To study the ways that new ways of sharing responsibility for teacher quality affect grading discourses and practices, more than 900 surveys were collected from student/intern teachers and cooperating teachers and supervisors over a 9-year period. Opinions about pass/fail grading for student/intern teaching were tabulated and analyzed. Additional documentary material and a survey were also reviewed. Differences among participants' reported attitudes and institutional practices are discussed. Analyses indicate that pass/fail grading is viewed in a positive light by a minority of respondents across types and programs and over time. In the 2002 survey, a majority of supervisors responding reported use of pass/fail grading and indicated support for rubrics and standards. Overall, secondary-level teachers and supervisors appeared to have slightly stronger preferences for grading systems traditionally used by colleges and universities. Respondents of all levels commented on the uses of traditional grades in communicating relative qualities of teaching. Appendixes contain the student teacher and supervisor interview surveys with results. (Contains 4 tables and 17 references.) (SLD)
Transforming Supervisory Grading:
Accountability in Supervising and Evaluating Beginning Teachers

Paper to be presented at
Session 57.012 of the
Annual Meeting of the
American Educational Research Association
Chicago
April 24, 2003

by

Henry St. Maurice
&
Amber Yudchitz

University of Wisconsin – Stevens Point
464 College of Professional Studies
Stevens Point WI 54481-3897

http://www.uwsp.edu/education/hstmauri

hstmauri@uwsp.edu

Abstract: Over 900 surveys were collected from student / intern teachers and cooperating teachers and supervisors over a nine-year period. Opinions about pass/fail grading for student / intern teaching were tabulated and analyzed. Additional documentary material and a survey were also reviewed. Differences among participants’ reported attitudes and institutional practices are discussed.
Transforming Supervisory Grading:
Accountability in Supervising and Evaluating Beginning Teachers

The purpose of this presentation is to respond to one question: how will new ways of sharing responsibility for teacher quality affect grading discourses and practices that have endured for decades? There are two aspects to this response: first, a brief report on a study of the history of grading academic performance, focusing on practices by university-based supervisors at one institution of teacher education; second, a description of attitudes reported by students, supervisors and cooperating teachers about changes in grading practices resulting from mandates for performance-based assessment.

Introduction

It is evident that mandated standards and performance-based assessments in k-12 schools have transformed grading discourses and practices in those institutions (Guskey 1996, Marzano 2000). Among many reasons, this transformation has occurred because it is impossible to assess multiple standards with one grade, especially an average. Not only does an average conflate incommensurable data, but no such grade can show differential rates of progress toward mastery of various content and process standards (Guskey 2002). By analogy, St. Peter of religious lore does not judge souls on their average scores for obeying the ten commandments. Nevertheless, a single grade – an average customarily summarized by one of five of the first six letters of the alphabet - has become a durable commonplace (Kirschenbaum et al. 1971, Kohn 2002, Riley et al. 1994). In 2003, almost every American ascribes significant academic and even moral meanings to grades such as A- or B+, although it is apparent that they are invalid measures of standards-based performance tasks.

It is also evident that assessment and evaluation of teaching is being transformed by the standards movement of the past two decades (Darling-Hammond 2001, Zeichner 2001). Nevertheless, in most institutions, grading by supervisors for beginning teachers has yet to fully address standards-based assessments. As an example of this situation, one institution enrolling over one thousand prospective teachers as part of a large state university system was studied in two ways: first, a history of grading by supervisors was conducted, showing that various grading systems were used over a century, including: percentages; five of the first six alphabet letters (A-F); and pass / fail (p/f). Second, survey data collected from supervisors over a ten-year period was tabulated and analyzed, showing trends in response to grading systems. These historical and opinion data are here summarized, along with a third survey of supervisors’ responses to a new grading system aligned with state standards.

