaluation explicitly stated?					
-Was the purpose recorded before the evaluation took place?					
-Was the purpose communicated to various concerned groups prior to the evaluation?					
-Does the report indicate who decided the evaluation was advisable?					
-Is the degree to which various affected groups agreed with the purpose(s) of the evaluation discussed?					
-Are the probable outcomes of the evaluation suggested?					
-Is there an indication of what actions might be taken based on the various possible outcomes?					
-Does the report indicate that these possible actions are realistic and practical?					

II. FOCUS

The project, program, or the materials being evaluated should be described clearly and completely. This description should include such elements as the philosophy, including goals, objectives and definition of reading; as well as specifics such as time, site, students, personnel, instructional methods, and materials.

	Planning		Report		Not Applic
	YES	NO	YES	NO	
-If a "reading program" is being evaluated, does the evaluation focus on all aspects (text, other material, personnel, methods, etc.) of the program?					
-Is the "program" defined or explained?					
-Is the primary focus of the evaluation clearly presented?					
-Does the evaluation state what aspects or portions of the program were not studied, and why?					
-Are reasons for selecting these aspects and not others explicitly stated?					
-Are different evaluation procedures applied to various portions of the program?					
-Does the report indicate the time span of the evaluation in relation to the time span of the program described?					
-Is a reason given for conducting the evaluation during this particular time span?					
-Is the particular philosophy of reading instruction embodied in the program stated? For example: Is reading seen as the recording process of turning graphic symbols into the sounds of our language? Is reading seen as the acquisition of a set of identifiable but discrete skills where mastery of lower-order skills is seen as prerequisite to mastery of higher-order skills? Is reading seen as a whole-language process in which meaning is central and particular skills are never isolated but become important only as they relate to this focus on meaning?					

	Planning		Report		Not
	YES	NO	YES	NO	Appli
A. FUNDING					
-Has the school district examined the funding pattern of its reading program to determine the sources of its funds? (Funding sources could include local dollars, compensatory education funds, grant money, allocations from state and federal agencies, etc.)					
-Has the district determined where and how those funds are spent? (Basal series, support materials, salaries of reading teachers, staff development activities, library books, building consultants, etc.)					
B. SITES					
-Are the sites (buildings or classrooms) in which the evaluation took place described (in terms of number, representativeness, reasons for selection, etc.)?					
-Is the "program" essentially identical across sites?					
-If not, is the evaluation separated by site?					
-Are the space/size and physical arrangements of the sites and their potential effect on the program considered in the evaluation?					
C. STUDENTS					
-Are relevant characteristics of the students involved (e.g., sex, age, entry capabilities, language diversity, past learning rate) described?					
-Are these characteristics analyzed in the evaluation?					
-Is the stability of the group over the period of the evaluation described?					
-If more than one site is evaluated, are students across the different sites "comparable" in terms of relevant variables?					
-Have the students previously been involved with this program?					
-Was feedback provided to students concerning the evaluation?					
-Are feedback procedures described?					

	Planning		Report		Not Appli
	YES	NO	YES	NO	
-Is the student-to-teacher or student-staff ratio included?					
D. STAFF					
-Are the general characteristics (number, role, age, education, special training, experience, etc.) of the staff presented?					
-Are any of these factors analyzed as part of the evaluation?					
-Is the "managerial" or organizational scheme from teacher to supervisor presented?					
-Was this scheme part of the evaluation?					
-Is the staff role in selecting and implementing the reading program described?					
-Are staff development activities based on relevant research and the identified needs of various levels of personnel described?					
-Are criteria for determining staff development activities for the district stated?					
-Is the role of the staff in the <u>selection</u> of staff development activities assessed?					
-In the <u>evaluation</u> of staff development activities?					
-Is staff self-assessment part of the overall evaluation?					
-Was feedback provided to staff concerning the evaluation?					
-Are feedback procedures described?					
E. INSTRUCTIONAL METHODS					
-Is there a description of all instructional methods being evaluated?					
-Is a rationale given for the use of those particular instructional methods?					
-Is an attempt made to match diagnosed student needs to appropriate instructional methods?					
-Are procedures for selecting methods to correspond to students needs assessed?					

