DOCUMENT RESUME

ED 105 443 CS 001 783

TITLE	Continuous Progress in Reading; Proceedings of a Conference of the New England Consortium for the Right to Read (Boston, January 8-10, 1975).
INSTITUTION	New England Consortium for the Right to Read, Kingston, R.I.
PUB DATE	Mar 75
NOTE	32p.
EDRS PRICE	MF-$0.76 HC-$1.95 PLUS POSTAGE
DESCRIPTORS	*Conference Reports; Reading; Reading Development; Reading Improvement; *Reading Instruction; *Reading Programs; Reading Tests
IDENTIFIERS	*Right to Read

ABSTRACT

The seven papers in this booklet include "Continuous Progress in Reading," which discusses managing continuing progress programs, clarifies the rationale underlying continuous progress programs, outlines a developmental checklist of objectives, and offers suggestions for arranging for continuous assessment of progress and providing for personalized teaching; "Criterion Referenced Evaluation Applied to Reading," which discusses a definition of criterion referenced testing, characteristics of criterion referenced tests, advantages and limitations of criterion referenced testing, how to choose published tests, and teacher observation; "Classroom Management of Continuous Progress Reading Programs," which presents sections on Project STRIDE, individualized basal reading programs, individualizing instruction through the use of a learning center, a Right to Read Program, and individualizing instruction in the content areas; and a list of conference participants. (WR)

ED105443

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The New England Consortium
For the Right to Read

Conference Proceedings

CONTINUOUS PROGRESS IN READING

This conference was held at the Copley Plaza Hotel,
Copley Square, Boston, Massachusetts,
January 8 through 10.

March, 1975

CS 001 783

CONSORTIUM LEADERSHIP

Commissioners of Education

Hon. Mark Shedd, Connecticut
Hon. H. Sawin Millett
Hon. Gregory R. Anrig, Massachusetts
Hon. Newell J. Paire, New Hampshire
Hon. Thomas C. Schmidt, Rhode Island, Project Applicant

State Right to Read Directors

Mr. Martin C. Gotowala, Connecticut
Mrs. Laurel Cyr, Maine
Dr. Joseph J. Tremont, Massachusetts
Miss Rosemary Duggan, New Hampshire
Mr. J. Robert Moynihan, Rhode Island

Commissioners' Committee

Dr. Olive S. Niles, Connecticut
Mrs. Mollie Reynolds, Maine
Dr. Natalie Norton, Massachusetts
Mrs. Joanne Baker, New Hampshire
Dr. Marion L. McGuire, Rhode Island

Contract Arrangements

The Curriculum Research and Development Center, Department of Education, University of Rhode Island holds the contract for this project. Fiscal management is under the direction of Dr. Theodore M. Kellogg. Dr. Marion L. McGuire is the coordinator of operations.

Table of Contents

Continuous Progress in Reading

— Dr. Harry Sartain

I was asked to talk about managing a continuous progress program in reading and, while I'm sure that anything I say to you won't be entirely new, because you are so competent and so experienced, I may package it a little differently and it may be of some help.

The management of continuous progress programs requires all kinds of administrative and supervisory functions, obviously, but tonight I'd like to discuss only a few of these. I'll say things that relate to in-service education, to the specification of instructional objectives, to the assessment of learning and to classroom organization. I'll discuss several aspects of these and in discussing them, I'm going to specify four of what I feel are primary phases of the management program: clarification of the rationale underlying continuous progress programs, specification of a developmental checklist of objectives, arranging for continuous assessment of progress and provision of personalized teaching.

Clarify Rationale Underlying Continuous Progress Plans

First, my suggestion in managing continuing progress programs is to clarify for the faculty the reasons for providing a continuous progress plan. A great many faculty members are not at all convinced that highly differentiated, personalized instruction is necessary. They succeeded well in school without it. It's a lot easier to just teach whole roomfuls at the same time, and so forth. One of the things that I usually ask along this line (and you may think it will help if you ask the same thing), "When you were in high school, which you can remember a little better than elementary school, how much of the time that you spent in class was learning time? How much of the time did you learn when you were in classes? Were most of your classes taught as whole classes, without any individualization unless you were very lucky and in a very unusual school; were you learning 100 per cent of the time you were in class? Were you? Seventy-five per cent of the time? Fifty per cent of the time?" All right, I find that most people say that the time they spent in class, in school, probably was less than 50 per cent profitable. The rest of the time was spent waiting for the teacher to come around to us. The teacher would go around questioning and quizzing, trying to develop ideas, struggling with individuals who didn't understand something. After our turn had come and gone we knew that the teacher wasn't going to get back to us for 15 or 20 minutes and we sort of "tuned out."

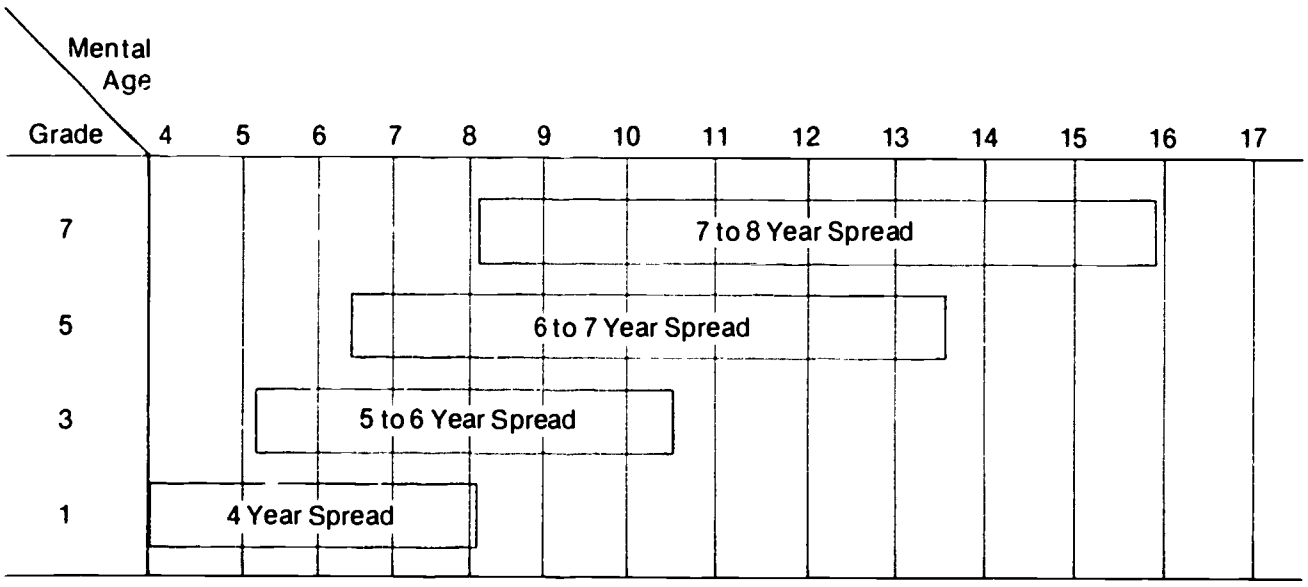
Do you remember all the things that happened during class time? You daydreamed a great deal, didn't you? You daydreamed about social matters, daydreamed about family matters, all sorts of things, and didn't make much use of the class time. Most of us did that. Then the teachers assigned us all kinds of work to take home, so that we couldn't make any use of our free time either . . . all that time we should have had for recreational purposes for music lessons and other things. And this is obviously the fault of the procedures for teaching. Teaching was in no way individualized, in no way personalized, and so the time was wasted. If the teachers had worked with us a few at a time in small groups or in some way to help each of us, each group, each individual in each group, to move at his own rate, then we could have profitably used the whole day because we would have known how to proceed, after our 5 to 10 minutes with the teacher, while the teacher was working with others. And so, for the sake of more efficiency in the time we have for education, it's important to provide an individualized or continuous progress, highly differentiated style of teaching.

Dr Sartain is Professor and Coordinator, Language Communications, University of Pittsburgh.

Another way that we can convince faculty members that this is necessary is through simply charting something like the mental age ranges of children within the various class levels. Of course, we have to recognize first that mental age is a rough indicator of academic achievement — there are all sorts of things wrong with intelligence tests — but still they do give us some picture of the academic ability of the children in the class. They may not be perfectly accurate for any one individual, but they give us an over-all picture You undoubtedly realize that when children enter the first grade they're about six years old chronologically and the range of chronological ages in the class is about 12 months from approximately five-and-a-half to six-and-a-half, depending on the admission date. The range extends from the child who was just too young to get in last year (and this year he's the oldest one in the class) down to the one who just got in this year. So, there's about a twelve-month spread, providing nobody's been retained. What is the mental age range in the class? Stanford-Binet data show us that the normal mental age range in the first grade class is about 4 years. You have children in Grade 1, then, who range mentally from four-year-olds to eight-year-olds. Now it immediately seems illogical — we can see the lack of logic — to try to teach kids who are mentally eight years old, who have the academic aptitude of average 8 year olds, the same thing as those who have the academic aptitude of 4 year olds. And, incidentally, that covers only the middle 96 per cent of the spread. There are still two per cent above that eight-year old mental age and two per cent below that four-year old mental age. About this spread, some teachers will say, "Well, I teach in *this* kind of neighborhood and so I have the more able kids." Or, "I teach in *that* kind of neighborhood, so I have all slower kids." If they'll look at the test results they'll find that generally they have almost the whole range, no matter where they teach. They have a bigger proportion in one area than in another area but they still have the whole range.