Background: A Brief Genealogy of Grades

Most grading systems are artifacts of so-called Progressive educational reforms, which began in the USA in the 1870s (Cremin 1988, Kliebard 1995), as schooling discourses and practices broke away from traditional humanistic models of teaching and learning. Progressive reformers claimed that their discourses and practices were based on scientific principles, particularly those of psychology, rather than so-called classical principles. As a leading Progressive psychologist, Edward Thorndike declared in 1918,
Whatever exists at all exists in some amount. To know it thoroughly involves knowing its quantity as well as its quality. Education is concerned with changes in human beings; a change is a difference between two conditions; each of these conditions is known to us only by the products produced by it - things made, words spoken, acts performed, and the like. To measure any of these products means to define its amount in some way so that competent persons will know how large it is, with some precision, and that this knowledge may be recorded and used. This is the general Credo of those who, in the last decade, have been busy trying to extend and improve measurements of educational products. (in Cremin 1988, p. 234)

Progressive educators declared that objective measures of intelligence, personality and skill would more fairly distribute the benefits of education. When these declarations were implemented in policy and practice, however, inequities did not diminish (Gould 1996, Lemann 1999). Nonetheless, progressivist discourses and practices proliferated and have endured. Throughout the world, nearly everyone in school gets grades.

Until the nineteenth century, educational assessments had always been in narrative form, such as essays, recitations, proclamations, debates, theses, dissertations and oral defenses (Barzun 1968, Curti et al. 1949). Evaluations of these narratives were also in narrative form. How they changed is a long story, but major changes evidently began in Europe and North America over a short period of time. In the mid-nineteenth century, railways, ship lines and telegraph networks coalesced into something new in human affairs: a network of social, political and economic systems larger than any empire. Within a single decade from 1840 to 1850, a patchwork of local markets merged into a global web tending to form monopolies (Hobsbawm 1975, pp. 48 ff.). As information, goods and services sped around the globe at unprecedented speeds, uniform standards emerged for quantities (such as time, distance, size, weight), as well as qualities. Environmental historian William Cronon (1991, pp. 114 ff.) has described the origins of quality grading in Chicago in the 1850s: as elevators commingled separate crops, farmers were given receipts showing quantity in pounds and quality on a three-part scale. These receipts themselves became commodities to be traded; soon, derivative instruments such as futures contracts were traded as well. In a world of mass marketing and standardization, anything that could be measured was graded and traded. If a wagonload of wheat could be assigned a grade, so could a can of milk or a portfolio of bonds. It was inevitable that the progressivist penchant for scientific management led to measuring and grading education (Callahan 1964, Tichi 1987).

Progressivist ideas and methods were incorporated in the institution under study when it was founded in 1894. From the outset, all courses were graded by percentages until 1936, when letter grades were introduced. Pass/fail grades were first approved in 1969. In 1991, grades for student teaching were changed from letters to pass/fail for the following majors: early childhood, elementary, exceptional, family & consumer, mathematics and computer science, physical education, health, and foreign languages. The following majors have retained letter grades for student teaching: art education, biology, English, and music. In keeping with national data (AACTE 1990, St. Maurice 2001), of about 1,000 grades posted annually for student teaching during the past decade, over 90% have been pass-fail. Of all grades, over 98% have been either “pass” or “A.”
**Method**

To show current uses of and opinions about grades, two surveys were conducted according to generally accepted survey research methodology (Jaeger 1997). The first was part of an exit interview (appendix 1) distributed every semester since 1994 by supervisors and returned voluntarily and anonymously. All responses \( n = 931 \) were tabulated by years, respondent types, programs, grading scales and opinions of pass/fail grading, yielding a total of eighteen variables (table 1).