	Planning		Report		Not
	YES	NO	YES	NO	Applic
-Is the degree of correspondence between the instructional methods used and the district's philosophy of learning to read described? Were a variety of instructional methods used with students?					
F. MATERIALS					
-Are the materials and their important characteristics (e.g., length, difficulty, format, previous use) described?					
-Are the procedures and personnel involved in developing or selecting the materials indicated?					
-Were the materials developed for use in programs similar to the one being evaluated?					
-If not, is a rationale provided for their use?					
-Were the materials used essentially in their entirety?					
-Were changes or adaptations made in the materials for program use?					
-Is there a description given of how the portions to be used were selected?					
-Are previous evaluations of these materails in similar settings available?					
-If so, are they summarized or referred to?					
-Is the relationship between instructional materials and strategies used and the district's reading program goals described?					
-Has the appropriateness of the materials to the goals of the program been determined?					
-Are the procedures used in the making of this determination discussed or referenced?					
-Are procedures described for selecting materials to correspond to student needs?					

III. LOCAL CONDITIONS

Exceptional conditions occuring during the evaluation period should be carefully described so the findings of the evaluation can be interpreted in light of them. Such internal descriptions should enable the reader to judge whether the findings might be similar in other contexts and should provide clear limitations to the generalizability of the results of other situations.

	Planning		Report		Not
	YES	NO	YES	NO	Applic
-Are any exceptional local conditions described that might have affected materially the evaluation?					
-Did the context of the evaluation provide a true picture of the reading program?					
-If not, was this accounted for in describing the limitations or interpretations of the study?					
-Were staff members and students aware of the evaluation while it was being conducted?					
-If so, is there reason to believe such awareness may have significantly altered the evaluation?					
-Were the significant internal "political" factors impacting the program at the time the evaluation was conducted?					
-If so, are the results of the evaluation discussed in terms of the potential impact of these factors?					
-Were there any special personnel factors that might have affected the evaluation -- e.g., a teacher strike, staff firings, student problems, new key administrators?					
-If so, is the potential impact of these factors discussed?					
-Were there any special physical factors that might have affected the evaluation -- e.g., facility problems, late delivery of materials, routine start-up problems?					
-If so, is the potential impact of these factors discussed?					
- Were there any important social or community factors that might have affected the evaluation -- e.g., school bond issues, exceptional local economic conditions, school legal issues?					
-If so, is the potential impact of those factors discussed?					

IV. PROCEDURES

	Planning		Report		Not Applic
	YES	NO	YES	NO	
-Is the design of the evaluation described in sufficient detail such that the study could essentially be replicated, given only the information presented?					
-Are limitations of the procedures discussed?					
-Did the evaluation include an analysis of special subgroups of students (e.g., control and experiental groups, Title I and non-Title I students)?					
-If so, is the evaluation designed to determine whether any results are a direct reflection of the reading program?					
-Would similar changes have resulted with no special program or from the instruction normally provided?					
-Is this possibility controlled for or discussed in the report?					
-Is the evaluation designed to isolate key components of the reading program for further analysis?					
-Was a timeline for the evaluation established during the planning stages?					
-If so, is evidence presented that the timeline was followed, or are significant departures noted?					
-Is (Are) the person(s) who conducted the evaluation identified?					
-Is a clear description of the role and duties of the evaluator(s) given?					
-Are reasons presented for the selection of the evaluator(s) presented?					
-Did staff have a role in selecting the evaluators?					
-Are there other data indicating support for the evaluation?					
-Are the evaluators' biases (methodological or instructional) noted?					
-Is there evidence that biases are controlled or minimized?					

IV. PROCEDURES (Continued)

Classify the evaluator(s) as:	Internal (permanent district employees) External (outside consultants) "Mixed" (Internal and External)

Now complete the appropriate sections below: Part A for Internal evaluators
Part B for External evaluators
Part A & B for "Mixed" evaluators

	Planning YES	Planning NO	Report YES	Report NO	Not Applic
A. INTERNAL EVALUATORS					
-Are the evaluators staff members of the program being studied?					
-Are their previous experiences with similar types of evaluations described?					
-Is evidence presented concerning their technical expertise (in evaluation and in reading)?					
-Is evidence presented concerning the ability of the person(s) to evaluate objectively the reading program?					
B. EXTERNAL EVALUATORS					
-Has the person had previous experience evaluating programs of a similar nature?					
-Is there evidence of qualification in terms of knowledge of reading, instruction, and evaluation?					
-Is the evaluator accessible for follow-up consultation?					
-Is an explanation given of why this person was selected for the project?					
-Does the report indicate who within the district had a role in choosing this evaluator?					
-Was the amount of time the person spent in the district studying the program specified?					