Well, what happens to this range of mental ages as children go through school? Does it stay the same? Does it get less? Does it get greater? Because the average child matures at an average rate, his growth is continuous in his mental aptitude, his academic aptitude. The above-average child, though, has the capacity to increase his capability more than a year every year and so he becomes increasingly more able, while the less able child grows less than a year in aptitude each year, so that your spread in mental ages becomes greater, the longer the youngsters are in school, up until they're around 16 or so. By third-grade level you have about a five and one-half year spread in mental age. When you get to the 6th and move into the junior high at 7th, you have between a seven and eight year spread in mental age. Well, if you put this on a chart or a graph using one axis for the mental ages and the other for the equivalent grade levels, teachers can see that when children are sixth or seventh graders, where they have a seven and one-half to eight year spread in mental age, some of these youngsters have the academic aptitude of third graders, while others have the academic aptitude of tenth graders; and it obviously becomes even more ridiculous to expect the same thing of children of all these abilities. This chart may illustrate what I've been saying:

Mental Age Range at Selected Grade Levels



Unfortunately, in most of the schools in the country, the practice has been to differentiate instruction fairly well at the primary levels because there the teachers are faced with the problem that pupils can't read at all. They have to group them in some way for initial instruction. But, then, as they go on to intermediate and secondary levels, they do no grouping or no differentiating — just simply take the easy way out and blame the teacher who had them the year before for not making them all alike. Obviously, with that range in capacity, there is no way in the world to make them all alike.

It was a New Englander, Horace Mann, who introduced the graded school in Quincy, Massachusetts approximately 130 years ago, as you know. Prior to that, instruction was quite differentiated; but he was looking for a cheap way — aren't we always, in education, looking for a cheap way to provide instruction? He was a lawyer, incidentally, not an educator; but he became the Father of American Education because of his interest. He went to Europe to study schools there and found a system that interested him in Prussia. When he returned, he influenced us to adopt a graded school as an economy measure, following the Prussian System. All seven-year-olds took the same curriculum; all eight-year-olds took the same curriculum; all nine-year-olds, and so on. This would also offset the need for much teacher preparation. A teacher could learn the curriculum for one age level and not have to know anything but that one age level, you see. It's contrary to everything we know now about individual differences, and the need for differentiating instruction.

But after the graded school had been in effect for a few years, some teachers learned that age was not a crucial factor in learning ability, and they started groping for ways of differentiating instruction. Homogeneous sectioning became very common and it still is common throughout the country. You are familiar with this — put all the so-called bright students in one room, the middle ones in another room, and slower ones in another room, and so forth. The research on homogeneous sectioning began in 1928 when Hull did his study on this and there have been studies ever since on homogeneous sectioning that show that it doesn't work. But we still have a large proportion of the schools in the country who practice it. We have moved a step further into Joplin-type planning, or cross-class planning, where we change the sections for each subject, knowing that the pupils are not all of the same ability in every subject. We do Joplin-type planning when we put them in one group for reading but in a different arrangement for math and shuffle them between rooms. How many of you are in districts where they have a cross-class or Joplin Plan? It's pretty common, now. It's assumed that you have X capability in reading, but you have Y capability in math, for example; and you re-shuffle among rooms to provide for different capabilities. The research on that, on cross-class planning, as probably most of you know, shows that the first year they get better results than they do with self-contained classrooms. In most studies thereafter, the results level off. They get no better results than with the self-contained classrooms and in some schools they get poorer results with Joplin planning than they do with self-contained classrooms.

Now we all know why we get better results the first year — because no matter what you do experimentally, you always get better results the first year due to the novelty effect. What we know, if we look at the research after more than one year, is that this just isn't doing any good. We're only kidding ourselves when we do homogeneous sectioning, whether it's the old homogeneous sectioning or the modern Joplin Plan sort of thing. A lot of us feel we still need more differentiated instruction.

One of the studies that I've quoted a great many times on the reasons why the Joplin Plan doesn't work, is one done by Irving Balow, who published, I think, in 1962. There are many others, but the one done by Irving Balow, and published in 1962 in the *Elementary School Journal,* is one of the best ones in recent decades. It is one of the most convincing pieces of evidence you will find if you want to point out to faculty members that you have to have highly differentiated or a continuous progress type of instruction.

Balow worked with four fifth grades in a California school to determine whether he could homogeneously section them for the teaching of reading under a Joplin Plan arrangement. He gave eight different reading tests, then averaged each child's reading test scores, and divided them into four sections accordingly. In other words, he said he was dividing them purely on the basis of reading ability. He had reduced the spread in average scores substantially. At the highest level, he had the least reduction because there was more variability among the higher pupils. The test scores in total ranged from about a second-grade equivalent to about a 12½-to-13th-grade equivalent before he did anything at all. When he averaged the tests and then grouped them on the basis of the averages, the top class, Class A, had average scores that ranged from about a grade level five-and-a-half to grade level nine. So he had about a three-and-a-half-year spread in this group instead of a ten-or-more-year spread for the four classes. Then he looked at their scores on the individual tests to see if he had only a three-and-a-half-year spread on the individual tests. Well, you know what happened. On the test of Reading Rate in this class where their reading averages ranged only three-and-one-half years, the individual pupils ranged from grade level two to twelve-and-a-half; because some of the youngsters in that top group on Reading Rate were very low, and of course,

some of them were at the top In reading Comprehension, based on only one reading comprehension test, or only one that was called that, the top group ranged from grade level three-and-a-half to eleven, a lot more than a three-and-one half year spread. In Word Meaning, their range was narrower. In Reading to Follow Directions, their range was almost as wide again as for the whole class.

Well, I could go through the others. Paragraph Comprehension, Sentence Meaning, Alphabetization, Use of the Index, and show the same results. Balow found that if you looked at the pupils individually, you still had almost as much spread in that fast group on their individual skills as you did in the whole group. Now this, of course, is easily explained Whenever we average scores we completely overlook the factor of variability We look at only central tendency, and that does not tell us much about any one youngster. Any one of us can take a number of reading tests, and get a certain average score but that doesn't mean that any one of us can't be quite low in one of these areas and quite high in others to balance them out.

Well, let's look at Class B to see if toward the middle of the distribution we have a more homogeneous class On the reading test averages, there was a range of only one year, in Class B, from about grade level four-and-one-half to five-and-one-half. They were bright-normal fifth graders. What happened when he looked at the individual skills test? Remember, he gave only eight skills tests, and there are so many more that could be given in reading Well, he found that again in Rate of Reading, he had not reduced the range at all In the middle, in this Class B, there were still some pupils who were clearly at the bottom in Reading Rate and some who were at the top in whatever they measured in Reading Rate. In Comprehension he had reduced the total range only a little more than a year In Reading to Follow Directions, he had reduced the total range for the middle group in comparison with the whole class — or all four classes — only about one-half year less than the total range of all four classes, and the same was true on most of the other skills.

So what this proved once again, and it's been done over, and over, and over, and over, is that if you really look at the different skills that we are trying to teach, there is no such thing as a homogeneous class. Our logic shows us that they all cannot be taught the same thing just because they happened to have the same average in reading. Well, if you put the results of this study on a transparency and compared the ranges for the different classes, you can very graphically show teachers that when we think we have a homogeneous class, we're just kidding ourselves. We can have some people who are homogeneous on one thing, like average reading scores, but that tells us nothing about all the other traits. The trait variability still continues to be very great, and that's what we have to deal with in our teaching.

Because of the high variability of traits in youngsters, our efforts to do the logical thing, with various types of homogeneous sectioning so we can teach them all alike, don't work. The research shows that those efforts don't really pay off. Yet there have been, for all these years since Horace Mann set up the first graded school, those that assumed everyone had the same learning capability and approached teaching as though they did.

Specify Developmental Checklist of Objectives

The second suggestion I would make on managing continuous progress plans, after you've convinced the faculty that you really need a continuous progress plan, is to specify your reading instructional objectives developmentally in a checklist form. Now probably all of you have done that already. The Right to Read Program, I think, provides some checklists of skills. But, lets say you have some questions about the adequacy of the lists of skills you are now using. If we are going to have developmental progress, continuous progress, whatever you want to call it, instead of graded progress where you just take one year's curriculum and then move onto the next year's curriculum, you've got to know exactly what you are trying to do in this sequence. There have been many different lists of skills provided, but I've never found one that I felt was adequate. Most checklists have objectives that are too general, or objectives that don't show progression as the pupil's move on. I suggest that you work on a checklist that does both. Let me give an example of some objectives that I feel are not adequate to show progress, an example of the sort of thing you frequently find in publicly available checklists.

"Recognizes common roots words", might be an objective that you will sometimes see in a list of objectives My question is, "Which root words?" A child cannot recognize all the root words at one stage, and so you can't check off that he recognizes common root words; you can only check off that he recognizes specific roots. you see. I believe that the checklist has to show which roots you expect him to recognize at the first stage, which at the second stage, which at the third stage, and so on.

Another example: "Reads social studies material critically." Now I've seen broad objectives such as that on checklists repeatedly. Well, critical or evaluative reading includes at least a half a dozen different qualities that are seldom mentioned on these lists of objectives, such as differentiating fact from fanciful information, differentiating fact from opinion, determining the relevance of material, determining the

completeness of the material, determining the accuracy of it, perceiving propaganda tricks. Those are just a half dozen of the specific evaluative reading skills that we need to teach; and one general objective is not enough. But often, that's the way it's stated.

Therefore, I would suggest that we group our objectives in some way, when we're specifying them — group them in some manner that makes it easier to deal with them, probably either by type or by function. Sometimes I group them by type this way: vocabulary and word-attack skills, comprehension skills; work-study skills in content reading including rate variation, literary appreciation capabilities, discovering language — where, through reading, they discover how sentences are patterned and things like this, language mechanics — where, through reading, they discover the functions of punctuation, the use of the voice, for example, and standard language forms — where they discover different styles of reading material, different dialects expressed, and so forth.

Or we can group our objectives by function: functions of using words, enjoying stories, enjoying poetry, enjoying drama, and enjoying using informative reading, and social interaction type of reading. If we group them that way, then we can specify more precisely the sub-points.

