Application of Performance Appraisal Systems to Evaluation of College Composition.

Drawing from a review of the literature and a comparative study of evaluation techniques, this paper offers a rationale for the use of performance appraisal as a method of evaluating student writing. The literature review covers: (1) various roles played by composition teachers, including the students' audience of one, writing coach, and Socratic tutor; (2) the art of rhetoric and the science of linguistics; (3) criteria and scales for evaluating student writing holistically, including Diederich's scales of general merit and mechanical skills and Lloyd-Jones's primary trait scoring method; (4) the importance of communicating expectations and evaluation criteria to students and providing feedback on how well their compositions meet these standards; (5) the distinction between grading and evaluating; and (6) the limitations of revision checklists. The next section of the paper explains and offers support for the use of performance appraisals, like those used in industry, to make students aware of the factors that are priorities for an assignment before they turn in their papers. The final section reports on a study comparing the effects on student writing of direct teacher instruction and collaborative learning in peer groups as means of providing feedback. Appendices include 76 references, a chart on the development of rating scales, a sample assignment and revision/grading sheet for the assignment, and two additional assignments with a list of grading points for each.
APPLICATION OF PERFORMANCE APPRAISAL SYSTEMS TO EVALUATION OF COLLEGE COMPOSITION

by

Roberta S. Boss, Ph.D.
Dept. of English
Professional Writing Program
2117 Taliaferro Hall
University of Maryland
College Park, MD 20742
Performance Appraisal

ABSTRACT

With the passage of the Civil Service Reform Act in 1978, the federal government was directed to acknowledge that a worker's awareness of expectations of the evaluating manager or supervisor for a particular job would necessarily result in better job performance. Not only would the new employee receive a detailed job description as in the past; along with the Position Description would be a copy of the Performance Appraisal checklist with the criteria and standards by which that employee would be periodically rated. Performance Appraisal, mandated by law, is a systematic procedure which is more structured than its predecessor, Performance Rating. Although there is some latitude allowed to the individual agencies in writing critical job elements and acceptable performance levels, managers must share these criteria with their staff members. If a worker fails to meet the standards, a plan for improvement, either by additional training or closer supervision, is presented so that the employee cannot be summarily dismissed. Performance Appraisal involves management by objective and is based upon the principles of behavioral psychology. Behavior can be modified when positive reinforcement and specific, timely criticism are given. Adequate feedback serves as a mirror to the individual. In education, behavioral objectives represent observable evidence of desired learning achievement. Application of Performance Appraisal Systems to evaluation of college composition can afford a clearer picture of "good writing" to both students and teachers than has thus far been imagined.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Problem</td>
<td>1</td>
</tr>
<tr>
<td>The Teacher's Role</td>
<td>1</td>
</tr>
<tr>
<td>The Art of Rhetoric &amp; the Science of Linguistics</td>
<td>2</td>
</tr>
<tr>
<td>Identifying Criteria</td>
<td>4</td>
</tr>
<tr>
<td>Communicating with Students</td>
<td>6</td>
</tr>
<tr>
<td>Process vs. Product</td>
<td>7</td>
</tr>
<tr>
<td>Revision Checklists</td>
<td>7</td>
</tr>
<tr>
<td>Toward a Solution</td>
<td>8</td>
</tr>
<tr>
<td>Performance Appraisal Systems</td>
<td>9</td>
</tr>
<tr>
<td>A Dual-Purpose Instrument</td>
<td>12</td>
</tr>
<tr>
<td>The Study</td>
<td>15</td>
</tr>
<tr>
<td>Flexibility</td>
<td>16</td>
</tr>
<tr>
<td>Scoring Guide</td>
<td>16</td>
</tr>
<tr>
<td>Table 1</td>
<td>17</td>
</tr>
<tr>
<td>Purpose and Procedures</td>
<td>19</td>
</tr>
<tr>
<td>Results and Conclusions</td>
<td>19</td>
</tr>
<tr>
<td>For Further Study</td>
<td>21</td>
</tr>
<tr>
<td>References</td>
<td></td>
</tr>
<tr>
<td>Appendix A: Developing Rating Scales</td>
<td></td>
</tr>
<tr>
<td>Appendix B: Sample Assignment Sheets with Grading Scales</td>
<td></td>
</tr>
</tbody>
</table>
APPLICATION OF PERFORMANCE APPRAISAL SYSTEMS
TO EVALUATION OF COLLEGE COMPOSITION.

The Problem

Whether teaching literature, composition, or rhetoric, most English teachers dislike grading papers. Their compromises run the gamut from no grades at all during the semester to grading only selected essays, to complex point or percentage systems applied to major themes, in-class exercises, and homework. Gene Stanford has compiled some of these alternatives to traditional grading methods in How to Handle the Paper Load (1979). Yet timely response to papers is necessary to mark and acknowledge students' progress, to ascertain whether what is "taught" is actually "learned." Otherwise, there is no counter-argument that improvement in writing comes inevitably through practice, regardless of teaching style, materials, or course content.

Peter Elbow represents the extreme: teachers should relinquish the responsibility of evaluation to the students themselves—often without guidelines (1973; 1981). His "Center of Gravity" response sheet inquires vaguely about a mood or impression that the reader can take from the essay.

The teacher's role. Donald Murray emphasizes that the teacher, however skilled as a literary critic, is still an audience of one. He suggests that perhaps students are more attuned to the writing of their peers (1968; 1984). Why force a teacher to be a surrogate reader when other students might provide a more authentic audience? Gibson's (1980) "teacher as dumb reader" is not only unrealistic but may also be unconvincing to students.

The trial-and-error workshop setting is effective for Murray, but he sees his role as "writing coach" rather than total abstainer,
a la Elbow. Murray's students experiment with journalistic "leads" into a story; Ann Berthoff reacts to students' journal entries and "lists" until order comes from what she sees as the natural state of chaos (1978). Lee Odell demonstrates that business correspondence entails intuitive rhetorical choices by untrained writers, and he tries to make these conscious and deliberate (Cooper & Odell, 1978). Flower and Hayes tape record "writing protocols" as students try to verbalize their mental processes while composing (1980a).