V. INFORMATION SOURCES

The evaluation should describe as completely as practical the sources of the information used in the report. Included should be a description of the various sources of data used in the evaluation, the characteristics of these instruments or procedures, and why they were included.

A. DATA COLLECTED

	Planning		Report		Not
	YES	NO	YES	NO	Applic
-Are interpretations in the evaluation based on multiple sources of information -- both quantitative and qualitative? (e.g., attitude, achievement, ability test results; interviews; questionnaires, rating forms)					
-Is quantitative evidence summarized clearly?					
-Is the meaning of any statistical tests discussed?					
-Is <u>educational</u> significance of the results discussed?					
-Are qualitative data included?					
-Are they summarized clearly?					
-Are conclusions based solely or primarily on single sets of data?					
-Do other data support the conclusions?					
-Are any longitudinal data provided to support the conclusions?					
-Are suggestions made for follow-up data collection to collaborate the results or to analyze long-term trends?					
-Were data collected from all involved groups (students, teachers, parents, various administrators, board of education)?					
-Were data gathered as near the beginning and end of the period to be evaluated as possible?					
-Would additional data concerning learning or attitudes -- of students or professional staff -- have been important to the evaluation?					
-If achievement test data are included, is a rationale given for the specific test that was used?					

A. DATA COLLECTION (Continued)

	Planning		Report		Not Applic
	YES	NO	YES	NO	
-Is evidence provided that the test content matches that of the instructional program?					
-Is this test series typically used in the system?					
-If not, is the reason for its use in the evaluation presented?					
-Were data analyzed and/or reported in terms of the important purposes of the evaluation?					
-Was the level at which results were summarized consistent both with the purposes of the evaluation and the instruments included? (e.g., a "reading comprehension" test was used, but results were applied to a phonics instructional program.)					
-Are summaries provided of the results of each set of information collected?					
-Is some evidence provided of "quality control" in the collection of the data?					
-In the analysis or processing of the data?					

B. TECHNICAL CHARACTERISTICS OF THE DATA

Validity

-Is a rationale provided for the selection and use of the instruments included in the evaluation?					
-Is some documentation given for the validity of the instruments used for the purposes they serve in the evaluation?					
-For achievement measures, is evidence of content validity (appropriate match of program and test content) provided?					
-For ability or aptitude measures, is there evidence or criterion or other appropriate validity?					
-For other measures (e.g., questionnaires, surveys, attitude instruments), is some evidence of validity or sound instrument construction procedures presented?					

B. TECHNICAL CHARACTERISTICS OF THE DATA

Validity (Continued)

	Planning YES	Planning NO	Report YES	Report NO	Not Applic
-For instruments constructed specifically for this evaluation, is there evidence of any validation or tryout procedures of the measure during its construction?					
-Was there a pilot testing or tryout of such instruments before their use in the program?					

Reliability

	Planning YES	Planning NO	Report YES	Report NO	Not Applic
-If any instruments used were constructed specifically or modified significantly for this evaluation, is reliability evidence presented?					
-If not, does the construction (question clarity, length, required tasks) suggest that the information obtained is reliable?					
-Were reliability data collected as part of the evaluation or in a separate study?					
-If in a separate study, is there evidence that the setting and groups used for calculating reliability were similar to those of the evaluation?					
-If published instruments were used, are reliability data available?					
-Were the data gathered on groups comparable to those being evaluated?					
-If not, were reliabilities provided for the evaluation sample?					

VI. CONCLUSIONS

The summary of the evaluation report should present the assumptions made in the evaluation and the conclusions reached. Such conclusions need not all be supported by all sets of data collected; nevertheless, conclusions should be referenced to the results wherever possible. Conclusions not supported by the data should be clearly identified. Finally, a good report should conclude with a proposed plan of action for either implementing the recommendations or initiating appropriate follow-up action.

	Planning		Report		Not Applic
	YES	NO	YES	NO	
-Does the report contain a complete list of assumptions made in designing and conducting the evaluation? (e.g., it is assumed that the XYZ Reading Test is an adequate and complete measure of comprehension.)					
-Does the report clearly state the conclusions reached?					
-Are the conclusions reached directly tied to the results obtained from the evaluation?					
-Are the evaluation results that yielded inconclusive findings clearly indicated?					
-If generalizations or interpretations are made that are not based on solid data, are they clearly labled as such?					
-For each conclusion drawn, are the supportive evidence -- qualititive and/or quantitative -- for this conclusion referred to?					
-Do the conclusions follow naturally from what was actually measured?					
-Do the conclusions include the reasoning that led to them?					
-Are inconclusive areas indicated and reasonable interpretations provided for these?					
-Are other reasonable interpretations provided?					
-Are limitations of the evaluation presented frankly?					
-Are these conclusions stated in a manner that leads to a plan of action?					
-Is evidence available to show that a relevant action play has been established and put into practice?					