As I mentioned, I would also list the objectives developmentally to show the continuousness of growth. As an example, in listening to evaluate presentations, I have a list of 25 objectives from kindergarten through junior high. (It happens that my list relates to listening which requires almost the same tasks as reading.) At the kindergarten level, one example is: *distinguishes between obviously real and fanciful elements in a presentation, oral or written.* That is the only objective at the nursery school — kindergarten level, but you could have more.

At the early primary levels (i.e., the years approximately equivalent to first and second grades; but again, if it's continuous, you are not too worried about grades), one example would be: *determines if information read or heard is useful for a particular project.* In other words, it determines the relevance of the information.

At the next level, which would be middle elementary, I have about four objectives. One is: *perceives one point of discrepancy in a presentation, oral or written.* So here they are checking accuracy.

At the upper elementary, or early middle school level, I would have in this case seven objectives: One example is: *identifies the speaker's personal motives and the direction of influence on the material;* so we are thinking about the intent of the speaker when we are evaluating. There are others at this level: one relating to fact and opinion; at a higher level, the relevance point; at a higher level, the point on accuracy, the point on completeness, and here's one on adequacy: *states whether or not information heard or read is complete enough for a given purpose;* a more advanced one than where the youngsters were determining whether something was relevant to a project. Now they are determining whether there is enough information, whether it is adequate, complete. There is another point on quality there, too.

Now at the later middle school level, I have about 11 objectives. On the intent of the speaker, a couple of the objectives would be: *identifies the speaker's or writer's biases* and again on the direction of influence of persuasive material: *identifies statements or phrases which illustrate propaganda techniques used by the writer,* such as inducement of fear or discontent, testimonials, half-truths, guilt-transfer by association, satisfaction transfer by association, and the slogan approach. There would be other objectives at a more advanced level on judging fact versus opinion, relevance, accuracy, completeness, adequacy, quality, and so on.

What I'm saying is that, again, too many lists of objectives have the objectives stated too generally so that you can't show growth. You have to have the objectives broken down more specifically in order to show what they've learned at the first stage, at the next stage, the next stage, and the next stage; because, in reality, we never reach the ultimate of having learned the whole list. Maybe we do on some simple things, like word-attack skills. Incidentally, some lists of objectives that I see published have nothing in them but word-attack skills. I think you've seen a lot of those, haven't you? And that certainly is a handicapping sort of thing.

Well, we must have the objectives in enough detail. This one example that I have with me on comprehension tasks, on listening or reading to obtain literal information, I would have, from kindergarten through junior high, 34 objectives; on listening or reading to interpret material, 32 objectives; on listening or reading to evaluate, 25 objectives; which makes 91 objectives on comprehension alone in kindergarten through junior high. Many lists of objectives are not 91-long, covering everything; they are nowhere near that long. In listening for literary appreciation, or reading for literary appreciation, I have 57 objectives; and that doesn't cover everything.

If after we have itemized our objectives so that we know where we are trying to take everybody, then we need to make some provision for them to go individually. When we have itemized our objectives clearly

enough, we can be sure that the children are going to travel at different rates. Having this list of objectives is absolutely essential to continuous progress plans. Their use substantiates the need for individualized instruction. Faculties are convinced that pupils are different when they have to check off what their patterns of learning are, and they can see that different youngsters have learned different things if they assess accurately. One caution: some reading programs have too many objectives listed; and sometimes the objectives are pretty meaningless or repetitive. So, while your objectives must be detailed enough, I would say to avoid needless repetition or needless objectives.

This concern for using objectives to guide teaching and learning reminds me of an unusual want ad. Somebody in our area once read this ad from a Michigan newspaper: "Wanted: Man to work on nuclear fissionable, isotope, molecular reactive counters and three-phase cyclotronic uranium photosynthetizers. No experience necessary." Could we be overly concerned about setting our objectives? Needless to say, we do have a concern about how well our programs function.

Arrange for Continuous Assessment of Progress

The third suggestion I would have is to arrange for continuous analytical assessment of progress. If we are going to record progress on some adequate checklist, we're going to have to have some way of assessing progress.

We have preached continuous assessment forever; they were preaching it when I was in college, and we've all preached it ever since. We are supposed to assess continuously, measure continually; but you know what usually happens, it's an annual achievement testing, a once-a-year sort of assessment with some possibly inadequate, occasional teacher assessments. And, more recently, we have introduced a second assessment wherein we buy some Stanford Diagnostic Tests, or similar instruments, and give them to everybody at the same time of the year; and then, supposedly, we plan our whole year's work on that one testing. Well, goodness, if we're doing any teaching, the testing that we do at one time is completely out of date a month later; and achievement tests aren't that appropriate for measuring objectives, anyway.

I have been surprised at the advice given to teachers by my fellow professors and supervisory workers in public schools — telling teachers to take the achievement tests and study the different items to determine the types of things pupils know and don't know; and then teach those things they don't know. Perhaps we've all been guilty of this at one time or another, but is this a valid use of standardized test data? Standardized achievement tests are norm-referenced tests. They do not have enough items of any kind to make any diagnostic assessment whatever. You've got to have more than one item of a kind to make a diagnostic assessment; and the standardized achievement test doesn't even have a single item of all the different types to cover all the different kinds of reading skills that we're trying to teach. So this business of analyzing the achievement test is wasted effort. Surely there is some information to be gained, but it certainly will in no way give you a measure of how they've grown in all of the objectives that you are trying to help them grow in.

Instead of the norm referenced achievement tests, there is a need for criterion referenced instruments. Now a criterion referenced instrument, as you all know, is one that tests every objective that you've itemized. It's necessary to have instruments that measure all of the objectives that are in your checklist or you can't check off whether the children have made progress or not. You can develop some . . . they don't all have to be published tests. They can be instruments that have been faculty-produced; but if they're tests, if they're observational, if they're directed observational checklists, I would advise having a faculty group work on them rather than individual faculty members. I'm not convinced that most faculty members can do a very good job of diagnosis by observation alone. I've seen too many students who have been checked into reading groups where their needs weren't being met at all to believe that we don't need some sort of help in detailed, analytical assessment. Call it diagnostic assessment, or call it analytical assessment. Some people don't like the term *diagnostic* but I'm not adverse to the term.

Continuous assessment, then, does not mean assessment once or twice a year. It means every day or two, or at least every few weeks. We must organize our instructional program so that the teacher can provide frequent analytical assessment — progress diagnosis for every child. It means that we need to provide inservice education on how to do it, and the tools for accomplishing it. (The tools may be published tests or committee developed.) It means that we must have criterion referenced tools in order to measure all the objectives.

Now, in my experience, many general classroom teachers have used the Spache Diagnostic Scales as a good way to start. There are many objectives that don't get tested with the Spache Diagnostic Scales, but I find that it's something that classroom teachers can use; and they can at least make progress on determining at what levels the children should be getting their general instruction. It does have some word-

attack skills tests There are also some different kinds of comprehension questions asked. If teachers use them well, they can start looking at comprehension differences, word-attack skill differences and rate differences. That's more diagnosis than most have been doing before. Then you'll need to use other published instruments to look at areas of weakness. There are a good number of them on word-attack skills that do provide pretty good coverage, but hardly any on comprehension. The Gates test used to have several subtests of different types of comprehension skills, but they are no longer available. Some of the Gates tests are useful; but again, they're norm referenced instead of criterion referenced.

The Developmental Reading Tests get at some different kinds of comprehension, or at least one kind of comprehension that's a little different from some of the others. Trigg's tests do test, at the secondary level, some specific objectives; but none of them do a very highly detailed job. One of the things, then, you'll need to do in managing continuous progress programs is to examine the instruments that are available, determine which ones really test your objectives at different levels, and then fill in the gaps gradually.

That's a huge order. Some of my recent experiences have made me acutely aware of that. I have been working for two years now with a group of people on listening tests for language testing. You know, there are some diagnostic tests in reading, but there are none in language. We've been trying to develop a series of language tests at different levels that are the analytical assessment type, not just achievement tests. We listed all the objectives first; now we're trying to develop instruments to assess them. And it's a very, very demanding job. The difficulty of writing your own tests shouldn't be underestimated.

In continuous progress teaching, not only the details of achievement need to be assessed; it's just as important that social and emotional growth be assessed, too, because so many learning difficulties are based on children's social and emotional, rather than physiological problems. Vision, as a factor, is pretty obvious; and we've looked at that and other physiological factors rather well in the past. But, there are social and emotional differences, too, because of living in different parts of the city, because of living in homes that have different aims, different values, and so forth, that must not be overlooked. There will be just as wide a range of values among the families as there will be of achievement levels among the pupils. We need to look at the different values of the families carefully if we are going to really personalize our teaching, really support the child where he needs to be supported. Different treatments are required in different situations.

Provide Personalized Teaching

The fourth point that I'd like to make is that we do need to provide personalized teaching. Of course, that's what continuous progress is all about. Some teachers, as you know, prefer to do completely individualized reading; that is, teaching one child at a time. This is something that I've spent quite a little time studying and I think the evidence points out clearly that an effective teacher can do a very good job with totally individualized reading. But, it takes someone who is willing to work 20 hours a day to do the job right. Many teachers who run individualized reading programs, the evidence shows, confer with students only once every two or three weeks. There isn't much teaching done that way. It's a nice, easy way to handle everything. The ones who do a really good job work like beavers.

I think it's more efficient in our utilization of time and effort to do some kinds of grouping within the classroom regardless of whether we have so-called homogeneous sections, which we know aren't homogeneous at all, or whether we have completely heterogeneous sections. I believe that some kind of grouping is preferable to a totally individualized program.