Teachers evaluate writing from an implicit definition of its functions. Writing as an expressive art (Coles, 1974, 1978; Irmscher, 1979) calls for a critic to appreciate innate talent. For the skilled craft of writing, masters train apprentices to use "tools" and techniques (Christensen, levels of generality & texture, 1978; O'Hare, sentence-combining, 1973; Hunt, T-units, 1977; Young, Becker & Pike, tagmemics, 1979). When writing is seen as an exercise in rhetorical problem-solving (Flower & Hayes, 1980a; Gere, 1980, 1985; Lloyd-Jones, 1981), the teacher becomes a Socratic tutor. Whatever the emphasis, some teachers retain the role of editorial advisor to former students well into the future.

The Art of Rhetoric and the Science of Linguistics

The idea of applying classical rhetoric to writing was proposed by Edward Corbett back in 1971, when his students engaged in stylistic imitation (new wine in old bottles). Since then, rhetorical principles have influenced many textbook writers, especially during the 80's. They are prefaced by a description of transactional writing (Britton, 1978) and audience accommodation (Bracewell, Scardamalia & Bereiter, 1978), the necessity of considering purpose and intended reader in producing effective writing. (See especially the rhetoric handbooks of Gere, 1985; Hairston, 1982; McCrimmon, 1980;
and Winterowd, 1981). In a series of experiments, Hirsch and Harrington (1981) were forced to discard their notion of "intrinsic communicative effectiveness," finally admitting that writing is "effective" only with respect to a specific reader. But here again, teacher-as-know-it-all intrudes. Producing what the teacher wants is proper according to rhetorical principles, for that is the student's intended audience. The purpose of a school paper is to get a high grade (see Bereiter, 1980, on performative writing). All protests to the contrary: that each type of writing carries a set of intrinsic criteria or that an argument rests upon a thesis supported by specific example, remain unconvincing to student writers (Murray, 1968).

Similarly, the field of linguistics has disappointed some teachers who expected definitive criteria to emerge from its study. Roger Shuy (1981) discusses "the past overpromise of linguistics" where the hierarchy of language components: semantics, coherence, syntax, and at the bottom, lexicon, morphology and inflection, and mechanics (Halliday & Hasan, 1976) gives so much information as to overload the fledgling writer. For the teacher, it is probably as easy to count superordinates and subordinates (Christensen, 1978) as it is to mark incorrect apostrophes, comma splices, and dangling modifiers. Perhaps after all of this linguistic data has been absorbed, teachers feel constrained to note and comment on everything. Murray empathizes with the instructor who goes error-hunting with a red pen through jungles of student writing, but more so with the student who receives the red-stained paper and is asked to revise it (Murray, Write to Learn, 1984; Odell & Cohick, "You mean, write it over in ink?" 1975).
Identifying Criteria

Any English teacher...will tell you how difficult it is to make a value judgment of a pupil's essay response. Yet criteria lurk whenever this teacher does make a judgment, and these criteria must be made explicit (to the student).

(Popham, reported in Kibler et al., 1974, 14.)

In a study preceding the adoption of the Composition Evaluation Scales by the ETS to rate the essay portions of the SAT's, 53 "expert readers" were asked to score a set of 300 college papers on a scale of 1-9.

101 papers received every grade from 1 to 9 on the scale; 94% from seven to nine different grades; and no essay received less than five different grades from the 53 expert readers.

(Diederich, 1974.)

Next, the readers were to write detailed rationales to explain their scoring methods. So diverse were the criteria named that a factor analysis was not possible until terms were homogenized under umbrella categories or "clusters." Even then, statistical reduction of components was difficult. It finally became apparent that the raters, depending upon their occupational fields and experience, were examining different factors or labelling the same ones differently, weighing them differently to arrive at scores, and even disagreeing as to the nature and significance of errors.

Mina Shaughnessy comments:

Definitions of proficiency in writing vary widely...with the [least] agreement at the upper rungs, where the stylistic preferences of teachers come into play. But even within the province of error, there are disagreements about the importance of different errors and about the number of errors an educated reader will tolerate.

(Shaughnessy, 1977, 276.)

The final Diederich scales list eight factors or "clusters" under "General Merit" and "Mechanical" headings: topic/ideas; organization; vocabulary/phrasing; persona/style; language use;
punctuation; spelling; format/handwriting (Composition Evaluation Scales, 1961). Double weight was assigned to "ideas" and "organization" at the insistence of English teachers among the raters.

With Diederich's scales, it is possible to score a large number of essays quickly and easily using holistic methods. For the purpose of ranking college applicants, a composite score derived from the eight separate scales is sufficient, since the essay is only a portion of the SAT battery.

Like the Diederich scales, many widely accepted analytical composition scales have been developed for administrative or research use in making group comparisons. (See especially Fagan, Cooper, & Jensen, 1975, for a handbook of composition scales.) Yet such grading scales are of little use to the classroom teacher who is concerned not with numbers or trends but who needs diagnostic information to help individual students (Charney, 1984; Coffman, 1971). The analytical scales seem to be misnamed, for categories are too general to indicate the strengths and weaknesses of a particular paper. Furthermore, with General Impression Marking (holistic rating) only a numerical score, usually a composite number, is reported, and papers remain unmarked.

The methods perfected by the ETS assume that excellence in one sample of one mode of writing predicts excellence in other modes—that is, good writing is good writing.

(Lloyd-Jones, 1977, 37.)

Dismissing the Diederich scales for failure to consider context, Richard Lloyd-Jones developed Primary Trait Scoring under the auspices of the National Assessment for Educational Progress (1969-1970). It is built upon a model of discourse that he credits to Aristotle: ethos, the writer's character; logos, the subject matter; pathos, audience appeal—all must be considered in evaluation.
For the test, the writing task is structured narrowly, to fit a rhetorical situation. For example, students are directed to select a consistent point of view, use dialogue, and remember tense for narrative, and raters are to consider only these "primary traits" in scoring. Levels of proficiency are limited to 0-2 or 0-3, in contrast to the typical five or six points on holistic scales (see American Council on Education, 1985, for GED Scoring Guidelines). Primary Trait Scoring is used to assess large numbers of compositions produced under controlled testing conditions at various age levels in both horizontal and vertical studies by the NAEP. It has potential for the classroom with some modifications.