<table>
<thead>
<tr>
<th>Item</th>
<th>Response Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year</td>
<td>9 = 1994 to 2002</td>
</tr>
<tr>
<td>Respondent Type</td>
<td>3 = student teacher (st), cooperating teacher (ct), university supervisor (us)</td>
</tr>
<tr>
<td>Program</td>
<td>4 = elementary, secondary, exceptional, other</td>
</tr>
<tr>
<td>Grading</td>
<td>2 = pass/fail or letter</td>
</tr>
<tr>
<td>Opinion of pass/fail</td>
<td>2 = positive or negative</td>
</tr>
</tbody>
</table>

Table 1. Response categories on 1994-2002 exit surveys

The second survey was distributed in late 2002 to over 60 supervisors via the Internet (appendix 2). All responses \( n = 34 \) were tabulated by programs, grading scales, rubrics, and opinions of rubrics and standards, yielding a total of sixteen variables (table 2).

<table>
<thead>
<tr>
<th>Item</th>
<th>Response Categories</th>
</tr>
</thead>
<tbody>
<tr>
<td>Program</td>
<td>5 = kindergarten, elementary, secondary, exceptional, other</td>
</tr>
<tr>
<td>Grading</td>
<td>2 = pass/fail or letter</td>
</tr>
<tr>
<td>Rubric</td>
<td>2 = y or n</td>
</tr>
<tr>
<td>Standards</td>
<td>2 = y or n</td>
</tr>
<tr>
<td>Rubric scale</td>
<td>3 = three-part, four-part, other</td>
</tr>
<tr>
<td>Opinion of rubric</td>
<td>2 = positive or negative</td>
</tr>
</tbody>
</table>

Table 2. Response categories on 2002 supervisors’ survey

Data was analyzed by ANOVA, subject to the following queries: in what groups of variables were significant variations among exit survey responses over nine years? Which respondents’ reported types were related to their opinions of pass/fail grading? Which respondents’ reported program areas were related to their opinions of pass/fail grading? Finally, were supervisors’ reported opinions of pass / fail grades in exit survey data related to responses to rubrics in a recent survey?

**Results**

Survey responses were reviewed and rated by two independent observers. Inter-rater reliability was 100%. There were no statements that defied categorization. Comments scored as positive included “I like it. Basically, it’s saying you either have what it takes or you don’t. Everyone improves with time and practice.” Comments scored as negative included, “There is no in-between. Some students may not necessarily be as strong as others but still have areas to work on. It doesn’t warrant an outright fail.”
On exit interviews surveys across all types and programs for all years, 38.4% (n = 358/931) responses were scored as positive toward pass/fail grading for student or intern teaching (table 3).

<table>
<thead>
<tr>
<th>Type</th>
<th>n</th>
<th>% favoring p/f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooperating Teacher</td>
<td>472</td>
<td>38.8%</td>
</tr>
<tr>
<td>Student Teacher</td>
<td>384</td>
<td>39.1%</td>
</tr>
<tr>
<td>University Supervisor</td>
<td>75</td>
<td>33.3%</td>
</tr>
</tbody>
</table>

Table 3. Percent positive responses on 1994-2002 exit interview items about pass/fail grading, by type.

ANOVA of reported types, programs and years show no significant variations. There was insignificant upward variation in positive opinion scores among kindergarten student teachers and cooperating teachers (n = 107), a group that has had exclusively pass/fail grading longer than any other at the institution under study.

Likewise, there were no significant variations among programs on the 2002 supervisor survey (n=38). Responses show general use and support for evaluation by rubrics based on state-mandated standards (table 4).

<table>
<thead>
<tr>
<th>pass/fail</th>
<th>rubric</th>
<th>rubric &amp; standards</th>
</tr>
</thead>
<tbody>
<tr>
<td>66% use</td>
<td>35% use</td>
<td>82% support</td>
</tr>
</tbody>
</table>

Table 4. Percentage of responses on 2002 supervisor survey.

**Discussion**

Analyses of these data indicate that pass/fail grading is viewed positively by a minority of respondents across types and programs and over time. No significant variations among these variables seem to occur. A possible shift in attitudes appears in the 2002 survey, which shows majorities of surveyed supervisors reporting use of pass/fail grading and stating support for rubrics and standards.