Incidentally, in our campus lab school, where I worked for quite a number of years, we found, as a few other schools have, that the most effective kind of grouping was a multi-aged sectioning. Instead of putting all the eight year olds in one room and all the 12 year olds in another room, and so on, we found it better to mix the ages by at least two years because it increases the range of reading achievement very little to have the seven year olds and the eight year olds together, or the eight year olds and the nine year olds together. That kind of grouping produces only a very slight difference in the range of achievement. But what it does do is to make us, as teachers, aware of the fact that the children are different. It helps us to get over the old graded school idea that because they're all eight-year-olds, they're ready for the same thing. Even in continuous progress plans it's hard to overcome this feeling. Because we've attended graded schools and we've taught in graded schools, it is hard to really think nongraded. But when we mix the age levels, that makes us think nongraded. It also makes it easier in working with the parents to show them we are working with what their youngsters are ready to learn, rather than with grade-level material.

There have been fifteen to twenty different kinds of grouping suggested in the literature, but I continuously recommend about three kinds that cover most problems. The three kinds of grouping that must be carried on concurrently are power, skills refinement and activity grouping.

Power Grouping

Many teachers, especially at the primary level, do a fine job with simple grouping—the simple grouping that I call power grouping, grouping by power or by general ability. I feel that to have a good personalized program, you do need power grouping. But, other kinds of grouping are needed, too.

In setting up power groups, we need more than three. Three is never enough. With the standard three groups, we've had youngsters falling out the bottom of every group, especially the lowest group. Unintentionally, we have left these students behind because there's no way for them to keep up with the groups we've placed them in. Now, the public is aware that about fifteen percent of the youngsters in the country have a literacy problem. It's a problem we must address at the classroom level. In our defense I have to say that 85 percent is a pretty good success rate, a lot better than the success rate in other professions, but it's not good enough for us. The 15 percent failure is something our society can't afford. Each child must be placed where he can succeed.

So, we need to have more power groups in order to really cover the whole range within a class. When I recommend this to inservice groups, I immediately get the cry, "But I'm worked to death trying to work with three different levels. How in the world could I possibly work with five, six or seven?" But I remember I started teaching in that rural school where I had to work with a wide range of levels, so I know it can be done, and I've seen good teachers in excellent schools working regularly with five to seven power groups within a class. One of my answers is that if you have enough different levels, you can introduce new material and new skills much more quickly than when you don't have enough levels. If you have ten or twelve youngsters in a group, there are going to be three or four of them who need so much extra teaching in order to try to learn the vocabulary, in order to try to handle the concepts, and so forth, that you have to spend too much time with that group. That's where a lot of your time goes. By having the children in groups where they are all ready to learn the new material that's presented and the new skills that are presented you can teach much faster. That was my experience in the rural schools. I had to average a class every seven-and-a-half minutes, but that is not as difficult as it sounds if pupils are well grouped.

A great advantage was that pupils did not sit around wasting time because of the graded teaching; they had something to do after I had gotten them working. I see the same thing happening in many classrooms; they have something to do all the rest of the time because of the way they are introduced to their work. We also have to avoid the luxury of sitting ourselves, you know. Many reading consultants have said, "When the group is reading a story, sit with them so you can help them with each word." You can't afford that luxury. You can have them help each other. If they're reading at the right level, and you've done some introductory vocabulary, there is hardly a word that someone in the group won't know. So, we have to introduce stories, get them to reading silently, move on to the next group, and so on, and keep the circus rings all operating at once instead of operating them only one at a time.

Another thing that we can do is to omit some of the lessons that have been suggested in the Teachers' Guides that the group doesn't need. I'm not suggesting that you not use the skill sequence in a guide of some kind, we definitely need it. But, again, some manuals have more reinforcement lessons than your group needs. We often have children doing the same things they did yesterday, the day before and the day or the week before that, just to keep them busy. We can cut out the busy work and save time. I've seen teachers doing that in recent years, and doing it very well.

So, our power groups are for presenting new materials and skills; and the children are grouped approximately for the level that they can handle, in the way we've usually done.

Skills Refinement Grouping

We also need skills refinement grouping. You know, when we do only power grouping, one of the big criticisms is that we develop a type of snobbery on the part of the more rapid ones and a defeatism on the part of the ones who are in the lower power groups all the time, and this is true. There is research that shows that this happens. One doctoral dissertation, for example, showed that if children were second or third to get a book in a basal series, knowing that another group had had it first, their achievement was significantly lower than if they got new materials that had not been used with other children before them. Children who always get the books second or third develop a second-class citizen feeling. That is one very strong argument in favor of having a variety of materials, instead of having a single basal series. Skills refinement grouping will provide instruction for those who didn't master the skills taught in the power groups, and also bring youngsters together on another basis so that they don't become permanently identified with a particular power group. Pupils from all of the power groups, even the top ones, will need skills refinement from time to time. Even the most highly educated adults occasionally have a problem learning something; surely seven-year-olds or ten-year-olds will need to be retaught something at some time. With such a large range of skills as are involved in reading, skills refinement grouping is essential to assure mastery of the skills at each level in a continuous progress plan.

Activity grouping

The third kind of grouping essential to meeting the instructional needs of youngsters is activity grouping. Interesting language arts experiences can be provided that will encourage reading. Elementary youngsters usually enjoy making puppets and putting on puppet shows In a more grown-up way, the junior high school student will find pleasure working with string marionettes. These activities stimulate interest in reading, foster the growth of positive attitudes toward reading, and help teachers provide the kind of personalized teaching I mentioned earlier. Pupils do need time to explore their personal interests through language arts activities.

In summary, before moving in to a continuous progress plan, I suggest you spend sufficient time developing a rationale so that teachers understand why it is necessary to change from whatever organization is presently used to a new and somewhat more complicated organization. Then, you will need to specify objectives clearly and develop or locate tests to measure each of the objectives. Finally, attention needs to be directed toward the classroom management aspect of continuous progress. Teachers need help with grouping youngsters to provide personalized instruction. I have suggested grouping by power or reader level for basic instruction, regrouping for skill refinement in areas of need and activity grouping to stimulate interest in reading.

Criterion Referenced Evaluation Applied to Reading

— Dr. John Pikulski

It seems an appropriate time to revisit the concept of criterion referenced measurement. The exact beginning of the use of the term is somewhat obscured in educational testing history; however, there are definite references to the term in the early 1960's and the term has become educationally very prominent in the 1970's. Its popularity is attested to by the fact that a number of other terms have been used as being synonymous. These include: domain reference testing, edumetric testing, mastery tests, maximum performance tests, content reference tests, and undoubtedly many others. We have reached the point where there is disagreement as to how criterion referenced measures should be defined, what form they should take and what value they have. There has been sufficient time for reading specialists to have become excited, confused and cynical about the use of criterion referenced tests. In short, there appears to have been a sufficient initiation period to allow us to examine the concept and its possible use with some perspective.

Before attempting a definition, it might be useful to expose you to some exercises involving criterion referenced statements and criterion referenced test items.

Consider the following statements. Which are made on the basis of criterion referenced measurement?

1. Marie scored at the 85th percentile in reading vocabulary, at the 88th percentile in speed of reading, and at the 80th percentile in reading comprehension.
2. Nancy can read a list of 20 one-syllable words which contain short vowel sounds with perfect accuracy.
3. John's score on the last reading test was 7.2.
4. Frank can identify 20 per cent of the words on a list containing words which begin with consonant blends of the "l" family.
5. James' overall performance on the reading test placed him within the second stanine.
6. Linda can read a paragraph from a seventh grade social studies book with 90 per cent word recognition accuracy.

Consider the following groups of questions. Can you identify those items taken from criterion referenced tests?

(C.R.) A. _____ puppies fell in the pool.
a. two b. too c. to

(N.R.) B. In building a nest, the mother bird may use twigs, mud, bits of straw, or pieces of string. When the nest is nearly ___1___, she may line it with feathers pulled from her own breast. These ___2___ make the nest a soft home for the baby birds.
1. broken forgotten finished empty missed
2. babies sticks feathers nests cats

Dr. Pikulski is Associate Professor of Education and Supervisor of Diagnostic Services, Reading Study Center, University of Delaware.

(C.R.) C. Read to child: Put your finger on the bell. Next to the bell is a ring. Which picture in this row rhymes with ring? Rope . . . sink . . . swing? Fill in the circle under the picture that rhymes with ring.
Child's copy shows pictures of:

ring	rope	sink	swing

In the spirit of criterion referenced measurement, if you answered all of the items correctly, you should be free to leave for the remainder of this presentation. Therefore, it seemed absolutely necessary to pose some additional pre-test questions.

A. Rocks in the region were shattered; trees were knocked down, uprooted, and thrown high into the air. The largest holes were dug by pieces of iron chunks weighing 30 tons and measuring about 6 feet across. The entire object that struck the earth weighed at least a thousand tons. It was a meteorite large enough to be called an asteroid or a small planet. As it entered the earth's atmosphere, it began to burn up and became a streak of light — a meteor — in the sky.

The best title for this article is:

1. One Morning in Russia
2. Unidentified Flying Objects
3. Very Large Meteorites
4. The Meteorite of 1947

B. Steve and Mary go to camp in the summer. The camp grounds are up in the mountains where tall, green pine trees grow everywhere. A sparkling stream runs between the trees where the campers may swim and fish. The air is clear and fresh and better than the air in the city. Steve and Mary are lucky to go to this camp.

1. The writer of the story thinks the camp grounds are _____.
   dirty pretty too far away
2. The writer thinks that going to this camp is a _____ idea.
   good poor fair
3. The writer thinks that the air in the mountains is _____.
   harmful healthy cold

The above should be challenging and vague enough to eliminate the possibility of anyone answering them. I suspect that some do not, in truth, have reasonable answers. However, being a most unreasonable person, I will take on the task of trying to provide answers to all of the above.