Communicating with Students

Charles Cooper and Lee Odell define holistic scoring as any method which stops short of counting linguistic features of text. They maintain that criteria do exist for evaluation, but that teachers need merely to "keep them in mind" while grading papers. They do not mean to suggest that students become literary mind-readers (or literal mind-readers)! Students do have a right to know by what means their assignments will be graded.

Any approach to evaluation must come to grips with the problems of grading; this is best done by discussing the problems with students. Initially, this requires that the teacher state the criteria. Unless the teacher quantifies every element in language skills, the teacher is responsible for either presenting students with criteria or encouraging student-generated criteria.

(Lague & Sherwood, 1977, 84, emphasis added.)

Even if it were possible, is it desirable to "quantify every element" for students? The issue of making criteria available to students is often misunderstood to imply that students must determine their own grades. Rather, self-analysis helps students to develop a sensitivity toward language alternatives.
Process versus Product (1985)

Anne Gere makes a distinction between grading (a finished paper) and evaluating (a draft). Only through between-draft revision can a student be provided with usable and timely information for improving writing (Beach, 1979). "The message of marking" (Searle & Dillon, 1980) is closure: the paper is finished once a grade is given. Yet teachers are understandably reluctant to read every student's drafts and final copy for every assignment throughout the semester (back to the "paperload" dilemma, Stanford, 1979). Those teachers whose painstaking interlinear corrections, marginal abbreviations, and long editorial comments seem to be ignored by students are discouraged. For students, the advice comes too late for the current assignment and is premature for the next. The next paper will present either opportunity for a fresh start or another futile attempt to read the teacher's mind (Emig, 1977; Sommers, 1980).

When students are cast into the editor's role, they may be able to gain from the "hands-on experience." This is not to say that "decomposing text" is the key to composing (Berthoff, 1978). It simply means that when students can judge objectively the work of others, they can often apply such objectivity to evaluation of their own writing. (See the peer group studies of Beaven, 1977; Benson, 1979; Danis, 1980; James, 1981; Pianko & Radzik, 1980.)

But peer groups must be trained to give (and receive) criticism, and they need guidelines for revision. Although teachers understand the importance of revision in improving writing, students need to be convinced that their "first shot" is not "the best I can do"!

Revision Checklists

Revision means word choice, spelling, and sentence structure--the simplest proofreading skills--to most students. Yet teachers
expect a second draft to represent significant improvement over the
"rough" one. Yet a checklist cannot include every "element of
language skills" (Laque & Sherwood, 1977). Sound educational prac-
tice dictates that a few serious problems be tackled at a time, so
that students are not overwhelmed by a gigantic task (Kibler, Cegala,
Barker & Miles, 1974). However, a teacher who has presented a
checklist specific to the current assignment and then encounters
unforeseen problems in the set of final papers ("Omigosh--they
still can't avoid fragments!") is forced to revert to intuitive im-
pressions to assign grades. Mixed writing products (perfect form
lacking a thesis, or insightful analysis with faulty grammar) may
get split grades. In this case, students become mistrustful (But
didn't she say that logic was more important than spelling?)

How can students be made aware of those factors which are
priorities for an assignment before handing in the paper for grad-
ing? How can a teacher intervene during the composing process so
that students do not have to sit back and await the omniscient
judgment? (Flower & Hayes, 1980b).

Toward a Solution

Suggestions for involving students in their own learning pro-
cess run into obstacles: both teachers and students may rebel
against the ostensible role reversal. Yet students can be led to
seek their own answers if both parties have the patience to refute
the myth of teacher-as-know-it-all. Gene Hammond (1985) presents
a long list of writing criteria which students must rank-order at
the very beginning of the semester. Amazing to both teachers and
students themselves, criteria are as available to any reader as to
"expert readers." Students are especially surprised that there is
not a definitive list--no one right answer (Perry, 1970).
Performance appraisal systems. What is the effect of "knowledge of performance" on a person's work? In a study made three decades ago, Ammons (1956) found that when an employee was given a copy of his job review, the supervisor could expect improvement in critical areas. The jargon is "feedback," and with timely and appropriate behavioral feedback, a speaker can gauge audience reaction, a computer student can repeat a tutorial, and a writer can make revisions suggested by an editor.

In the federal government service, personnel management systems depend upon efficient evaluation and review procedures which have their foundation in behavioral sciences research. From the employer's perspective, differences between expected and actual performance can be minimized by explaining tasks and responsibilities to the new employee. Often, a deficiency can be traced to lack of knowledge of requirements. Insufficient feedback from the supervisor may be the culprit, rather than a lack of ability to undertake and complete tasks (Rummler, 1972).

The Civil Service Reform Act of 1978 states:

Each agency shall...provide for periodic appraisals of job performance by employees; encourage employee participation in establishing performance standards; use the results...as a basis for training, rewarding, reassigning, promoting, reducing in grade, retaining and removing employees.

(U.S. Code, Title V, Ch. 43, Sect. 4304, emphasis added.

Morrisey (1972) advocates personnel development through systematic appraisals. Mager and Pipe (1970) claim that behavioral objectives in education can aid both the teacher (in the manager's role) and the student (as promotable employee). A review of the literature in "Behaviorally Anchored Rating Scales" prefers performance to be measured in multi-dimensional terms rather than a single overall...
composite score (Schwab, Heneman & Decotis, 1975). Jacobs (1977) criticizes management in that clear performance standards applicable to specific occupations are not always provided in advance of supervisory observation. In this case, employees are unaware of the basis for ratings.

Although the terminology of educational evaluation and job appraisal differs, both are rooted in behavioral psychology. Personnel managers try to identify "critical elements," those tasks which are necessary to obtain "an acceptable level of competence." (Federal Personnel Manual, Ch. 531). Teachers assess skills and competencies in terms of numerical percentages or letter grades.