Classroom assessment has been transformed in the past decade by the implementation of standards and associated criteria for performance tasks, but the transformation process has only begun to affect assessments of teaching, especially by cooperating teachers and university supervisors. There appear to be strong attachments to cumulative grades and averages by letters and numbers among all teachers at all levels of instruction. Secondary-level teachers and supervisors appear to have slightly stronger preferences for grading systems traditionally used by colleges and universities. Respondents all levels commented on the uses of traditional grades in communicating relative qualities of teaching.

**Implications**

Two centuries of accumulated practice have embedded deep meanings into the letters and numbers of the academic grading system most used by secondary and tertiary schools, propagated by media and enshrined in folklore. The onset of mandated standards, criteria and rubrics has created a conflict of values as intense as any since the end of World War II, when post-secondary education became more widely available than ever before. This study shows evidence of conflict at one institution between its participants' attitudes and its institutional practices. One implication is that continued longitudinal studies are warranted to track ongoing changes in attitudes and practices. Another implication is that more comparative and general studies are warranted, espe-
cially among states and countries where standards implementation is transforming teacher education and schooling. One further implication might be that supervisors in leadership roles at k-12 schools as well as at institutions of teacher education need advocacy for standards-based assessment of performance tasks.

Conclusions

Standards-based assessment came to k-12 schools as educators faced limitations on traditional systems for grading learning and certifying performance. Despite similar limitations and after more than a decade of implementation efforts by policymakers, standards-based assessments at the tertiary level have yet to render obsolete cumulative grades by instructional supervisors for initial teachers, particularly not for the minority of student and intern teachers whose assignments are still letter-graded.

This study is part of a process of implementing standards-based accountability at one institution among thousands in the nation. Its findings indicate that difficult processes of transformation lie ahead for these institutions, at the end of which most supervisors will have reconstructed their ideas about assessment and evaluation. It is to be hoped that these reconstructions will involve ongoing analyses and continuous inquiry into the origins, purposes and methods of evaluation. Concerted efforts to foster authentic assessment and educative evaluation could lead to improved teaching and learning.
References


Appendix 1: Exit Interview Survey & Results, 1994 – 2002

University of Wisconsin - Stevens Point Office of Field Experiences
Exit Questionnaire for Student Teacher ___ or Intern ___

INSTRUCTIONS: Please give answers to the items below to help evaluate these field experiences. Additional comments are welcome; please use the reverse side of this sheet. All responses will be kept confidential, in accordance with UWSP rules governing program evaluations.

1. District in which this assignment took place (optional):

2. Grade level(s): **173 elementary, 107 secondary, 78 exceptional**

3. Subject(s):

4. Class size(s):

5. Amount of contact with cooperating teacher per day as a %:

6. Were there written Statements of Expectations? If so, please comment on their strong and weak points for you.

7. How many visits were there from the university supervisor in this assignment?

8. Please comment on the supervision process, noting both strengths and weaknesses.

9. Was this assignment graded by letter or pass-fail? **337 p/f**

10. Please comment on the grading system used for this assignment, noting both strengths and weaknesses. **150 positive**

11. Were there regular evaluation conferences prior to the final one? If so, how often were they?

12. Was the evaluation form used, or was a letter substituted for it?

13. Please comment on the evaluation process, noting both strengths and weaknesses.

14. Please comment on the process by which your assignment was made by the Office of Field Experiences, noting both strengths and weaknesses.

15. Please comment about the length of this assignment, noting its duration (6, 9, or 18 weeks).

16. Please comment on the process by which you were oriented to this assignment by the cooperating school, noting both strengths and weaknesses.

17. Please comment about your pedagogical methods courses, noting any that you feel were most or least useful in this assignment.