VI. CONCLUSIONS (Continued)

	Planning YES	Planning NO	Report YES	Report NO	Not Applic
-Is program staff aware of the intended implementation of any findings including new goals, programs, procedures, materials, appropriate timelines, staff development activities, etc.					
-Are conflicting points of view concerning the program and the evaluation results presented fairly?					
-Are "minority" views represented?					
-Does the report explain why these views are not considered "correct," without attacking these views in a non-objective manner?					
-Is some assurance given that the report was not modified against the best judgment of the evaluator to reflect the views of school staff with vested interest concerning the program?					
-Have the evaluators analyzed and interpreted specific results in light of the broader perspective of reading and the teaching of reading?					
-Is there a plan for the dissemination of information derived from the evaluation?					

APPENDIX B. Summary of Comments on Field Test Guidelines by Local District Reviewers (N = 19)

(OPTIONAL)

Name: ______________________

Title: ______________________

COMMENTS & REACTIONS

I. In order to be useful for different types of evaluations, these Guidelines must be clear, concise, and comprehensive. Please give your opinion on the following points. Feel free to suggest changes or additions or deletions.

	Very	Somewhat	No
A) Are the Guideline questions readable?	5	13	
B) Is the intent of the questions clear?	6	12	
C) Is the organization of the topics clear?	16	2	
D) Are the individual Guidelines sufficiently non-technical?	[illegible]	8	4
E) Are the criteria realistic for a typical school evaluation?	4	9	2
F) Are the Guidelines comprehensive?	18		
F) Is the coverage <u>too broad</u> for a typical school evaluation?	3	8	7

Please comment on any of the above: ______________________

II. Indicate below your judgment of the usefulness of the Guidelines in your own situation.

A) To what extent would you find the Guidelines useful for:	Very Adequate	Adequate	Inadequate
1. Planning an evaluation?	14	2	1
2. Conducting an evaluation?	9	5	[illegible]
3. Interpreting an evaluation?	11	4	2

Comments: ______________________

B) Rate each of the major sections of the Guidelines in terms of its usefulness in planning or interpreting an evaluation report:

	Very Adequate	Adequate	Inadequate
I) Purpose	15	2	
II) Focus	12	5	
III) Local Conditions	6	8	3
IV) Procedures	11	6	
V) Information Sources	9	7	1
VI) Conclusions	12	5	

III. What would you consider the major strengths and weaknesses of the Guidelines in their current form?

Strengths: provides a good base; thorough; Sections I and II are excellent; well-organized; relevant; very systematic

Weaknesses: language and wording; technical; time-consuming; criteria too broad; too complex for typical school; neglect of teacher/process variables

IV. Can you suggest any additional questions or important areas that should be included in the Guidelines?

Needs section on use of accepted practices

Needs more examples

Follow-up should be included

How are student needs assessed

V. As a result of using these Guidelines, what specific changes, if any, would you make in any future evaluation activities in which you might participate?

Look at more aspects of the reading "program"

Better dissemination procedures

Demand a more thorough and systematic job by the outside group--now I know what to look for in a good job

Get more input from all involved groups

THANK YOU VERY MUCH FOR YOUR COMMENTS.

APPENDIX C. Draft of Final IRA Guidelines

I.R.A. GUIDELINES FOR REVIEWING EVALUATION REPORTS

The purpose of these guidelines is to assist individuals in the analysis of reading program evaluations. It is important to realize that not all sections of this document will necessarily apply to a given evaluation. Conversely, certain evaluations might include content not addressed by these guidelines. It is important, however, that the reading program being evaluated is described in terms of the program scope and sequence, and also in terms of its key elements. Inherent in this description would be a statement of the:

- purpose(s) of the evaluation
- focus of the evaluation
- procedures followed in conducting the evaluation
- sources of information used in the evaluation
- conclusions/recommendations based on the evaluation results

Responses of "No" to questions in the guidelines do not imply weaknesses in the evaluation being reviewed. Even in excellent and comprehensive evaluations, many questions might be answered "No", depending on the purpose and scope of the given evaluation.