Management by objective is similar to use of instructional objectives: both begin with a task analysis to derive critical factors of performance. Rating is always criterion-based, as contrasted with norm-referenced ranking in screening applicants for a position or for college placement, for example.

In the federal service, position analysts and classification specialists collect job information from position descriptions, organizational charts, training manuals, on-site interviews, and direct observation of workers. Critical elements are listed for specific jobs and are clustered with "desirable" elements, weighted for each position, and narrative descriptions of Fully Acceptable are written. Outstanding and Unacceptable levels are later extrapolated. These "behaviorally-anchored standards" are usually placed on a five-point scale. Elements and standards are written together, because employees need to know not only what is expected of them but how well they must perform. New employees receive not only a detailed job description but also a copy of the review sheet by which their job performance will be evaluated. Current employees are reminded of the critical elements and standards of their posi-
tion before the evaluation is scheduled, and they receive a copy of the supervisor's report when it is completed, with recommendations for appropriate job actions. Under this system, an employee may not be summarily dismissed or downgraded; suggestions for more training or closer supervision are given with the report.

The participatory method. The Civil Service Reform Act of 1978 specifies:

Each appraisal system shall provide...written Performance Plans to employees at the beginning of each appraisal period. An agency shall encourage employee participation in establishing performance standards.

(Code of Federal Regulations, Part 430, emphasis added.)

Employees who help to set standards and choose critical areas of their own work assignments are likely to view them as fair and reasonable. A research study made two decades ago has not yet been refuted in its findings that goal-setting is related to subsequent job performance improvement (Mayer, Kay & French, 1965). The law requires that critical elements and standards be communicated to them upon appointment; at the beginning of each appraisal period; and during and at the end of the appraisal period. The Performance Improvement Plan is developed by a manager with an employee who has received a "less than"fully acceptable" rating, so that they both have "input" in the plan.

When performance standards are applied to education, the terminology but not the intent changes. Levinson suggests "a means of giving management and employees information that they both need...and a method for improved communication (between them)" (1979). Substitute "teacher and students" and it still makes sense.

...students learn more rapidly and reach higher levels of performance when provided with knowledge of performance...upon completion of a formative test, a student should be provided with immediate feedback of the
objectives answered correctly. Also by identifying unmastered objectives, the student may then be directed to corrective activities."

(Ammons, 1956)

Feedback of correct responses provides positive reinforcement, which is necessary to sustain motivation and build confidence. Often, students ask only, "What did I do wrong?" not "What have I learned?"

If instructors can objectify their criteria and standards for writing in advance of the due date of the paper, and especially if students can participate in identifying criteria and priorities, they are more likely to understand the basis for their grades. Students can be trained to use those criteria in drafts and in revisions when those "critical elements and standards" are predetermined and predefined. This is a given condition in Performance Appraisal Systems, but students too often receive information after the fact of grading.

The chart in the Appendix was developed to illustrate the application of Performance Appraisal Systems to the evaluation of college composition.

A Dual-Purpose Instrument

...analytical scales (used) in the classroom...provide a public statement of the general features of writing stressed in a writing course, and they provide a focus for students' examination of their own writing and the writing of others.

(Cooper & Odell, 1977, 21.)

However, the examples given for adapting composition scales for student use are ambiguous. For instance, in a Personal Narrative Scale, under "General Qualities," the author's role is described for teachers:

...the relationship of the author to the subject, incident, or person, (using) pronouns, keep[ing] his/her correct role of either participant or observer throughout.
As a dichotomous checklist, the category reads, "Author's role consistent," with space provided for "yes" or "no." Thus the full description, meant for teachers, provides more information than needed, while the checklist for students offers too little.

Usually, when students are permitted to use checklists for revision of drafts, no grading scale or scoring guide is provided. And teachers typically do not assign grades by means of the checklist. Those who suggest adaptation of an analytical scale for student use (notably Cooper & Odell, 1977; Larson, 1968; Sager, 1973) do not intend that the revision checklist and grading scale be identical. Yet presentation of the instrument complete with scoring guide in media res can give direction to students while they are composing. The checklist can function in setting priorities for a particular paper. Student grading of drafts need not be seen as usurping the teacher's authority. Rather, this could provide information to the teacher for lesson-planning, whether on particular grammar points or on more sophisticated language use or problems of logical organization. The checklist serves as a teaching tool, explaining criteria and standards in objective terms for the students. In a study of methods of grade reporting, Stanton (1974) found no differences in student achievement whether the teacher wrote on papers, conducted in-class question-and-answer sessions, or used a checklist, except that the last seemed to "keep teachers more reliable (in grading)."

This researcher's instrument (Boss, 1986) is designed for a dual audience, one party of which is motivated by grade and the other by writing behavior. The major concern (and sometimes the sole consideration) of students is grade improvement, and that is the fulcrum about which their perceptions swing. The central motive of teachers is to help students to develop strategies and options
in writing, a difficult task to accomplish with a summative letter grade on a theme. Whether the checklist is used in self- or peer-critique, it is an indicator of both strengths and weaknesses in the draft: what areas have been competently handled and what problems still need to be addressed before the paper is due for grading.

The grading scale which has first been used by students for revision is no longer a post-mortem summary of deficiencies. Students can be trained to use the criteria (or better still, to develop them in collaboration with the teacher) from the very first assignment, where "grading points" are listed at the bottom of the assignment sheet. Some teachers may prefer to withhold the checklist until the first draft has been produced. Either way, the checklist is used as a grading sheet by the instructor, who remains the final judge and critic. However, that judgment is now more meaningful to students than are the usual written comments and corrections. No longer is there reason to wonder, "What does the teacher want?" or "How can anyone get an 'A' on this paper?"

The scale is used twice for each assignment: once by students to revise first drafts, and then by teachers for grading final copies.

Peer workshops are normally constituted to provide critical feedback to the student writer, not to usurp the teacher's authority in awarding grades. Long ago, Maize recommended workshop grading in class rather than teacher grading alone at home or in the office, so that students could be more involved in the process (1952). Since then, several studies comparing peer critique with traditional teacher grading found that when goals were pre-set and students trained in the critical skills, writing improvement resulted in both situations (Beaven, 1977; Lagana, 1972).