18. Please comment about your cooperating teacher, noting both strengths and weaknesses.

19. Please comment about your cooperating school, noting both strengths and weaknesses.

20. Please add any additional comments you may have for university or school administrators about your student teaching or internship.
University of Wisconsin - Stevens Point Office of Field Experiences
Exit Questionnaire for Cooperating Teacher

INSTRUCTIONS: Please give answers to the items below to help evaluate these field experiences. Additional comments are welcome; please use the reverse side of this sheet. All responses will be kept confidential, in accordance with UWSP rules governing program evaluations.

Student Teacher ___ or Intern ___ Assignment

1. District in which this assignment took place (optional):

2. Grade level(s): 197 elementary, 150 secondary, 40 exceptional

3. Subject(s):

4. Class size(s):

5. Amount of contact with student teacher or intern per day as a %:

6. Were there written Statements of Expectations? If so, please comment on their strong and weak points for you.
7. How many visits were there from the university supervisor in this assignment?

8. Please comment on the university supervision process, noting both strengths and weaknesses.

9. Was this assignment graded by letter or pass-fail? 414 p/f

10. Please comment on the grading system used for this assignment, noting both strengths and weaknesses.

11. Were there regular evaluation conferences prior to the final one? If so, how many?

12. Was the evaluation form used, or was a letter substituted for it?

13. Please comment on the evaluation process, noting both strengths and weaknesses.

14. Please comment on the process by which this assignment was made by the UWSP Office of Field Experiences, noting both strengths and weaknesses.

15. Please comment on the process by which you were oriented to this assignment by the cooperating school, noting both strengths and weaknesses.

16. Please comment about the manner with which this student teacher or intern began this assignment, noting both strengths and weaknesses.

17. Please comment about the manner with which this student teacher or intern ended this assignment, noting both strengths and weaknesses.

18. Please comment about your perceptions of the student teacher or intern's pedagogical and content preparation for this assignment, noting both strengths and weaknesses.

19. Please comment about the length of this assignment, noting its duration (6, 9, or 18 weeks).

20. Please add any additional comments you may have for university or school administrators about student teaching or internship assignments.
INSTRUCTIONS: Please give answers to the items below to help evaluate these field experiences. Additional comments are welcome; please use the reverse side of this sheet. All responses will be kept confidential, in accordance with UWSP rules governing program evaluations.

1. District in which this assignment took place (optional):

2. Grade level(s):

3. Subject(s):

4. Class size(s):

5. Number of supervisory visits to student teacher or intern:

6. Were there written Statements of Expectations? If so, please comment on their strong and weak points for you.

7. Were your supervisory visits scheduled in advance? Please comment about the load and schedule that you followed.

8. In how many three-way conferences did you participate? Please comment about them, noting both strengths and weaknesses.

9. Please comment on the supervision process in general, noting both strengths and weaknesses.

10. Was this assignment graded by letter or pass-fail? 64 p/f

11. Please comment on the grading system used for this assignment, noting both strengths and weaknesses. 25 positive

12. Were there regular evaluation conferences prior to the final one? How much, if any, did you participate in all of them, including the final one?

13. Was the evaluation form used, or was a narrative on letterhead substituted for it?

14. Please comment on the evaluation process, noting both strengths and weaknesses.

15. Please comment on the process by which this assignment was made by the UWSP Office of Field Experiences, noting both strengths and weaknesses.

16. Please comment about your perceptions of the student teacher or intern's pedagogical and content preparation for this assignment, noting both strengths and weaknesses.

17. Please comment about the length of this assignment, noting its duration (6, 9, or 18 weeks).

18. Please comment about the cooperating teacher, noting both strengths and weaknesses.

19. Please comment about the cooperating school, noting both strengths and weaknesses.

20. Please add any additional comments you may have for university or school administrators about student teaching or internship assignments.
INSTRUCTIONS: Please give answers to the items below to help evaluate these field experiences. Additional comments are welcome; please use the reverse side of this sheet. All responses will be kept confidential, in accordance with UWSP rules governing program evaluations.

