Evaluation of Children’s Responses to Literature

DORIS YOUNG KUHN, NCTE Research Foundation

The Challenge of the Foundation

The subtitle of this report might well be “Unwillingly to Test.” The steering committee and larger item writing committee expressed so many concerns over the dangers of such a test and the limitations of a paper-pencil test that it seemed at times the task might not be completed. Yet all have recognized the importance of the project initiated by the Research Foundation of the National Council of Teachers of English. The process of developing an instrument to evaluate some facets of children’s responses to literature has raised important questions about appreciation of literature, development of taste, and, in effect, teaching materials and methods.

Some educators object to the attempt to evaluate emotional responses, attitudes, beliefs. In the recent Taxonomy of Educational Objectives II: Affective Domain Krathwohl (10)* pointed out the erosion of educational objectives which were not evaluated for grading purposes. Have we already lost sight of the goal of teaching so that citizens appreciate literature? Problems of evaluating appreciation relate directly to preservice and inservice education of teachers of English at all levels.

The Research Foundation had announced the availability of funds for research. Surprisingly, only a few proposals were received. At that point, the Foundation became aware of the problems in planning research and discovered the need for evaluative instruments. A conference was called to discuss some guidelines for development of an instrument to test literary appreciation. Perhaps the development of such an instrument would lead to more fruitful research proposals. It was decided to ask a committee to develop a paper-pencil test for children in grades four to six. Margaret Early accepted the Foundation request that she direct the project.

A stream of questions about the proposal came to my mind as I considered the request to serve with Helen Caskey and Norine Odland on the steering committee. Should we test children’s appreciation of literature? Can it be done by a paper-pencil instrument? When we begin testing, may we not destroy the very thing being measured? How can we really separate the literary text and the child’s response to it from his background of experience, which is indeed part of that response? I reread the statement of the task—“a paper-pencil test for grades four, five, and six.” But one test alone is not enough, I argued. Then I recalled

*Dr. Kuhn’s notes appear on pages 74-75.

Copyright © 1966 by the National Council of Teachers of English. All rights reserved.
the many times I had spoken of appreciation as a goal of teaching, the many paragraphs I had read about its importance, the number of statements in curriculum guides which referred to appreciation of literature as an objective of the language arts program. Such statements certainly point to the need for means of evaluating the progress of children. How does the school determine its effectiveness if it does not gather evidence about the objective? It was time to try.

Review of Research

Using a bibliography prepared by Dr. Early, the committee reviewed research and wrote summaries to present at the 1964 Cleveland work conference.

The basic problem of defining appreciation was revealed by comparison of statements about appreciation. Dewey (6) had listed four levels of appreciation: mere vicarious experience, mere attention to form and technique, concern wholly for meanings incorporated in the work of art, identification of oneself with the author's creative experiences. Carroll (5) included "sensitivity to style, appreciate deeper meanings, and emotional capacity to respond to fine shades of meaning." Pooley (16) had written of "... the emotional responses which arise from basic recognitions, enhanced by an apprehension of the means by which they are aroused." Pooley had also written that "any measuring instrument which rests exclusively upon the student's ability to identify and explain the sources of appreciation without first measuring his inarticulate sensitivity to them is not a valid test." Can some means be determined to measure this nonverbal response? Smith (19) had described appreciation as "the personal acceptance of worth." How is personal acceptance reflected in the behavior of the child?

How had others attempted to measure these emotional responses, recognitions of merit, sensitivitics to style? Several studies were based upon the idea that an individual's appreciation can be measured by noting how well he agrees with elements in the original work or those selections deemed best by a panel of experts. Thus, Ruhlen (17) asked subjects to choose one of three couplets which sounded best; Carroll (5) had children rate prose from good books, poor books, and mutilated versions; Eppel (7) had students select the best line to fill the gap in a poem. High school students in a study by Burton (2) were asked to select the best of three endings given for a short story; Cannon (3) asked students to match two prose passages of ten authors.