Definition

Although there is not universal agreement about the definition of criterion referenced testing, there are some generally accepted parameters for the concept. The first step that is usually taken in defining it is to contrast it with norm referenced testing and sometimes with diagnostic testing. With norm referenced evaluation the basic question asks how well an individual performs when he is compared with other individuals. Under most circumstances a student's performance is compared with the group of individuals who comprised the standardization population for the test that is being administered — they are the norm. As a result of the test, the examiner can draw comparisons with local, regional or national groups, depending on the population to which the individual is being compared. The focus throughout is upon inter-individual differences. The greater the variability among the individual performances, the better the test is evaluated in most circumstances.

The distinction between diagnostic tests and criterion referenced measures is largely a matter of emphasis, and many writers would prefer not to make a distinction between them. However, some feel that in diagnostic testing the emphasis is upon an evaluation of an individual's strengths and weaknesses in skill areas with attention to the possible causes of problems that exist. The primary purpose is often to distinguish between individuals who do and do not have a learning deficiency. Criterion referenced testing is less concerned with definition of disability and does not emphasize the etiology, i.e., the cause of problems. If criterion referenced testing really operates, there are no student problems, only instructional challenges. The question here is: Is the child learning disabled or teacher disabled?

Prescott introduces another consideration that deserves attention. He maintains that it is more accurate to discuss criterion referenced interpretation of test scores rather than criterion referenced tests. If you will recall the test items presented earlier, you would probably readily agree that it is extremely difficult to distinguish between those that were criterion referenced and those that were norm referenced. A number of writers suggest that a test can be administered and then interpreted either from a criterion or norm

referenced point of view. This position suggests that there are, in fact, no identifying characteristics that separate the two types of test items. Some would hold that the only necessary characteristic of a criterion referenced test is that some standard be established for determining whether or not a particular skill has been achieved. In other words, the score necessary for passing must be set ahead of time without reference to the scores obtained by others taking the test.

Characteristics of Criterion Referenced Tests

Rather than attempting a simple definition it seems more appropriate to try to enumerate several dimensions that some people have felt are necessary — hallmarks of criterion referenced measures.

1. *There must be a clear definition of the task, objective or skill that is to be measured.* This is the most obvious and essential prerequisite for a criterion referenced test. The term "criterion" obviously means a standard and such a test must be one where the item is referenced to or refers to a clearly stated standard. One of the consequences has been that the criterion referenced measures have not attempted to be global measures of reading, but have instead centered on specific skill areas. Most advocates of criterion referenced measures maintain that the objectives must be stated in behavioral terms and tested in a fashion that allows for clear evidence of whether a skill does or does not exist.

Consider the following illustration:

Which item is a clear definition of the task, objective or skill that is to be measured?

1. Has a sight vocabulary.
2. Given a maximum one second exposure per word, the child recognizes preprimer and primer level words from the adopted Dolch sight vocabulary list.

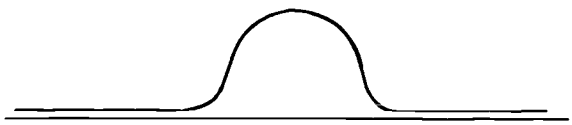
The item is stated in a form that will allow for a determination of whether a reader has or has not the particular skill without student confusion.

This dimension does present problems at times. For example, would the item: The student will be able to read a newspaper, be an acceptable one? Carver (1972) cites this as an illustration of an important piece of information about a student's reading ability in his discussion of criterion referenced measurement. If one were to take a newspaper, select an article and ask a student to read it, would it be a criterion referenced test? In its present form it probably is too much like a "general test of reading." However, practical information such as this can rather readily be transformed into a format that would allow for the designation of criterion referenced. A clearer specification of the source and length of the article along with the definition of an acceptable level of performance would certainly go a long way toward making this a "respectable" criterion referenced test.

2. *Items in a criterion referenced test are chosen for mastery rather than for obtaining a normal distribution.* This determination is one that clearly divides criterion referenced from norm referenced tests. Items are chosen for norm referenced tests so that the results form what is called a normal distribution. This means that writers of test items include many items that are of a moderate level of difficulty, some that are reasonably easy and a few that are very difficult. A test is considered to be a good one when the scores of a group of individuals distribute themselves normally.

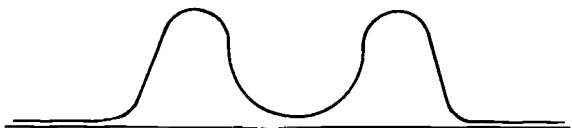
Notice the difference in the distribution of scores for norm referenced (1) and criterion referenced (2 and 3) tests.

1.



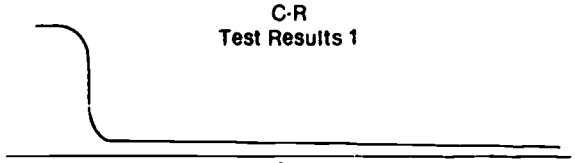
N-R
Test Results

2.



C-R
Test Results 1

3.



C-R
Test Results 2

The anchor point for a norm referenced test is the middle — the average level of performance for a particular group of individuals.

The situation is quite different in criterion referenced testing. Rather than looking for the normal distribution with the anchor in the middle, criterion referenced tests are anchored at the extremes. At one end are those individuals who have mastered the skill under consideration; at the other extreme are those individuals in whom the ability is absent.

A criterion referenced test would be considered a poor one if there were many individuals who had some of the items right and some of them wrong. Remember presence or absence of the skill is what the test is trying to determine.

In norm referenced measurement the goal for choosing items is to discriminate among and compare individuals; in criterion referenced measurement the goal of each item is the same — determining whether a skill has or has not been mastered. Items are chosen because they are considered to be representative of essential skills needed by students. The goal of the teacher with regard to such a test is to have all students respond to all items with perfect or near perfect accuracy.

There is fairly good agreement on the part of test constructors that items for norm referenced and criterion referenced tests are chosen by very different criteria.

3. *The items should be representative of skills that are essential to learning to read.* This is also a very vital characteristic of a useful criterion referenced device in reading. One question that is always raised with regard to a test is: Is it valid? A test is considered to be valid if it measures what it says that it is measuring. The questions related to validity could involve many hours of discussion. Let me simply say that the primary validity question with regard to criterion referenced tests is the question of content validity. Content validity asks if the test items are an adequate sampling of the thing that they are supposed to be measuring. All testing is simply sampling from a great body of behaviors and a test constructor must always face the question of whether the sample that he proposes for a test is a fair sample.

A major problem is that there is far from perfect agreement as to what is and is not an essential skill for reading. Since in criterion referenced testing the goal is mastery of a skill by all students, it is particularly wasteful in time and effort for a group of students to engage in achieving a minute goal that is not essential to some larger, practical goal.

Consider these items:

Objective: Given a list of words, the pupil identifies the three sounds of "ea" with 95 per cent accuracy.
Test Item: Choose the answer that tells how the letters "ea" sound in each word.

1. death
   a. long a
   b. long e
   c. short e

Objective: Given specified two-syllable words and counterparts divided into syllables, the pupil classifies words by syllable stress with 95 per cent accuracy.
Test Item: Read each list of words. Locate the word that has the accent on the wrong syllable.

1. a. pe'ri od
   b. stran ger'
   c. sub scribe'
2. a. ex'act ly
   b. law'yer
   c. hel'met

4. *The items should be arranged in an established hierarchy or should follow the sequence decided upon for reading instruction.* Prescott (1971) is very firm regarding the first half of this characteristic of criterion referenced testing. He writes, "Basically, the criterion-referenced approach is of little value unless the assumption is made that mastery of one skill or bit of information is essential for mastery of another skill or bit of knowledge of a somewhat similar character at a higher level of difficulty or complexity" (p. 352). He goes on to acknowledge that there is no well accepted hierarchy of reading skills. Different programs follow different sequences. Many of the skills are not hierarchical in the sense of depending on lower-level skills. Some children may use context clues effectively before demonstrating a knowledge of phonic skills or vice versa. One could therefore conclude that using criterion referenced testing at this stage of knowledge regarding the reading process is inappropriate. In truth if we insisted on a hierarchy, we could throw criterion referenced testing out. The

more reasonable approach, however, seems to be to arrange the criterion referenced testing of skills in the order in which they appear in a program by which a child is being instructed. This suggests that criterion referenced testing operates most efficiently when considered in relationship to some program. The program need not be a published one. It can be totally teacher devised, but it does require that the teacher designing the program have some suggested sequence of skill development in mind.

Skills such as the following may be in a teacher-established sequence:

Place Relationships: Choose the phrase for a picture.
Synonyms: Identify the synonym for a word in a sentence.
Multi-Meaning Words: Identify the meaning intended in a sentence.
Pronoun Referents: Identify the person or object referred to by a pronoun.
Details, Paragraphs: Identify details.
Main Idea: Choose a title for a story.
Draw Conclusions from Clues: Identify a character or an object.

If skills are sequentially arranged, there is much greater potential for determining necessary steps to overcome problems.

5. *The tasks to be accomplished must be evaluated in the context of a normal reading situation.* Johnson & Kress (1971) challenge the position that most tests that are norm referenced can be used for criterion referenced interpretation for a different reason then others have presented. They maintain that a multiple choice format is contradictory to the basic purpose for criterion referenced measures, which they maintain is finding out if the student can accomplish the reading task.

Consider this item:

Objective: Given a list of words, the pupil will recognize silent consonants with 95 per cent accuracy.

Test Item: Choose the word that has the silent consonant.

1. a. gnat
   b. gallant
   c. given
2. a. grind
   b. gnarl
   c. glory

It is not possible to determine if a reader could have figured out these words in context by administering this test item.