Benefits accrue to both parties when a dual checklist/grading
scale is employed. Goals and objectives specific to the current assignment are kept actively in mind by the teacher while grading. The response is more directly related to what has been taught in class about that particular type of paper. Grading itself is less time-consuming when guided by specific criteria. For students, instead of merely handing in their papers and waiting for a grade, they become actively involved in determining the criteria. Far from the traditional passive dependence upon teacher reaction, students come to an appreciation of good writing through their own efforts as proofreaders, editors, critics, and finally, as self-evaluators. At last, the "keeper of the mystery" can be persuaded to share it with every student (Perry, 1970).

The Study (Boss, 1986)

The revision checklist/grading scales were developed under the following conditions:

1. Criteria and standards were objectified for each assignment, matching the rhetorical requirements specified on the assignment sheet.

2. The high performance level was described in each category, referring to "critical components" of each.

3. Criteria were chosen to combine the thoroughness of the Diederich scales with the specificity of Lloyd-Jones' Primary Trait Scoring.

4. The scoring system was simplified so that the scales could be used for revision independently from teacher direction.

The scales were thus meant to provide each student with a rubric and vocabulary to discuss and view his/her own work. For the purpose of the study, the scales were constructed by the researcher. Ideally, categories would be derived inductively by the students themselves guided by their teacher.
Flexibility. For each assignment, six categories were chosen to represent both Primary Traits for that writing task (Lloyd-Jones, 1977) and more general criteria such as mechanics and word choice/diction applicable to all good writing (Diederich, 1974). It was hoped that a judicious mix of criteria would help students to concentrate upon a few strategies or skills at a time. (In contrast, see Olson's "grading slips" 1982, which have no logical basis for selecting criteria.)

Scoring guide. A problem was encountered in approaching teachers with yet another composition scale. There was resistance to incorporating a complex system of percentages into a writing course. Teachers were also concerned about conversion of percentages to letter grades.

Since composition grades are derived not from objective measures but rather from the subjective judgments of teachers, a means of applying holistic scoring to analytical reading was developed. The simplified system is based on Lloyd-Jones' which awards a maximum of 3 points within a category and totals the points to rank papers.

Holistic scoring mandates that features of text must not be counted or deeply analyzed. In fact, this stipulation is part of the definition offered by Cooper and Odell (1977). Instead, in each category, a plus (+) for superior is worth 2 points; a check (✓) satisfactory, 1 point; a minus (-) unacceptable, zero. This is the extent of the discrimination made, much like the "high, middle, low" placement on the Maryland Functional Writing Test (1982). It involves less deliberation than the five levels of the Diederich scales. Holistic scoring normally ranges through 5 or 6 levels (American Council on Education, GED Scoring Guide, 1985).
Following the federal government's performance level guide, three levels were chosen as the ideal in combination with an analytical instrument.

Most of the composition scales examined by the researcher (see especially Fagan, Cooper, & Jensen, 1975) do not include scoring guides. Where directions for grading are given, they are complicated by the means of conversion to percentages and by weighting of factors. On the present scales, the maximum score is 12 points (6 categories x 2 points). Points convert to letter grades as follows:

<table>
<thead>
<tr>
<th>Points</th>
<th>Grade</th>
<th>Points</th>
<th>Grade</th>
<th>Points</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>A+</td>
<td>8</td>
<td>B</td>
<td>4</td>
<td>C-</td>
</tr>
<tr>
<td>11</td>
<td>A</td>
<td>7</td>
<td>B-</td>
<td>3</td>
<td>D+</td>
</tr>
<tr>
<td>10</td>
<td>A-</td>
<td>6</td>
<td>C+</td>
<td>2</td>
<td>D</td>
</tr>
<tr>
<td>9</td>
<td>B+</td>
<td>5</td>
<td>C</td>
<td>1</td>
<td>D-</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0</td>
<td>F*</td>
</tr>
</tbody>
</table>

* Students who received an F by not handing in a paper were not counted in the grade breakdown report for the study.

This scoring system did not affect an individual teacher's policy on late papers, which the grade report for the study excluded. The investigation did not extend to determination of course grades, and students were advised that participation in the study would affect neither individual assignment nor final course grades.

The simplified scoring system is one major advantage; the other is the flexibility of the categories and critical components chosen. For example, for Description of an Unusual Place, the first category reads:

**Dominant Impression/Mood**: Unity of details, vivid and concrete sense images without distracting associations, liberal use of picture words.

The "high" level is described, the quality deserving of a plus (superior, 2 points). The categories and their descriptors are...
specific to that rhetorical mode. Rather than a broad heading such as "ideas/content" as on the Diederich scales, Description calls for "dominant impression/mood." Instead of "organization," the arrangement of details is spatial. For the Process paper, order of details is chronological.

Categories are not listed in order of importance, and there is no weighting of factors. Each criterion is considered a major contributor to overall writing quality, reflecting the "juggling of constraints" upon the writer (Flower & Hayes, 1980b). However, to avoid giving students the wrong impression, "correctness" categories are not first on any of the scales. In later assignments, as mechanical aspects have been practiced and remediated, "local" categories are combined to allow for more complex criteria (Freedman, 1981; Nold, 1981).

The six criteria on each scale include cognitive, linguistic, and rhetorical goals, representing the terms of a learning contract: "This is what must be done at this level of proficiency to earn this grade." The scales serve not only as an instrument of formative summative evaluation (writing product) but also as a teaching tool (composing process). One form to guide both revision and grading renders much deliberation about writing quality unnecessary. Teachers can mark final papers quickly, with confidence that students already know the terminology through which their faults and strengths will be expressed. Papers need have no marking or notations, for the grade sheet outlines priorities for that assignment, and the "critical components" are circled in those categories receiving a minus.

Scoring has been adjusted to accommodate special needs as well. In a philosophy course, the instructor used five criteria for a
term paper, with 10 points maximum, and converted to percentages. Another variation is to choose 12 categories for a research paper, assign 0-2 points in each as usual, and divide by two for the grade.