A number of investigators used the galvanometer and pneumograph to measure physiological changes which might reflect emotional responses. Broom (1) used a galvanometer to determine emotional reactions as certain words were read. Hruza (9) measured changes in pneumograph and galvanometer records as college students listened to fifteen poems. She also used introspection to elicit awareness of emotional responses, and found some subjects had large and continuous galvanic reactions without expressed awareness of emotional reaction. The problem of isolating the many factors such as personal experience and recent events was recognized. I... Smith (20) checked overt responses of two- to four-year-olds as they listened. Laughter, smiles, clapping hands, nods, or expressions of annoyance were kinds of behaviors analyzed. Other investigators used methods
of verbal introspection to assess responses to literature. Piekarz (15) asked sixth
graders to read a selection and then talk about words and phrases from the
selection which were written on cards. Letton (11) also had ninth grade students
report what they thought as they responded to cards with "thought units." Sev-
enth graders were interviewed by Foreman (8) to determine student's "con-
cept of the reality and humanness of the characters in a story; the awareness of
the story purpose and trend; the pictures which are stimulated by the author's
description and completed through the child's experiences." Skelton (18) analyzed
responses to poetry as fourth, fifth and sixth graders wrote "what the poem
means." In 1948 Loban (13) had suggested several approaches as a plot comple-
tion test to identify improbable coincidences or sudden changes of character,
a scale of sensitivity to feelings of a character, or a Social Distance Scale related
to a character, asking students to give reasons for the character's choice. His 1953
study (12) of secondary school students' social sensitivity to literature followed
one of these earlier suggestions.

Many studies of pupil interests and preferences for prose and poetry selec-
tions have been reported. For example, Cappa (4) had analyzed responses of
kindergarten children, noting interests in storybooks read by the teachers.
Mackintosh (14) was one of the early investigators of poetry preferences and
Wells' (21) study of taste in humorous literature revealed preferences for
absurdity, slapstick, satire, and whimsy, in respective order. If a child tends to
like more poems of "literary merit," has he reached a higher level of appreciation?
What is available for him to read is another significant question related to
development of appreciation.

Examination of tests for secondary school students revealed types of items
which might be considered for the instrument we hoped to create for inter-
mediate grades. These tests asked subjects to select the "best poetry" of a group
of stanzas, to identify theme, to indicate sensory appeals, to indicate awareness
of figurative language, etc.

Developing the Instrument

The stream of questions emerged as a flood as the committee worked in
Cleveland; yet we felt the task should be attempted. In its report to the Founda-
tion, the committee stressed the need for research which would include varied
approaches to evaluation. We were, and continue to be, especially concerned
that the study include analysis of materials available and the child's actual
reading. The use of interviews and evidence regarding nonverbal response was
suggested. Recognizing the problem of response to short selections, the com-
mittee recommended that some instrument be developed to assess responses
to longer works of fiction.

Teachers, principals, college teachers, and librarians met in Champaign,
Illinois, in June to construct test items. The participants in this conference
underwent stages of thinking similar to those of the members of the steering
committee. Fears and doubts about paper-pencil tests were expressed. It took
much time to explore the specific behaviors we hoped to measure.

One approach to the study was seen to be the presentation of a selection
followed by a study of the child’s response to the whole. His degree of enjoyment, awareness of meaning, and recognition of value to him were identified as factors to consider. As so many teachers and librarians have noted, observation plus knowledge of the child gives evidence of his response. They say, “There’s a gleam in his eye,” “I just know by his expression that he understands,” or “Through discussion he reveals his understanding and appreciation.” The group discussed the importance of such observations and some considered preparation of an instrument which might be used.

Another approach suggested was that of identifying components of appreciation with their related behaviors. Interests, attitudes, and preferences were seen as one component of appreciation evidenced by such behavior as choosing to listen, listening with attention, choosing reading as an activity, reading a variety of materials, giving attention to different forms. This behavior could be assessed by keeping records of the child’s reading, by anecdotal records, or by checking expressions on a preference scale. Comprehension was seen as another component of appreciation. Reading tests would give information regarding vocabulary and general comprehension skills. Comprehension of such elements as plot, conflict, and character development in each selection becomes the basis for appreciation. Knowledge of standard works (plot, characters, situations), familiarity with authors and illustrators, understanding of literary types or genres, and knowledge of methods of study form another component of appreciation. Written tests, use of allusions in his own compositions, analysis of the child’s book reviews would provide data regarding knowledge necessary for appreciation of literature. Emotional response was identified as a further component of appreciation. The nonverbal feelings may be expressed through overt behavior or may be recalled through introspection. A smile, holding the breath, tears may indicate emotional response. We may ask the child to recall how he felt when he read a particular passage. Can he remember? Does he describe what he thinks the teacher wants to hear? What is the effect of background of experience? What is the effect of immediate events, the mood of the day? The imaginative response and sensory response components are closely related. Verbal comments such as, “I know how it feels,” “You could smell the salt air,” “I really felt the spookiness,” reveal these kinds of responses. Sensitivity to sound effects of language, mood, figurative language, recognition of plot structure, awareness of character delineation, or effect of setting are examples of the component of appreciation termed awareness of literary devices. Written tests could be devised to assess this awareness. Apprehension of meanings would include recognition of values expressed, theme, moral, or universal questions, while the component application of meanings would relate meaning to the individual’s life experience. Judging literary quality is expressed as the child determines literary merit in terms of other books he knows and criteria he is developing. Preferences could be studied, choices determined, book reviews analyzed, and a written test devised to obtain evidence about these aspects of appreciation.