Returning to the illustration cited above (Determining the sound of "ea" in death), these authors would most likely prefer to have the student *read* the word *death* as a way of determining his knowledge of this pronunciation of the "ea"digraph rather than using the multiple choice format shown earlier. Critics of this position would probably raise the objection that the student might be able to read the word "death" as a sight word and not know that the "ea" digraph has a short *e* sound. Quite likely those in harmony with the Johnson & Kress position would ask if he really needs to know that "ea" in death has a short *e* sound, so long as he can read the word correctly. Bleismer (1972) in advocating the greater use of more "natural" evaluation techniques raises the question of whether it isn't possible for a child to be able to accurately syllabicate a list of thirty words and yet not be able to pronounce any of them. He asks if it wouldn't be a more valid procedure to observe what the student does when he meets an unknown polysyllabic word in his reading material.

As you consider the approach you might take to criterion referenced testing, consider the five characteristics presented above or, at least, those that seem important to you.

Advantages of Criterion Referenced Testing

The advantages of criterion referenced testing are:

1. They are relatively free from one of the frequent criticism of tests, i.e. that they are given without reason. This is especially true if you have clearly stated goals.
2. Cultural bias is of less significance.
3. They tend to avoid labeling and categorizing.
4. They are a better gauge of progress because they are given more frequently, they measure shorter steps and they instill a sense of accomplishment — a feeling that certain things are done.
5. It is possible to change, add or delete items without invalidating the test.
6. They are closely tied to instruction and have immediate implications for teaching.
7. They force us to examine our goals. If we have none, they force us to set some.

Limitations of Criterion Referenced Testing

Criterion referenced tests do have some limitations. They are listed here to indicate things you should watch out for in writing or selecting tests for your own use.

1. Some criterion referenced tests include unnecessary items.
2. There is no established hierarchy of skills. This may be due to an inadequate theory of comorehension.
3. There is a question of reliability for some individuals.
4. The specificity of the items is important. If they are too general, the test is not useful; yet, if they are too specific, it is very destructive to the process of reading.
5. Reading cannot be broken down into a list of hundreds of skills. Criterion referenced tests tend to emphasize that which is measurable and overlook aspects of reading that are more difficult to measure.

Published Tests: What's Their Place?

Published tests are a helpful but small portion of the reading program. One program describes itself as "limited to the skill development aspect of reading . . . we have made no attempt to describe a total instructional program in reading, the assumption being that viable reading programs are best worked out at the local level . . . our assumption is that skill development is best fascilitated when teachers accept the responsibility for directing learning experiences which suit their pupils' characteristics and needs."

And again: "Teachers who know their pupils are in the best possible position to guide instruction. For this reason (name of program) offers neither a definitive reading program nor specific instructional prescriptions for any child or school; it offers possibilities . . . Teachers, not the system or the materials, do the teaching."

A position such as this program takes is realistic. Don't try to kid yourself. Bad teaching cannot be corrected through a set of criterion referenced tests. Goals and the scope and sequence of skills must be devised or at least understood by teachers. A program can be good only if the teachers understand it and implement it.

Published Tests: How to Choose Them?

Here are some criteria you might use in evaluating published tests:

1. Claims are modest.
2. Acknowledgement is given to the primacy of instruction by first presenting a minimal number of goals and objectives with a manageable evaluation framework.
3. The sequence closely approximates the preferred program of reading or sequence of skill development. (If there is a commitment to a published program and that program has accompanying criterion referenced tests, these should receive prior
4. Unnecessary skills testing is avoided.
5. Skills are tested in as natural yet efficient means as possible.
6. An efficient record keeping system is provided.
7. Acknowledgement is given to the importance of goals that cannot be measured in a conventional way. For example:
   a. Enjoys reading
   b. Reads independently
   c. Enjoys sharing reading experiences with others
   d. Responds to the mood evoked by a book
   e. Empathizes with characters
   f. Appreciates character description
   g. Appreciates the humor in materials

Published Tests: What Do You Do With Them?

1. If they're part of an instructional package you bought, there's no problem. Follow the directions and *then modify. Remember* you can change, omit and add to the tests, as necessary.
2. If you feel that the goals and philosophy are suitable for your education situation, you can organize reading instruction around the management system suggested by the criterion referenced testing program. But don't do it until you visit someplace or places where they're using this program to see how it works.

3. You could read and study the tests and then write your own to fit your own objectives. An option many schools have chosen is to set objectives, write criterion referenced tests, and then tear up available materials to fit the tests and objectives. It is a tremendous undertaking to establish a management system. Don't start anew.

Teacher Observation and Criterion Referenced Testing

Some teachers are poor diagnosticians. But, I would maintain that, in many cases, those teachers who are poor diagnosticians are poor teachers.

Those who are poor diagnosticians but good teachers can be helped and usually want to be helped to become good diagnosticians. I'd be happy to work with these teachers. The difficulty is more with individualizing instruction to provide the skill development needed.

The term *diagnostic teaching* is hackneyed, overdone, old fashioned and yet good. It's basic. We need more of it. But again, emphasis must first be upon *teaching*. If you've got a weak teacher, forget about criterion referenced testing. *Be realistic*. It's great to talk about criterion referenced testing, continuous progress, individualized instruction, and skills groups, but where I come from there are still teachers who can't handle 3 groups in a traditional basal program.

Use some diagnosis *yourself*, and be sure that the teacher is doing a good job.

But as indicated before, there are teachers who simply aren't sensitized to diagnostic teaching who could easily be. Let's consider a very simple situation, a group of 12 youngsters in a basal reader.

1. Don't preteach. That *often* wastes a lot of time. Many students know the controlled vocabulary.
2. Know what the objectives for the lesson are. Get rid of those who know what you are teaching and keep a record of that.
3. Observe them reading, talk to them, and test them as you move around. (e.g., four new compound words)
4. Present skill activity instruction.
5. Make note of those still not achieving mastery. Record keeping is important.
6. Use criterion referenced testing in its most significant, practical and efficient form. Here you'd follow the goals of the program.

To review the important characteristics of a criterion referenced test, the following points are worthy of your consideration:

1. The criteria for success and failure are set ahead of time. You know in advance what score is needed for mastery of the objective. As mentioned earlier, some hold this as the only necessary characteristic of a criterion referenced test.
2. Clear objectives are set. The questions on the test will parallel the questions the teacher uses in class. When the objectives are clear and instruction improves, diagnosis will improve as well.
3. Criterion referenced tests are constructed to determine mastery of specific skills rather than to show a normal distribution on a variety of items.
4. Test items should cover only the essential skills.
5. A hierarchy or sequence of skills should be decided upon so that tests can be administered a few at a time as instruction progresses.
6. Testing should be done as much as possible in the context of normal reading.

In conclusion, the business of education and learning to read are too serious to leave to chance. We've got to set goals for reading and help children achieve them through individualized continuous progress education. To do otherwise is to deny them that right you're all so very concerned about — *their right to read*.

Classroom Management of Continuous Progress Reading Programs

Section A: Project STRIDE

— Mrs. Catherine Brine

Mrs. Catherine Brine, a third grade teacher at St. Colman's Public School in Brockton, Massachusetts, outlined Brockton's own reading management system: S.T.R.I.D.E. Project S.T.R.I.D.E. is basically a diagnostic/prescriptive approach for individualized instruction. Many materials, keyed to specific objectives, are used.

The program, authored by the city's reading department and a representation of classroom teachers, is a child centered one which holds two major goals: to bring children to the point where they are responsible for their own learning and to teach children how to learn. In other words, S.T.R.I.D.E. tries to develop self-directed learners.

Through a well-developed in-service program, S.T.R.I.D.E. teachers were trained in their new additional roles of constant diagnosticians and prescribers for each child's individual needs.

The components of Brockton's individualized reading system include:

1. 600 Performance Objectives covering the areas of Pre-Reading, Word Identification, Word Meaning, Comprehension, Study Skills, and Literary Analysis. Each objective is cross-referenced to specific learning materials (ie. workbook pages, filmstrips, tapes, overhead transparencies, and games). Resulting are Prescription Sheets (contracts) for each of the 600 objectives.
2. Diagnostic Instruments to assess specific needs. Standardized tests are given in September. Books of Pre-tests and Post-tests are available, containing a pre and post-test for *each* contract (for each performance objective).
3. Alternative learning resources are listed on each Prescription sheet (contract) to accommodate the different learning styles of individual students.
4. Classroom management system which includes:
    a. Prescription Sheets (contracts) which, after an introduction by the teacher, guide students through the listed alternative learning resources.
    b. Individual student record sheets to record individual progress.
    c. Class record chart which gives an overview of the Prescription Sheet on which each child in the class is working.

In sum, Project S.T.R.I.D.E. is not only a written program, but also a way of thinking about teaching. It provides concrete learning activities and a philosophy for providing a self-directed learning environment in the classroom.

Section B: An Individualized Basal Program

— Mrs. Lynn Lanier

Mrs. Lynn Lanier from the Annie Fisher School in Hartford, Connecticut, presented a reading program at the fourth-grade level which is an individualized basal program monitored by the Croft word attack and comprehension materials.

The Ginn 360 basal series is used as the core of the program. Pupils are given the Croft criterion referenced diagnostic tests and are grouped according to their instructional reading level for guided reading and discussion of the stories in the reader. Questions for the stories are teacher-made and based on the Croft comprehension question strategies. The skills development part of the program is organized to coordinate with the objectives of the Croft word attack program. Diagnostic testing is done and the results are used to form ad hoc groups for work on a specific skill listed in the Croft objectives. Ginn 360 workbook and skills-book exercises are interwoven through the program to provide reinforcement on the objectives.