**Purpose and Procedures**

The purpose of the study was to explore the comparative effects on student writing of peer critiques versus teacher response to first drafts. The study investigated whether direct teacher instruction or collaborative learning in peer groups was the more effective in providing feedback to improve writing.

There was actually no control group, in that both groups received the revision checklists and were told that teachers would grade on the basis of these and no other criteria. The hypothesis was that no matter the goals of the teacher for each of the six major assignments, the sharing of those goals in advance of grading would result in grade improvement far beyond that expected through practice alone.

The direct instruction group received the checklist as a handout, but drafts were collected, read, and commented upon by the teacher, both in writing and in class discussion. Final papers were graded as usual, with papers receiving written comments as considered necessary (after drafts had been thoroughly marked).

The peer critique group used the checklist first on sample papers, then on classmates' drafts, with two editors agreeing on a grade. Teachers checked on the peer grading sheets when final papers were collected; drafts were not seen by teachers. Final grades reported by teachers on grade sheets only; papers remained unmarked.

**Results and Conclusions**

All of the treatment groups but one registered significant grade improvement (t-scores above critical 1.67, p < .05) as measured by
pre- and post-test sample writing scored holistically by independent raters (using ACE guidelines, 1985).

Teachers had no problem incorporating holistic rating into their essay reading. Many instructors recognize their own intuitive response in the procedure: reading quickly through a set of papers for "norming"; separating into "poor, average, and good" stacks; grading on a curve. Yet several of the teachers had to be reassured that reporting grades on the grading sheets for the peer critique groups would be sufficient--they still worried about lack of written comments and corrections on the papers until late in the semester (after four of the six major assignments had been completed).

Students in the peer groups complained about "no written comments from the teacher," but from class observations, the researcher soon realized that the topic was corrections, not comments, on papers. No matter how skilled the peer editors, students continued to look to the teacher for all guidance.

Additionally, students were fixated for a long time on the quantitative grading points. Training them in holistic scoring helped to shift focus from error-hunting at the word and sentence levels to consideration of whole discourse quality. As recommended by the ACE (1985), "Is this an upper-level or lower-level paper?" should replace, "Why does this category deserve only a 1?"

Student attitudes toward the grading procedure were checked through questionnaires and interviews at the end of the semester. Many had expressed concern during class observations by the researcher that the teacher was not grading on a curve, and that they had difficulty in comparing themselves to the rest of the class. There was no significant difference between the responses of the two groups, and so all answers were pooled. 78% of all students felt that usually or often "requirements for 'A' papers were made clear for
75% said that usually or often "the teacher did not show favoritism in grading." 70% "knew throughout the semester of my standing in class" (even though they had complained about this very thing during class observations!) 60% said, "I was kept informed of my strengths as well as my weaknesses." 53% believed that "my grades on final papers have been what I expected."

One of the most ambitious and largest-scale studies, begun in 1973, corroborates the findings of this researcher:

The basic premise behind this use of small groups is that, to provide a student writer with a sense of audience, he must receive audience reactions while engaged in the process of writing, not at the end when the paper has been handed in, days have gone by, and the piece is handed back, minutely evaluated by the teacher.

Bay Area Writing Project, (Healy, in Camp, 1983, 166; author's emphasis.)

For Further Study

This study was not concerned with validating the scales, but rather with the effects of advance knowledge of writing criteria, and different ways of communicating those criteria to students. A new investigation might compare grades awarded by the holistic method or the Diederich scales with grades on the researcher's scale. A control group might be introduced which does not receive a revision checklist, while the teacher continues to use the researcher's scale to determine (but not to report) grades. A departmental- or district-wide program in high schools might guide all grading with the researcher's scales without affecting teaching or classroom practices. Instead of peer groups, the scales might be used in conjunction with personal conferencing.

With a participatory system of criteria-seeking and goal-setting (Beaven, 1977), and with revision opportunity guided by the same scale that the teacher uses to assign grades, perhaps students can move from total dependence on teacher as proofreader, editor, critic.
REFERENCES


Code of Federal Regulations. Civil Service Reform Act of 1978, Title V, Ch. 43.


Emig, J. *Writing as a Mode of Learning.* College Composition and Communication, 18 (1977) 122-128.


Maize, R.C. A Study of Two Methods of Teaching English Composition to Retarded College Freshmen. Unpublished PhD dissertation, Purdue Univ., 1952.


Sager, C. Improving the Quality of Written Composition through Pupil Use of a Rating Scale. Univ. Microfilms #AAD 73-23-605. Sager's scale also available from ERIC #ED 091 760 (1973).


APPENDIX A. Developing Rating Scales

<table>
<thead>
<tr>
<th>Task/Outcome</th>
<th>Job</th>
<th>Essay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data-gathering by managers, teachers, researchers, for &quot;critical components.&quot;</td>
<td>Analyze position descriptions, training manuals, classification specs.</td>
<td>Read widely authors, student essays, composition theory.</td>
</tr>
<tr>
<td>Validation of criteria.</td>
<td>Consensus reached by managers &amp; position classification specialists.</td>
<td>Consensus reached by expert readers &amp; raters.</td>
</tr>
<tr>
<td>Weighting factors, writing standards.</td>
<td>Narrative description of fully successful with other levels extrapolated.</td>
<td>Narrative descriptors of high/middle/low with benchmark models chosen.</td>
</tr>
<tr>
<td>Performance Appraisal Report</td>
<td>Essay Grading Sheet</td>
<td></td>
</tr>
<tr>
<td>3-4 achievement levels as basis for pay increase, promotion, adverse job actions.</td>
<td>5 achievement levels numerical score converted to grades A-F.</td>
<td></td>
</tr>
<tr>
<td>Advantages:</td>
<td>Motivation, employee involvement, feedback for self-correction, individual improvement plans shared.</td>
<td>Revision opportunity, knowledge of basis for grades on individual essays, skills transfer.</td>
</tr>
</tbody>
</table>

The Civil Service Reform Act of 1978 tied Merit Pay to a Performance Appraisal System based upon periodic employee review guided by a checklist incorporating critical elements and performance standards for jobs. An analogy can be drawn between job review and essay evaluation for grades. Both systems depend upon timely feedback and involvement in the evaluation process: formative evaluation.
APPENDIX B. Sample Assignment and Revision/Grading Sheets

Process/Instructions: Explanation to a Child, 8-12

**Purpose:** To explain how something works or happens, or how to make or do something in an understandable and enjoyable way.