Meaningful evaluation always results in the improvement of instruction for students. A useful evaluation is essentially a plan for action that gives perspective to the curriculum.

International Reading Association - Study of Reading Program Evaluations Committee

Michael Beck (Chair), Alden Moe (Board Liaison), Rebecca Barr, Roger Farr, Frank Ferris, Gerald Maring, Edward Smith, Susan Stuber

I. PURPOSE

The report should provide the reasons for the conduct of the evaluation. Included should be a statement of why the evaluation was conducted and who decided on the advisability and the purposes of the study.

	YES	NO	NOT APPLIC.
A. The reason (s) for conducting the evaluation is (are) explicitly stated.			
-The purpose was recorded before the evaluation took place.			
-The purpose was communicated to various concerned groups prior to the evaluation.			
B. The report indicates who decided to conduct the evaluation.			
C. Probable outcomes of the evaluation are suggested.			
-Action that might be taken based on the various possible outcomes are indicated.			
-The report indicates that these possible actions are realistic and practical.			

II. FOCUS

The project, program, or the materials evaluated should be described clearly and completely. The description should include such elements as the philosophy, including goals, objectives and definition of reading; as well as specifics such as time, site, students, personnel, instructional methods and materials.

	YES	NO	NOT APPLIC.
A. The primary focus of the evaluation is clearly presented.			
-If a reading program is being evaluated, the evaluation focuses on all important aspects (text, personnel, methods, etc.) of the program.			
-The "reading program" is defined or explained, detailing all components.			
-The evaluation states what aspects or portions of the program were not studied, and why.			
-Different evaluation procedures are used with various portions of the program.			
B. A reason is given for conducting the evaluation during this particular time span.			
C. The particular philosophy of reading instruction embodied in the program is stated.			
D. FUNDING The funding pattern of the reading program as been examined to determine the sources of its funds. (Funding sources could include local dollars, state/federal allocations, etc.)			
-The evaluation includes an analysis of where and how funds are spent (e.g., basal series, support materials, salaries, staff development activities, library books, consultants).			
E. SITES Building or classroom sites in which the evaluation took place are described in terms of number, representativeness, reasons for selection, etc.			
-The "program" is essentially identical across sites.			
-If not, the evaluation is separated by site.			

	YES	NO	NOT APPLIC.
F. LOCAL CONDITIONS Exceptional conditions occurring during the evaluation period are described so the findings of the evaluation can be interpreted in light of them. Such descriptions should enable the reader to judge whether the findings might be similar in other contexts and provide clear limitations to the generalizability of the results to other situations.			
The setting in which the evaluation took place provided a true picture of the reading program.			
-If not, this was accounted for in describing the limitations or interpretations of the study.			
Were staff members and students aware of the evaluation while it was being conducted?			
-If so, might this awareness have significantly altered the evaluation?			
Any exceptional local conditions that might have affected the evaluation are described. (E.g. personnel factors such as strikes or firings; physical factors such as facility or start-up problems; or social/community factors such as bond issues or exceptional economic conditions.)			
-The potential impact of any of these personnel, physical, or social factors is discussed in the evaluation.			
G. STUDENTS Relevant characteristics of the students involved as sex, age, entry capabilities, language diversity, etc. are described.			
-These characteristics are analyzed in the evaluation.			
The stability of the group over the period of the evaluation is described.			
Students across the different sites are "comparable" in terms of relevant variables.			
-The method for determining comparability is described.			

H. STAFF

	YES	NO	NOT APPLIC.
The number, role, education, experience, other relevant characteristics of the staff are presented.			
The "managerial" or organizational scheme from teacher to supervisor is presented.			
The staff role in selecting and implementing the reading program is described.			
Staff development activities were provided for personnel involved with the program.			
-The role of the staff in the <u>selection</u> of staff development activities is discussed.			
Feedback was provided to staff concerning the evaluation.			
-Feedback procedures are described.			
I. INSTRUCTIONAL METHODS			
Are instructional methods being evaluated?			
-If so, these methods are described.			
-A rationale is given for the <u>use</u> of those particular instructional methods.			
-An attempt is made to diagnose student needs and match students to appropriate instructional methods.			
A variety of instructional methods were used with students.			
-The correspondence between methods used and the district's philosophy of reading is described.			