Small work groups proceeded to construct items based upon short selections. It was proposed that a longer selection be presented on one day with a follow-up test on the next day, but there was not time to develop this.
EVALUATION OF CHILDREN'S RESPONSES TO LITERATURE

Material

It was agreed that selections should represent good literary quality. The committee decided that selections should include myth, fable, folklore, realistic fiction, rhymed verse, unrhymed poetry, and both lyric and narrative poetry. It was decided that the component of appreciation which we could consider in this instrument was awareness of literary devices and meanings. It was agreed that the manual should indicate that other components were as important as this one, they should be evaluated, and suggestions for use of the test should be included.

In November the steering committee met in Boston to classify items on prose selections according to the types of responses each seemed to measure. Awareness of meaning was divided into literal and beyond the literal. Now we realized we needed items on awareness of setting, plot, conflict, mood, identification of the narrator. We found there were items for awareness of sound effects, figurative language, identifying genre and dialect, but we needed more items that test awareness of diction, dialect, consistency of characterization, elements of humor. A review of the selections indicated we needed more selections which were easier. There was a great need for more poetry selections. To test awareness of form (including rhythm, assonance, "eye shape," alliteration), symbolism and metaphoric language, and appeals to senses, we needed more poetry and more items.

Work continued and more materials were sent to Educational Testing Service, which is cooperating in the development of the test. ETS will assist with writing items, trying out forms, and standardizing the instrument. It is still in process.

One problem not yet solved is that of oral reading of poetry. Should there be a recording of the poetry so children hear it instead of reading silently?

Another question relates to preference items. Should we ask children to express preference for one poem over another? To what degree would differences in content, length of poem, rhyme, etc., influence the preference? How could we obtain the child's reason for selecting one poem instead of another? How many similar choice items would be needed? These are the kinds of problems we continue to meet.

Implications for teacher education

What are some of the implications of the work of the Foundation in trying to develop an instrument for research related to appreciation of literature for education of teachers of English? I have identified three questions which may be discussed:

1. How do we help preservice and inservice teachers become aware of the behavioral aspects of such objectives as "appreciation of literature"? Appreciation is closely related to the cognitive domain and affective domain, as pointed out by Krathwohl et al. (10). We have assumed that developing cognitive behavior will also develop affective behavior, but this is not true. Knowledge of selections of literature has not produced
interest in literature. How does this problem affect practices in literature classes, methods classes, and student teaching?

2. How do we prepare teachers who are growing in their own appreciations? Are they aware of the process of literary criticism, or methods of literary criticism? Or do they follow the practice of a bright student completing a 24-hour "second major" in literature who says, "I try to tune in on the professor's method of study; I really haven't been taught any principles of literary criticism. I think you get methods of study in graduate school." How do we teach "structure of the discipline" as suggested by Bruner? Related to this is the problem of relating study of literature to the materials designated as children's literature.

   Another aspect of the problem is development of taste. Do teachers understand the process of valuing? Do they continue to believe exposure to good literature is enough? How do we prepare teachers who are skilled in guiding pupils so they develop criteria for making value judgments?

3. How shall we help preservice and inservice teachers develop skill in evaluating growth of students? We need to help them develop methods of evaluation, to go beyond marking compositions. If we do not consciously evaluate the many components of appreciation, will we continue to fail in this objective?

Several members of the committee expressed concern that teachers would misuse the instrument we were developing. Teachers have misused intelligence tests, achievement tests, and sociograms, but the fault lies in not understanding the limitations of the instrument. How do we prepare teachers who are capable of using instruments effectively? What is done to engender interest in research, both their own classroom research and research findings of others? What instruments are used in college literature classes to evaluate growth in students? Future teachers do learn about teaching from the teaching-learning experiences they encounter. Is the cognitive domain the only area for evaluation? In methods classes and student teaching, what experiences should students have in constructing tests and using a variety of instruments to gather information about pupils' progress in both the cognitive domain and affective domain?

The need for research in the teaching of literature is one of the major needs in teaching English language arts. The development of one paper-pencil instrument to measure awareness of literary devices and meanings is but one small part of the major task of developing means to study children's responses to literature.

References