In order to individualize as much as possible, Mrs. Lanier developed a unit package that is given to the pupil to guide him/her through his/her prescribed program. As each pupil, or group of pupils, is ready to begin a unit, she gives each child the unit package. She and the children look over the unit directions guide and their readers. Then she introduces the first story. After the introduction, the children follow the unit directions guide of their unit package. An example of a set of directions for one story in the unit follows:

I. "Space Monkey"
   - A. listen to story word tape
   - B. read the story
   - C. answer story questions
   - D. skills-book 50, 51, 52, 53
   - E. workbook 34, 35, 36, 37, 38
   - F. read the poem "Three Skies"

Usually, there are several stories in the unit. Each story has a set of directions and the child receives all of them at the beginning of the unit. Also, at the bottom of the unit directions, the test for the unit is given so that the child knows what he is moving toward.

The other part of the unit package consists of the new words and story questions for each of the stories in the unit. See the example below:

Level 9 — Unit 4

"Space Monkey"

New Words

hours	minutes	capsule	couch	stroked
medal	instruments	passenger	ribbon	Cape Canaveral
tender	searchlight	aiming	delighted	retired
skyward				

Story Questions

1. Why did the American scientists want to send an animal into space?
2. Why did they choose a monkey?
3. What kind of man did Mr. West seem to be? How do you know?
4. Why did the monkeys have instruments placed under their skin?

5. In the story what did T.L.C. stand for?
6. What was the capsule lined with?
7. What time of day was the flight?
8. After everything was over, what seemed to make Miss Baker the happiest?

The pupil looks at the new words for the story and listens to the story word tape to hear them pronounced. Next he reads the story and answers the story questions. The story questions are based on the literal, interpretive, analytic, critical, and creative categories of the Croft Comprehension Model.

After answering the questions, the pupil continues on doing the assigned skills book and workbook activities.

He/she also meets with an *ad hoc* skill group to work on the objectives indicated by his/her diagnostic test results. If the pupil needs still more help, a pass (see example below) is filled out indicating the specific objective with which he/she needs help and the teacher arranges a time for him/her to meet with the reading specialist in the reading resource room.

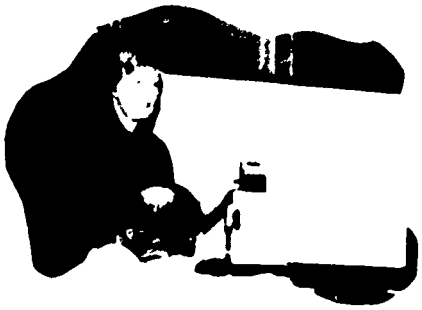
PRESCRIPTION

Name _____ Room _____

Program _____

Unit _____

Objective _____



Section C: Individualizing Instruction Through the Use of a Learning Center

— Mrs. Evelyn Lerman

The Learning Center described here serviced individual children at each level of instruction in a K-8 school. The rationale was to individualize learning for remediation, reinforcement, enrichment and acceleration and to maximize the use of materials housed there. This Center provided materials in a central area where children could be sent by teachers to be serviced by aides. The materials had to be self correcting, multi-media, and span all the levels needed. They also had to encompass a wide range of skills in the language arts.

The Team which set up the Center consisted of teachers, student teachers, and the director of language arts. Their job was to delineate the skills, establish the levels, code the materials and produce a manual which teachers could use to determine which materials to prescribe for their children. A comprehension skills outline was agreed upon along with a decoding skills outline. These consisted of a logical sequential development of skills in which each skill became the foundation for the next. A segment of the skills outline is shown here:

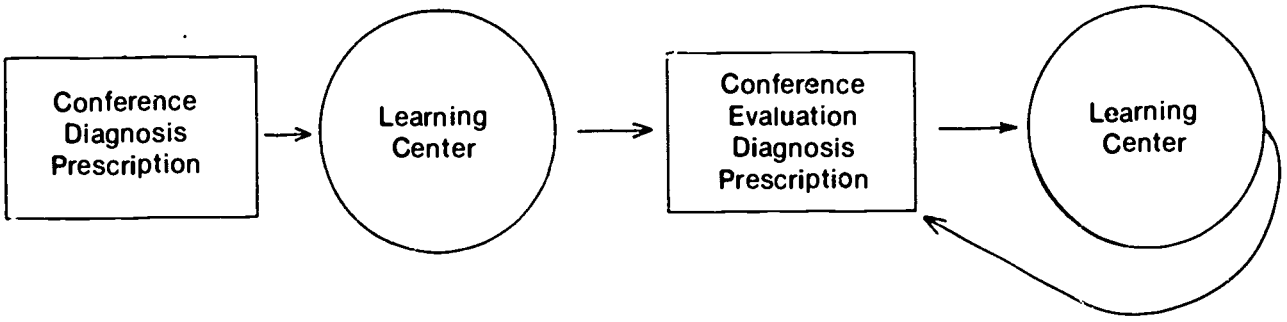
Comprehension Skills Outline

1. Understanding Facts and Concepts (Seeing and Knowing)
   - a. Finding context clues (syntax)
   - b. Using word signals (lexical)
   - c. Recognition of parts of a book
     - Title
     - Title page
     - Copyright
     - Acknowledgements
     - Bibliography
     - Lists of maps, illustrations
     - Table of contents
     - Headings and subheads
     - Captions
     - Illustrations
2. Application (Abstract and Concrete)
   - a. Drawing conclusions
   - b. Preface
   - c. Footnotes
   - d. Diagrams
   - e. Glossary
3. Analysis (Breakdown)
   - a. Making comparisons
   - b. Solving problems
   - c. Judging relevancy
   - d. Noting significance
   - e. Recognition of fact and opinion
   - f. Finding key sentences
   - g. Outlining
   - i. Charts and tables
   - j. Time lines

4. Synthesis (put together)
   a. Making inferences
   b. Finding proof
   c. Determining sequence
   d. Summarizing
   e. Paraphrasing
   f. Classifying (objective)
   g. Alphabetical order
   h. Graphs
   i. Index
   j. Distinguishing fact from opinion
5. Evaluation (Judgment)
   a. Judging validity
   b. Verifying accuracy
   c. Classifying (subjective)

An Inventory of Code Charts was an integral part of the manual so that teachers could be aware of the materials available, and the manual itself was set up so that teachers could look up skill, level, and material needed to fill a need.

At a conference between the student and teacher the teacher determined the prescription, filled out a Learning Center Admission slip, and sent the child to the Center. Once there, an aide helped the child by finding the material, and the child worked independently. Upon his return to the classroom, the conference cycle continued as the teacher held a conference with the child on his progress and recorded the work on the Learning Center Progress Card. The Conference cycle might be graphically represented as follows:



The Progress Card contains information, as shown below:

LEARNING CENTER-PROGRESS CARD-LANGUAGE ARTS NAME: _____ _____

CLASS: _____

DIAGNOSTIC AND PRESCRIPTIVE DATA EVALUATION

Date	Skill Needed Level	Materials— Code No.	Page or Color	Complete or Incomplete	Level Mastered	Comments

Training sessions were required for children, teachers, student teachers, and aides, and these needed to be continued throughout the year as new materials came into the Center. The key to the operation seemed to be a sense of ease on the part of the teachers and students using the manual and the Center. The Center Coordinator, a paraprofessional, found that once teachers started sending their children and were satisfied with the work they were doing there, they continued to send children throughout the year. To assist with record keeping an easy-to-use admission slip was devised, as shown:

Learning Center Admission Slip

Student's Name _____

Class _____ Date _____

Materials Prescribed: (Code and pages)

Time of departure from room _____

Time of arrival at Center _____

Time of departure from Center _____

Aide's initials _____

The initial concern that children would consider it a place for remedial work was obviated by the tactic of sending children to the Center for enrichment and acceleration as well as reinforcement before sending children for remedial activities. This built up the feeling among the children that the Center was for all levels of ability. It was viewed as an interesting place to do individual assignments.

Section D: A Right-to-Read-School Program

— Mrs. Diane McNamara and Mrs. June Lawton

The R_2R program of Coggeshall School, Newport, Rhode Island uses Scott Foresman Reading Systems as the basal program in conjunction with the Croft Diagnostic Tests. The Croft materials are used to monitor student progress and aid in prescriptive planning. Besides using these two tools as the core of the reading program, the teachers rely heavily upon audio visual equipment and teacher-made instructional materials.

Many of the teacher-made materials were developed and produced in CAM workshops. Project CAM is a Title III project based in Portsmouth, Rhode Island. Its purpose is to help teachers develop and make materials to meet the specific skill needs of children. The CAM process is as follows:

1. Teachers identify a specific skill need of the child.
2. Materials are made to match the skill objective.
3. Instruction or reinforcement takes place using the material.
4. Testing takes place to see if the child has acquired the skill and met the objective.
5. The results for the child and the appropriateness of the materials are evaluated.
6. The process begins over again with the proper adjustments made, if any.

The use of these materials enables the classroom teacher to set up listening and learning centers for certain groups of children while freeing the teacher to work with individuals or small groups that may need reteaching of a particular skill. A wide variety of tapes, filmstrips, records, transparencies, puzzles, and games can be used to reinforce each of the skills. A varied approach to learning is used in order to reach *all* of the children, particularly those with specific learning disabilities.

Ad hoc groups of children are formed according to pupil weaknesses in particular skills. Each child's needs are determined by using the Croft Diagnostic Tests and groups are formed on the basis of results recorded on the class record charts.

To meet the children's needs, teachers have set a central resource file that contains worksheets, visuals and games, all labeled and filed according to the hierarchy of skills specified in the Croft program. The resource file is invaluable to teachers in their efforts to provide individualized practice to all pupils to assure skills mastery.

This system allows the teacher to find the child's immediate need, set up *ad hoc* groups, locate the necessary materials easily, and finally, determine by post-testing whether or not the child has fully grasped that concept or skill.

All of these materials and methods are used by the teacher who receives support and help from the paid aides and volunteers who are an integral part of the R_2R program.