J. MATERIALS

	YES	NO	NOT APPLIC.
Materials used in the program and their important characteristics (e.g., length, difficulty, format, previous use, commercial/unpublished) are described.			
The procedures and personnel involved in developing or selecting the materials are indicated.			
The materials used were developed for use in comparable situations and with comparable student groups.			
-Any changes, adaptations, or use of selected portions of the materials are described.			
Previous evaluations of these materials, if any, in similar settings are summarized or referred to.			
The relationship among instructional materials, strategies used, and the district's reading program goals is described.			
-The appropriateness of the materials to the goals of the program has been determined.			
-Procedures used in making this determination are discussed or referenced.			

III. PROCEDURES

The procedures and/ or experimental design used in the evaluation should be discussed in enough detail so readers can judge the appropriateness of the conclusions reached. Evaluators should be named, along with a description of how they became associated with the evaluation and their relevant background and experiences.

	YES	NO	NOT APPLIC.
A. The design of the evaluation is described in sufficient detail to enable its replication.			
-Limitations of the procedures are discussed.			
B. The evaluation includes an analysis of different groups of students (e.g., control and experimental groups, Title I and non-Title I students).			
-Would similar changes have resulted in the absences of the special program or as a result of the instruction normally provided?			
-Is this possibility controlled for or discussed in the report?			
C. A timeline for the evaluation was established during the planning stages.			
-The timeline was followed or significant departures are noted.			
D. The evaluator(s) is (are) identified.			
-A clear description of the role and duties of the evaluator(s) is given.			
-Reasons for the selection of the evaluator(s) are given.			
-The evaluators' biases (methodological or instructional) are noted.			

III. PROCEDURES (continued)

Classify the evaluator(s) as:		
	Internal	(permanent district employees)
	External	(outside consultants)
	"Mixed"	(internal and external)

Complete the appropriate sections below: Part E for Internal evaluators
Part F for External evaluators
Part E & F for "Mixed" evaluators

	YES	NO	NOT APPLIC.
E. INTERNAL EVALUATORS			
Are the evaluators also staff members of the program being studied?			
Their previous experiences with similar types of evaluations are described.			
-Evidence is presented concerning their technical expertise (in evaluation and in reading).			
-Evidence is presented concerning their ability to be objective in their evaluation of the program.			
F. EXTERNAL EVALUATORS			
The evaluators have had previous experience studying programs of a similar nature.			
-Evidence of expertise in reading, instruction, and evaluation is presented.			
-The evaluators are accessible for follow-up consultation.			
-An explanation is given of why the evaluators were selected for the project.			
-The report indicates which district personnel had a role in choosing the evaluators.			
-Data indicating staff and administrator support for the evaluator are presented.			
The amount of time the evaluators spent both in the district and in total studying the program is specified.			

IV. INFORMATION SOURCES

The report should describe as completely as practical the sources of the information used in the evaluation. Included should be a description of the data used, the characteristics of these instruments or procedures, and why they were included.

A. DATA COLLECTED	YES	NO	NOT APPLIC.
Interpretations in the evaluation are based on multiple sources of information -- both quantitative and qualitative (e.g., attitude, achievement, ability test results; interviews; questionnaires; rating forms).			
-Quantitative evidence is summarized clearly.			
-The meaning of any statistical tests is discussed.			
-Educational implications of the results are discussed.			
-Qualitative data are included.			
-They are summarized clearly.			
-Conclusions are based solely or primarily on single sets of data.			
Suggestions are made for follow-up data collection to collaborate the results or to analyze long-term trends.			
Data were collected from all involved groups (e.g., students, teachers, parents, various administrators, board of education).			
The time at which data were collected is specified.			
The manner in which results are summarized is consistent both with the purposes of the evaluation and the instruments included (e.g., a "reading comprehension" test was used, but results were applied to a phonics instructional program).			
Summaries are provided of the results of each set of information collected.			
Evidence is provided of "quality control" in the collection of the data.			
-In the analysis or processing of the data.			

B. TECHNICAL CHARACTERISTICS OF THE DATA	YES	NO	NOT APPLIC.
Validity			
A rationale is provided for the selection and use of the instruments included in the evaluation.			
Documentation is given of the validity of the instruments used for the purposes they serve in the evaluation.			
-If an achievement test was used, evidence is provided that the test content closely matches that of the instructional program.			
The test series used in the evaluation is typically used in the system.			
-If not, reasons for its use in the evaluation are specified.			
For locally developed instruments, evidence is provided of validation procedures of the measure during its construction.			
-Such instruments were pilot tested before their use in the program.			
Reliability			
For instruments used that were constructed specifically or modified significantly for this evaluation, reliability evidence is presented.			
-If not, does the construction (question clarity, length, required tasks) suggest that the information obtained is reliable?			
If published instruments were used, reliability data are reported.			
-Data were gathered on groups comparable to those being evaluated.			
-If not, reliability data are provided for the sample of students used in the evaluation.			