**Audience:** A child aged 8-12 whom you know well. If not, you will have to interview him/her to get acquainted quickly.

**Procedure:** Be clear whether you are writing a how-to paper (instructions) or a how-it-works paper (explanation). Pay attention to the differences. Do you want the child to go through the process or merely understand it?

**In-class:** You have written a profile of a child, about his personality and interests (or hers). Your topic will be selected especially for him/her: explain what you know she would like to learn. If she is "into" a hobby, take her from where she is now to a new level. You must provide new information, even if you have to go to the library to get it. (The children's non-fiction section can help.) Be sure to provide enough background information for a young child.

**Caution:** If this is a how-to paper, the best way to teach anything is by demonstration. Resist the urge to show him or draw a diagram. You must convey the information in writing only. **DO NOT READ ANY PART OF THE PAPER TO HIM/HER.**

**Write** your first draft. Let the child read it alone. If you read it to him, you will find yourself explaining as you go. Tell him to ask questions, write them down, and answer them afterward, so that you will know what and how to revise.

**Report:** The Process paper is addressed to the child. Write a report addressed to me to tell me his questions, comments, and reactions. You will have given him the answers orally and discussed the parts he didn't understand or given more information about the parts he liked. Write down all the changes (revisions). To check his understanding, ask questions, or in the case of a game, play it with him. Write all this down in your report. Include the following:

1. How did he like your explanation? What words or ideas did he have trouble understanding? Did he already know most of it or was it too vague (not enough detail)?

2. Were you satisfied with his reaction? If you meant to amuse him, did he laugh? In the case of no reaction, or "It was good," you must ask him questions that cannot be answered yes or no.

**Note:** The report must be submitted with the child profile and the Process paper, because none of your other readers will be children. The characteristics of this child (in-class profile) and reactions to your paper (as reported by you) are my guidelines for grading. The paper must match the interests of the child. You will not know how to improve it until he has read and reacted to it.

**Submit:** Child Profile--Reaction Report--First Draft--Final Copy
Performance

Revision Checklist/Grading Sheet: Process Explanation to a Child

Name_________________________ENGL 101 Section_______Late_____

Check one: first draft___ final Copy___ If first draft (for revision workshop) list names of your peer reviewers________________

Instructions: The high level of performance + for 2 points is described in each of the categories. Quality levels are + for 2 points, for 1 point, or - for zero. In those categories marked 0, you must circle the critical components that apply, or make a notation at the bottom of the sheet of areas that need attention. DO NOT write on the paper itself. See scoring guide for conversion of points to letter grades. Remember, you are not grading your classmate's final paper—I will do that, noting whether your draft grade provided guidance for revision.

1. **Purpose**: Intent to instruct (how to) or explain indicated in introduction. Setting or circumstances, necessary conditions clearly given.  
   Quality___ Points___

2. **Word Choice**: Precision in vocabulary, diction, usage, appropriate tone/level of formality for audience, terms defined.  
   Quality___ Points___

3. **Chronological Order**: Steps in process separated, described in proper sequence. Short sentences & paragraphs unified by transitional words.  
   Quality___ Points___

4. **Conclusion**: Outcome or desired result clearly described, whether for instructional or explanatory paper.  
   Quality___ Points___

5. **Sentence Structure**: Proper punctuation to avoid run-ons & fragments, no misplaced modifiers, proper parallel structure, use of either statement or command form consistently.  
   Quality___ Points___

6. **Mechanics/Grammar**: Correctness in spelling, Capitalization, subject/verb agreement, pronoun case, verb tense, apostrophe use, essay format.  
   Quality___ Points___

   Total Points___
   Letter Grade___

Additional Comments:
Assignment #3: EDITORIAL with CATEGORICAL PROPOSITION ARGUMENT

Preparation: Topic Summary

You have been collecting articles on a current news event that interests you as a concerned citizen. (It need not be related to your major field of study.) Normally, when an item has been on the front page for a while, an editorial is written, or at least comments begin to come in in the form of Letters to the Editor. Collect these as well. If you have found an editorial on the day you began collecting news articles, go back a few days to the start of the series of events. (Public library has back issues of newspapers.) Write a summary of the news articles only. Do not include information from the editorial page in your summary.

Editorial Format

Note differences in organization and objectivity between editorial writer and news reporter. Editorial has title instead of headline, organization in climactic order (least to most important), summary of news event at the beginning, and thesis statement and concluding "call for action." News report has 5 w's first, then background and/or quotes from eyewitnesses or authorities, sometimes historical or other comparisons at the end.

Categorical Proposition

Read chapters in your textbook and study the examples of CP's. A CP argument is a natural one: we classify or categorize an individual item to show that a situation exists, and then develop an opinion or thesis statement or proposition to determine the nature of the situation (what is it?) In order to convince others of the validity of the assertion we have made (conclusion in logical terms), we must gather evidence to support the statements leading up to our conclusion (premises) and also identify our assumptions—which we expect the readers to share without proof. Besides verifiable facts, we need to define our terms (what do we mean by ...), sometimes both the subject and predicate of our CP.

Topic Refinement

The topic should be suggested by the news articles; however, the current event is only an example of the whole situation or problem which you will be analyzing. Often, words like "issue" or "affair" will appear, which indicates controversy. The difference between reporting and editorializing is that you must go beyond what has already happened to project a trend and predict possibilities for the future. The emphasis in this assignment is on careful definition and categorizing and gathering evidence to support your assertions.

