A quality evaluation program is needed to determine student progress in the language arts. Evaluation procedures emphasized should appraise in terms of stated objectives. A variety of procedures should be used since no approach is perfect. Approaches include: instructor observation of student progress; completion of checklists or rating scales; journal entries written periodically on each learner; and instructor-developed tests. True-false tests are not recommended due to the possibility of a student's guessing many correct responses. Standards used to evaluate student progress should possess clarity and meaning for the learner. An observation made or test given should also possess the length necessary to determine what students have learned that is relevant and salient. (RS)
EVALUATION, THE LANGUAGE ARTS, AND THE STUDENT

by Dr. Marlow Ediger, Professor of Education
Northeast Missouri State University, Kirksville

A quality evaluation program is needed to determine student progress. Evaluation procedures emphasized should appraise in terms of stated objectives. Methods of evaluation then tend to be valid. Consistency of results in using the evaluation procedures is desired to stress reliability. For example, if two or more evaluators use the same criteria to assess a student, the results should be consistent. Should the evaluators not agree on their results, particularly if the differences are great, the evaluation process has no value.

Evaluating Student Achievement

A variety of procedures should be used since no approach is perfect. Each will have its standard error of measurement. The greater the standard error of measurement, the less reliability in the test and in test results. One approach to use in appraisal is instructor observation of student progress. The instructor must have criteria which represent excellence in student progress in the language arts. Each standard in the criteria needs to possess clarity and worth. Based on instructor observation, the needs of students can be diagnosed and remedied. The instructor may also observe learner interest and responsibility as attitudinal objectives. Politeness, cleanliness, and punctuality are further facets of student attitudes that need evaluation in the attitudinal dimension. With quality instructor observation, language arts students might be monitored more carefully and assisted to attain in a more optimal manner.

It is good if more than one instructor apprises using observational methods. Observations made may then be compared to notice inter-observer reliability. With quality inter-observer reliability, the two or more observers agree rather thoroughly on observations made.

Second, a checklist may be used. The checklist is based on
teacher observation. However, the results are recorded. It is relatively easy to forget or be hazy about a student's progress unless records are kept. Results from several checklists for a student will indicate a pattern of behavior. The language arts instructor can then notice in which area(s) a student needs more guidance, as well as where the strong points of a learner's behavior are in evidence. A checkmark is placed next to the behavior on a checklist where the instructor is satisfied with the student's progress. Those items not checked indicate areas where more help must be given to the student. Checklists can be dated and filed for future use to notice growth on the part of the student.

Instead of checking which behaviors in language arts are acceptable to the instructor, a rating scale may be used. For each behavior, the instructor rates a student on a five point scale. The ratings can go from a five being the highest to a one being the lowest. Low ratings indicate a student needs more assistance and practice in one or more areas. Comparisons may be made of earlier and later ratings given to each individual student to notice progress.

Third, journal entries may be written periodically of each learner noticing strengths and weaknesses in achievement. Each entry should be dated and comparisons made with later accounts of the same student in terms of achievement. Each entry written must be stated precisely and omit vague sentences not based on instructor observation.

Fourth, instructor developed tests can be a source of information to evaluate student progress. With multiple choice test items, the stem and each distracter should make for a grammatically correct sentence. All distracters must be plausible. If one or two distracters are ridiculous, it becomes much easier to secure the correct answer since the absurd distracter is immediately eliminated from consideration. The four distracters should be of somewhat even length; otherwise, it may give clues as to which response is not plausible.

Essay items administered in testing language arts students provide a more unstructured situation as compared to the use of multiple choice tests. There needs, however, to be adequate structure to measure if learners have attained the objectives of the course. Each essay item
must be directly related to the course objectives. Items must be clearly written so that vagueness is not in evidence. Thus students in language arts know what is wanted when responding to each essay item. Vagueness should be eliminated and clarity of content should be salient concepts to emphasize when writing any kind of instructor written test items. One problem in appraising essay test responses is reliability. Thus if a colleague can also appraise the same responses, one can notice inter-scorer reliability. The scorers should come up with the same or similar results for the essay test to have been reliable. If one scorer gives an A grade and the second scorer of the essay test gives a C grade to the same student, the grades given may mean very little, if anything, with such a range.

Reliability may be increased by spelling out rather thoroughly what is wanted by the instructor, prior to testing, in terms of desired correct answers from learners. Essay item one results should be scored separately for all the the students in class before item two is scored followed by item three and others in sequence. Thus the instructor may focus upon each test item results for all students before going on to evaluating the next test item's results. Otherwise, the instructor may be inconsistent in scoring when comparing student's results from each test item.

The writer does not recommend using true- false tests due to the possibility of a student guessing many correct responses. If true- false items are used to measure student progress, learners should correct that which is incorrect in any single test item. This eliminates chances for guessing on test items and their responses.

If matching items are used on a test, they should measure important factual learning of students. There are facts that are basic to the language arts. Each fact on a matching test must be relevant and useful. Trivia must be eliminated in teaching- learning situations. The instructor must attempt in all cases to emphasize cognitive and skills goals of problem solving, critical thinking, and creative thinking. When writing matching items, column one or column two should have more responses than the other to eliminate using the process of
elimination in responding correctly to test items. One of the two columns should have concepts consisting of single words or phrases; the second column may have complete sentences or generalizations. It is difficult for the responder to match the two columns if both have lengthy sentences.

In Closing

Language arts instructors must be good evaluators. He/she needs to appraise student learning to notice if sequential progress is being made. When student achievement is not in adequate evidence, the instructor should diagnose to determine reasons for the occurrence. Remediation must then follow to insure more optimal learner progress.

Evaluation procedures used should measure what students have learned in terms of the stated objectives of the course. Standards used to evaluate student progress should possess clarity and meaning for the learner. An observation made or test given should possess the length necessary to determine what students have learned that is relevant and salient.