V. CONCLUSIONS

The summary should present the assumptions made in the evaluation and the conclusions reached. Conclusions need not all be supported by all sets of data collected; nevertheless, conclusions should be referenced to the results wherever possible. Conclusions not supported by the data should be clearly identified. Finally, a good report concludes with a proposed plan of action for either implementing the recommendations or initiating appropriate follow-up action.

	YES	NO	NOT APPLIC.
The report contains a complete list of assumptions made in designing and conducting the evaluation (e.g., it is assumed that the XYZ Reading Test is an adequate and complete measure of comprehension).			
The report clearly states the conclusions reached.			
-Conclusions reached are limited to the results obtained from the evaluation.			
-Results that yielded inconclusive findings are clearly indicated.			
-Generalizations or interpretations made that are not based on solid data, are clearly labeled as such.			
For each conclusion drawn, the supportive evidence -- qualitative and/or quantitative -- is referenced.			
-The conclusions follow from what was actually measured.			
-The conclusions include the reasoning that led to them.			
-Inconclusive areas are indicated and reasonable interpretations provided for these.			
-Limitations of the evaluation are presented.			
-Conclusions are stated in a manner that leads to a plan of action.			
-Evidence is provided to show that a relevant plan of action has been established and implemented.			
-Program staff is aware of the intended implementation of any findings including new goals, procedures, materials, timelines, activities, etc.			
Conflicting points of view concerning the program and the evaluation results are presented fairly.			

V. CONCLUSIONS (Continued)

	YES	NO	NOT APPLIC.
-"Minority" views are represented.			
-The report presents evidence as to why these views are not considered "correct," without attacking these views in a non-objective manner.			
Some assurance is given that the report was not modified to reflect the views of school staff having vested interest concerning the program.			
The is a concrete plan for the dissemination of results of the evaluation to all concerned groups (e.g., students, staff, the public).			

INTERNATIONAL READING ASSOCIATION, 800 Barksdale Road PO Box 8139 Newark Delaware 19711 USA
ASSOCIATION INTERNATIONALE POUR LA LECTURE
ASOCIACION INTERNACIONAL DE LECTURA
Telephone 302 731 1600
Cable Reading Newark Delaware

September 2, 1981

MEMO TO: IRA Committee Members

FROM: Michael D. Beck, Chair MB
Committee for the Study of Reading Program Evaluations

RE: Guidelines for Reviewing Evaluation Reports

Over the past several months, the Committee for the Study of Reading Program Evaluations has been working on a set of "Guidelines" for reviewing evaluation reports. This work, which included a field test in four school systems, was supported by a grant from the National Institute of Education. The feedback we received from the field test participants was used to revise an earlier draft of the enclosed Guidelines.

We would greatly appreciate your assistance at this stage of our work. Will you please critically review the enclosed draft, making any comments or suggestions directly on the document? Your suggestions might be of an editorial or wording nature, or more substantive—either are welcome. After you have completed the review, please fill out the short "Comments & Reactions" at the end of the document. The total amount of time this review will take from your busy schedule should not exceed 60 minutes. Again, feel free to respond both on the pages of the Guidelines and on the "Comments & Reactions" section at the end.

It is important that we receive your reactions no later than September 24, so they can be taken into account in our final draft. Please use the enclosed envelope to return the document before that date. Thank you for your valued assistance.

MDB/ss

Enclosures

APPENDIX E. Summary of Comments on Draft of Final Guidelines by IRA Committee Members (N = 41)

Name: _______________ (Optional)

Committee: _______________

COMMENTS & REACTIONS

I. In order to be useful for different types of evaluations, these Guidelines must be clear, concise, and comprehensive. Please give your opinion on the following points. Feel free to suggest changes or additions or deletions.