-INTRODUCTION: Condense the summary, and then write a thesis in CP form. Show that the news points to an existing situation, not an isolated event. (Ex.: The sum of all the shooting incidents in DC lately is violence and the mood to accept and expect more and more of it, not merely scattered killings by a few maniacs or betrayed lovers.)
-BODY: Define all your terms, showing how the current events are part (or 1 example) of the larger issue. Find other evidence from the past &/or from other places about similar events that fall into the same category. You will be quoting from and discussing these other sources and documenting your references. Ex. Jones was upset by this development, but others were not. "It isn't very important, as the media have exaggerated it" (Smith, 1964). However, Jones (1965) remained unconvinced. Note that you still must tell "how you know" even when not directly quoting. Using these outside sources to supplement your own knowledge is crucial in research.

-CONCLUSION: Although most editorials close with a "call to action," some recommendation to be done NOW, your editorial will end with only consciousness-raising. You want to be sure that your readers will agree that "a situation exists" and that you have answered the questions, "Why do you say that?" (definition and category) and "How do you know?" (current news articles and other sources). Your research will not be complete at this point, only begun, as the final three papers will continue to analyze the issue which you have identified here. Please resist the urge to include causal analysis, evaluation, or proposal here. Check with your textbook for lines of demarcation.

Submit: Grading Sheet with 4-line heading upper right corner. Copy only capitalized words in grading points below.

GRADING POINTS

1. THESIS: opinion or viewpoint on arguable issue
2. VALIDITY OF CATEGORIZATION (CP statement)/LOGICAL ORDER
3. SUPPORT: SUMMARY/EVIDENCE from sources
4. DEFINITION of subject & predicate terms
5. WORD CHOICE/Definition of Terms
6. CORRECTNESS: grammar, mechanics, sentence structure, punctuation
Subject: REQUEST FOR PROPOSAL ON RESEARCH PROJECT

Proposals, whether solicited or unsolicited, generate a great deal of work in the business world. They are the means through which projects are initiated, funds allocated, and programs begun. The proposal becomes a contractual agreement about the work to be done: scope, methodology, time frame, work schedule, and about the end result: feasibility study, recommendation report, or suitability study (recommendation for change).

PURPOSES: From your study of a current journal article, your preliminary reading, oral presentation to the class, and personal consultation with your instructor, you will
1. Identify a research problem and describe its importance to non-majors.
2. Demonstrate the feasibility of your study and your ability to complete it in the allotted time (credentials, work schedule).
3. Show how you plan to find a solution (methodology) and anticipate objections.
4. Develop a reading chain from references in original journal article, call numbers of "foundation reading" (books and textbooks by recognized authors), and important journals in the field.
5. Include "non-print" media in your library search.
6. Decide whether to conduct two "expert" interviews (letter of request for each) OR an opinion survey (questionnaire) involving 20 subjects (sample population). For a survey, the questionnaire consisting of 5 multiple-choice questions must be included.

AUDIENCES: The proposal will be made in a memorandum form addressed to your instructor. DO NOT identify your instructor or classmates as the "expert" audience, but as readers with little background who need terms and information translated. In a real-life situation, your proposal would be directed to someone with authority (or funds) to take action on your recommendations. The final Letter of Transmittal will address your paper to that individual.

PROCEDURE: Refer to your textbook on Proposals and Memoranda; Oral Presentation feedback.
1. Define the problem: "A situation exists calling for action." Remember that the topic of a journal article is not a suitable problem for your own research (too specialized). With audience in mind, explain the significance of the problem (why it needs a solution; why it is important). Relate the problem to classmates' interests by explaining how it might affect their daily lives. Use both feedback from oral presentations and your own preliminary reading. Use analogies (familiar examples) to show this connection. Unless your topic changes a lot, this will become the INTRODUCTION of your final paper. Write a complete problem statement, including possible causes, so that the reason for your study is clear. Anticipate classmates' disinterest, and supply background to involve them.

2. Prepare in bibliographic form (MLA style) your preliminary reading list. Simply use the alphabetized notecards. Your subsequent reading should develop as a chain suggested by prior sources, NOT as a library catalogue or index search on a pre-determined topic. Indicate other "leads" you intend to follow, both theory (foundation) and application (research).
REQUEST FOR PROPOSAL ENGL 393 Fall 1981

3. Demonstrate feasibility of your study by defining the scope and methodology to be followed. Break the project into tasks such as data-gathering (searching and recording sources on notecards), organizing material (coding notecards), studying and drafting the paper in sections, and finally, revising, typing, and proofreading.

a) Make a work and time schedule for each phase of your research by setting dates according to your work load in other classes, your job, and personal obligations. You must adhere to due dates for #5 The Proposal, #6, The Progress Report, and #7, The Final Project. For example, if you decide to conduct two interviews, you will need to make appointments by letter. Even if you conduct a survey, you have to schedule administration of the questionnaire and produce the typed questions. This decision must be made in the proposal.

b) Discuss your credentials: major field of study, pertinent coursework and perhaps a set of class notes, your qualifications to conduct this study. If the topic is outside your major, show how you became interested and what "foundation reading" (textbooks or major works) you intend to do.

/SPECIFICATIONS/ Each section of the proposal memorandum report will begin on a separate page with a heading in caps and centered. Double-space Intro. & Feasibility.

INTRODUCTION. Written from your own experience before research. In explaining the problem and its importance to non-majors, you need not have proceeded very far in your reading chain. The Intro. has no references. It should be a problem of current concern (that's why you were to browse in current journals). This will be Intro. of your final paper, unless your topic changes radically.

METHODOLOGY: Develop Preliminary Reading list from the tops of your notecards. Give full bibliographic info (MLA style). Indicate leads from oral presentations, references still to be followed (not full bibliography). State whether you will conduct 2 in-depth interviews (Name, Position, Organization, Location, Date) or survey of 20 subjects (describe subjects by common characteristics, demographics).

FEASIBILITY. Give credentials (see above 3b), work and time frame (3a) answers to questions from oral presentation (to show suitability of topic for audience).

GRADING POINTS
1. Identification of Problem/Causes
2. Background for Audience
3. Translation of Technical Terms
4. Bibliography Form (MLA)
5. Feasibility: Scope, Credentials, Sched.
6. Format/Headings/Mechanics of English