Section E: Individualizing Instruction In The Content Areas

— Ms. Leslie Anderson
Mrs. Diane Bleakney
Mr. Jeff Day

Miss Leslie Anderson, Mrs. Bleakney and Mr. Day from the Right to Read School- based site in Lewiston, Maine, presented materials on grouping and individualizing in the content areas. The materials were developed in a summer staff development workshop. The thrust of the workshop was to provide opportunities for Lewiston Junior High School teachers to develop model learning experiences for students in group situations. A wide variety of instructional units in all subject areas were created. Miss Anderson said that "even more significant was the realization that instructional grouping was indeed a feasible and valuable way to meet the needs of secondary students in every content area."

The following rationale, extracted from the works of authorities in this field, was adopted as a basis for grouping and individualizing in the content areas: "Teachers should be more than information dispensers. Rather, they should set the environment to help students develop the means for acquiring, interpreting, and using knowledge independently." (Harold Herber) "The teacher's responsibility in a grouping situation is to maintain a flexible approach and allow freedom for both pupil and teacher to follow problems. Such an approach values a pupil's ability to think rather than emphasizing the content of what he is supposed to know. Research showed (Revans, 1965) that the effectiveness of the teacher was related to the ability of the students to ask questions and understand how to search for answers." (J. Foster)

Various possible procedures were suggested to aid the teacher in selecting a particular way of grouping appropriate to both the rationale and the specific objective of a classroom activity. Their guidebook includes the following management procedures and grouping criteria:

Classroom Management

Possible Grouping Procedures

1. Thematic/Problematical
   a. Class-wide orientation (concept development) activities
   b. Small group activities based on established criteria
   c. Whole-class sharing of interests/learnings through small group presentations
   d. Personalized activities identified during participation in small groups
2. Class Project
   a. Class orientation according to demonstrated *student* interest in a project
   b. Small group activities to complete specific segments of class project
   c. Small group presentations to whole class
   d. Completed project shared with other classes
3. Interest Group
   a. Small group identifies and pursues project based on shared interests and/or needs
   b. Demonstration activities developed for sharing with entire class if desired
4. Skill Group
   a. Teacher identifies and groups students needing help in a particular skill (reading, study, or mechanically related)
   b. Teacher assists group until skill is sufficiently mastered. Group then disbands

Summary of Grouping Criteria

Type	*Sources/Materials*	*Participants**
Random	Multi-Level Books, Media sources, Comprehension guides	Heterogeneous
Achievement	Multi-Level books, Comprehension guides	Homogeneous/ Heterogeneous
Skill	Instructional Materials for specific skills	Homogeneous
Interest (similar or diverse)	Multi-Level books, supplementary material of student choice	Heterogeneous
Social (similar or diverse backgrounds for purposes of exploring concepts)	Multi-Level supplementary materials	Heterogeneous

ALL GROUPS ARE TERMINAL

*Participants are grouped by reader level and ability.

Record keeping procedures can be found in Foster's book. (See bibliography)

Guides were developed for planning an individualized approach, and for implementing intraclass grouping in content areas. They are written in a question format to both guide and stimulate the teacher to develop strategies that take into account student needs, teacher needs, teaching style, and curriculum demands. The Preparation Guide is excerpted below:

Preparation Guide

Preparing to Group and Individualize in Content Areas

1. Pre-Evaluation
   a. What are the needs of my student population in terms of achievement and motivation, interest, and talent?
   b. How would I evaluate my own needs, abilities, attitudes, and teaching methods?
2. What goals and objectives do I wish to set for this instructional program?
3. What activities can I use to establish concept relationships, promote background knowledge, and develop attendant skills?
4. Record keeping — how will I keep track of student progress?
5. Evaluation — how will student progress be evaluated?
6. What methods of student and teacher assessment will be used? How will I determine the extent to which my program objectives have been achieved?

After answering the questions on the Preparation Guide, the teacher will have much information on what she wants to do, why she wants to do it, what she might use to do it, and how she might evaluate whether or not she has accomplished it. The next step in the process of grouping and individualizing is to organize the information to formulate a plan of action — the "instructional unit plan." The "Implementation Guide" (see below) is used to aid in this task. Examples of each step, taken from a language arts unit, have been provided.

Implementation Guide

Steps for Implementing Intraclass Grouping in Content Areas

Step 1. Concept Identification
   a. What ideas and concepts do you hope the students will discover?

Step 2. Rationale
   a. What is my rationale for presenting these concepts?
   b. How will it be of value to my students?

Example (Language Arts)

Topic: Characterization and Setting in the Short Story

Introduction: The grouping situation described by Herber seems highly appropriate in the development of a language arts unit on the short story. Using the elements of characterization and setting as examples, this project attempts to show how the unit would be planned, taught, and evaluated.

Characterization: In stories, as in life, there are many different kinds of people interacting in many interesting ways. (Identified concept)

Step 3. Resources and Materials

a. What do I need to meet the individual needs of my students?
b. What materials do I have in my class? What is available from other sources?

Collections and anthologies of short stories, (all-class example: Poe, "The Telltale Heart") a video-tape system, a tape recorder, files and pamphlets on television techniques, tape recordings of previous examples of TV shows and radio plays are some examples.

Step 4. Concept Development

a. What background knowledge do my students need to understand these concepts?
b. What is my method of developing these concepts and helping students to draw relationships among ideas?
c. Do I need to take any further concept development steps before grouping?
d. Have I helped my students to set purposes for the reading of related materials?

Example (Concept Development)

1. Method: Word association, class reading, discussion, teacher lecture on definition and terms, group project choices.
2. Purposes for reading related materials: Students will attempt to locate in subsequent stories the various kinds of people mentioned in word association and exemplified in reading, discussion, and projects.

Concept development activities:

1. Word Association
   Students, in groups, devised lists titled "Types of People" (90 seconds)
   Lists compiled as follows:

Good	Bad	Misc.

Students compare lists, compile, and discuss

2. Class work:
   Story reading by teacher, discussion, teacher lecture, definition of terms, class questions.

Step 5. Follow Up Activities

a. What group and individual activities might interest my students?
b. How can I assess the interest of my students?

Example (Student/Group Activities)

Directions: Select one of the following as a group project on characterization.

1. Based on your reading, produce a horror TV show being sure to base some of your story on Poe's example, especially your characters. Build your set and plan your production.
2. It isn't really necessary to see "The Telltale Heart." Use your imagination and create with the tape recorder a radio play of the story. Be sure to stimulate our imaginations by using many sound effects.
3. Write the script for #1. Use elements of characterization learned from reading of Poe. Invent your own character types. Remember, the best characters come from real life.

4. Write some character sketches and give them to another group. That group must then act out the characters described and the class will guess who it is.
5. Stereotypes are easily recognized characters. Some are the bossy type, the quiet type, and others. Write descriptions of these character types and present them to the class.
6. Impersonation is acting like someone else. They are live characters, usually famous people. Write down the easily recognizable character traits of five famous people and then have someone impersonate them in class.
7. Create your own project. Check with your teacher to make sure that it is acceptable.

Step 6. Grouping

a. What is the best method of grouping in this classroom situation?

Other Considerations

Record Keeping

Groups will keep all materials in a folder. Chairman and each person will keep daily accomplishment record. Teacher conferences are the end result.

Method of Evaluation — Based on:

1. Creativity
2. Following Directions
3. Class rules
4. Technical skills
5. End result of project
6. Provision for positive reinforcement

Students will evaluate project at the end. End results will determine the extent to which the program objective is achieved.

Time Schedule

Six class periods in two weeks
Teacher will divide time, after initial lecture, to visitation during each period with groups.

Period I — Word Association, Reading of Story, Discussion, Group Choices
Period II — Group Work, Teacher Circulates
Period III — Group Presentations, Evaluations

Contingency Plans

Other short stories would be available to students.

Bibliography

Foster, John. *Recording Individual Progress.* London: Macmillan. Education Limited, 1971.

Herber, Harold. *Teaching Reading in Content Areas.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1970.

Conference Participants

RIGHT TO READ OFFICE, USOE
Mr. Thomas Keyes, *Project Officer*

Consortium Committee

Dr. Olive S. Niles
Mr. Martin C. Gotowala
Mrs. Laurel Cyr
Dr. Joseph J. Tremont
Dr. Natalie Norton
Mrs. Joanne Baker
Miss Rosemary Duggan
Mr. J. Robert Moynihan
Dr. Marion L. McGuire

Connecticut

Marsha Hilsenrad
Irving Ouellette
William Hracyk
Joan Avitabile
Mary Brewer
Marcella Vreeland
Joan Chernauskas
Roberta Roy
William Baker
Martita Wezowic
Joanna Rosenberg
Gertrude Morrison
Sister Mary Kelly

Maine

Sister Jacqueline Ayotte
Elizabeth Bois
Gloria Ferland
Jean Perkins
Nancy White
Caroline Sturtevant
Barbara Moody
Rena Grey
Beatrice Cobb
Phyllis Small
Sister Rita Martin
Roger Grondin
Rachel Stevens
Leslie Anderson

Massachusetts

Marilyn Power
Mitchell Afrow
Eleanor Magane
Anita Gerrein
Carlene Flannery
Elsie Miller
Polly Furbush
Mary Lou Fishman
Carolyn MacDonald
Sister Francis Regis
Sister Paul Marie Dougherty
Sister Mary Ridge
Mary Cook
Marilyn Flaherty
Dorothy Kron
Raymond Matthews
Marion Cushman
Marti Smith
Janice Connors
Dorothy Lally
John Murray

New Hampshire

Ann Yuscavage
Dorothy Kaffel
Wesley J. Rooker
Gregory Root
Donna Alexandre
Alphonse Soucy
Ramona Dearborn
Charles Yeaton
Iris Bochmann

Rhode Island

Mary Macioci
Roberta Szlatenyi
Patricia Alger
Edward Mitchell
Phyllis Vuono
Sister Mary Francis Ryan
Kathleen Ball
James Emidy
Sharon Capobianco
Joseph De Guilio