	Very	Somewhat	No	?
A) Are the Guideline questions readable?	29	9		
B) Is the intent of the questions clear?	27	10		1
C) Is the organization of the topics clear?	32	5	1	
D) Are the individual Guidelines sufficiently non-technical?	28	9	1	
E) Are the criteria realistic for a typical school evaluation?	18	15	3	2
F) Are the Guidelines comprehensive?	32	6		
G) Is the coverage too broad for a typical school evaluation?	7	18	11	2

Please comment on any of the above: May be too comprehensive (6)

Stylistic/editorial/format suggestions (5)

II. Indicate below your judgment of the Guidelines' usefulness in a typical school situation.

A) To what extent would you find the Guidelines useful for:	Very Adequate	Adequate	Inadequate
1. Planning an evaluation?	27	11	0
2. Conducting an evaluation?	20	15	3
3. Interpreting an evaluation?	23	12	2

Comments: This is a good "Model" for an example of the "ideal" that should be kept in mind while planning/executing an evaluation (3

B) How might the practicality of the Guidelines for a typical school be improved?

Reduce the length (2)

Set/indicate priorities on the list (e.g. "essential", "suggested") (3)

C) Rate each of the major sections of the Guidelines in terms of its usefulness in planning or interpreting an evaluation report:

	Very Adequate	Adequate	Inadequate
I) Purpose	18	15	
II) Focus	18	15	
III) Procedures	20	13	1
IV) Information Sources	23	10	
V) Conclusions	23	9	

III. What would you consider the major strengths and weaknesses of the Guidelines in their current form?

Strengths: Comprehensiveness/scope (16)

Clarity (4)

Weaknesses: Length (4)

Format, wording (4)

IV. Can you suggest any additional topics or important areas that should be included in the Guidelines?

None (5)

Individual comments and suggestions (5)

V. Can you suggest any particular section of the Guidelines that should be eliminated, rewritten, or modified substantially?

None (4)

Individual comments and suggestions (6)

THANK YOU VERY MUCH FOR YOUR COMMENTS
Please return the Guidelines and this form to Mary Seifert at IRA headquarters.

APPENDIX F

Field Test Reviewers' Comments on Question 5: "As a result of using these Guidelines, what changes, if any, would you make in future evaluation activities in which you might participate?

I think this would have been exceptionally valuable if we had had it before we began. It could serve as an outline of things to be sure were considered -- even if we excluded them.

Everything in this document should be included in the planning of an evaluation.

If we had an external evaluation, this would be very valuable in evaluating its quality.

I would be sure we recorded our plans for followup before we started; and I'd record specific plans to implement (or refute) the reports' recommendations.

This provided a good base to analyze our reading program evaluation.

I would demand that our report be more specific and look at more aspects of our "program."

I failed to realize the many things we have not been including in our plans to date. These are far more comprehensive than we were being. I think we can do a better job now.

I'd realize better how limited we've been in the past -- definitely conduct a thorough internal evaluation, since our external once leave much to be desired. I guess we sort of knew that before, but this shows us clearly.

I would demand a more systematic, thorough evaluation job by the external contractor -- now I know what to look for in a good job

We would more seriously consider long-term followup and a broader look at the overall reading program and we'd get more imput from all involved groups.

More comprehensive planning and reworking; better dissemination of information; more use in a formative sense to revise ongoing programs.

Excellent planning tool -- I'd use it mostly as we planned the effort so we didn't leave out important elements.

I'd give more attention to equitable input from those involved in preparing the evaluation, those whose program is being evaluated, and those who will implement changes based on the evaluation.

I'd feel comfortable judging the quality of an external evaluation -- or one I didn't work on -- if I could use this document. It provides a focus or outline of what to look for.

APPENDIX G

Interview Comments During Field Test: Recommendations Concerning Local Uses of the Guidelines

I would select portions of the Guideline -- not try to use them all -- Choose the sections most relevant to our study before beginning.

I'd use all but the technical section (Information Sources) as a guide in conducting the evaluation.

We probably don't have the technical-expertise to ensure it, but I'd try to take as much of this into account in our planning as possible.

I would have our evaluation committee screen the set to select the most appropriate questions for us.

I'd select the sections we'd consider to be "high priority" and concentrate on them rather than attempting to use it in its entirety.

I think these are best used as a laundry list that our evaluation committee can use to select a subset of items that fit our needs best. I'd suggest such a step in the Introduction.

How could we get some training on how to use this? If one of our group was trained, he or she could help the rest of us.

I'd mark the key parts and leave the rest as backup -- not irrelevant, perhaps, but less important for our purposes.

Without someone who had a better technical and evaluation background to help us, I doubt that we could make optimum use of the Guidelines.

One set of written Guidelines will never provide all the answers, but this gives us a good checklist.

I'd ask any outside contractor to work through this with us so we could reach agreement on what they'd do for us.