Number of Student or Intern Teachers in your building:

Building & District in which these assignments took place (optional):
Grade level(s) or subject(s):

Please comment on the process by which these assignments were made by the UWSP Office of Field Experiences, noting both strengths and weaknesses.
Please comment on the process by which student or intern teachers were oriented to the cooperating building, noting both strengths and weaknesses.

Please comment on the university supervision process, noting both strengths and weaknesses.

Did you observe any student or intern teachers? If so, please state how many observations and whether you provided written feedback, formal evaluations, or letters of reference.

Please give your opinions the quality of curricular and instructional preparation of UWSP student or intern teachers.

Please add any additional comments you may have for university administrators about student teaching or internship.
Appendix 2. Supervisor Survey & Results 2002

Please give us information and opinions as a supervisor of student & intern teaching. Thanks.

1. Program type: I supervise student & intern teachers in (check all that apply):
   - early childhood (13/34)
   - elementary (21/34)
   - exceptional (11/34)
   - middle - any major (24/34)
   - secondary – any (22/34)

2. I grade student & intern teachers on the following scale (check all that apply)
   - pass/fail (25/38)
   - letter A-F (7/38)
   - both (6/38)

3. I also use a rubric in evaluating student & intern teachers:
   - Y (12/34)
   - N (22/34)

4. If Y to #3, my rubric is based on the Wisconsin Teaching Standards:
   - Y (5/12)
   - N (7/12)

5. My rubric has the following scale:
   - Unacceptable/ Acceptable / Exemplary (10/12)
   - Unsatisfactory/Basic/Proficient/Distinguished (1/12)
   - Other (please specify): (1/12)

6. Whatever I do now, I support a grading scale based on a rubric and WTS:
   - Y (23/28)
   - N (5/28)

7. Comments:
**Reproduction Release**

*(Specific Document)*

### I. DOCUMENT IDENTIFICATION:

| Title: Transforming Supervisory Grading: Accountability in Supervising and Evaluating Beginning Teachers |
| Author(s): Henry St. Maurice & Amber Yudchitz |
| Corporate Source: University of Wisconsin – Stevens Point |
| Publication Date: 4/24/2003 |

### II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, Resources in Education (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic media, and sold through the ERIC Document Reproduction Service (EDRS). Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following three options and sign in the indicated space following.

<table>
<thead>
<tr>
<th>Level 1</th>
<th>Level 2A</th>
<th>Level 2B</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="Sample1.png" alt="Sample" /></td>
<td><img src="Sample2.png" alt="Sample" /></td>
<td><img src="Sample3.png" alt="Sample" /></td>
</tr>
</tbody>
</table>

- **The sample sticker shown below will be affixed to all Level 1 documents**
- **The sample sticker shown below will be affixed to all Level 2A documents**
- **The sample sticker shown below will be affixed to all Level 2B documents**

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Check here for Level 1 release, permitting reproduction and dissemination in microfiche or other ERIC archival media (e.g., electronic) and paper copy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE, AND IN ELECTRONIC MEDIA FOR ERIC COLLECTION SUBSCRIBERS ONLY, HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Check here for Level 2A release, permitting reproduction and dissemination in microfiche and in electronic media for ERIC archival collection subscribers only.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Check here for Level 2B release, permitting reproduction and dissemination in microfiche only.
Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but no box is checked, documents will be processed at Level 1.

I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche, or electronic media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries.

Signature: 

Organization/Address: 464 College of Professional Studies
Stevens Point WI 54481-3897

Printed Name/Position/Title: Henry St. Maurice

Telephone: 715 346 2449 Fax: 715 346 2549
E-mail Address: hstmauri@uwsp.edu

Date: 4/24/2003

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:

Address:

Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant this reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:

Address:

V. WHERE TO SEND THIS FORM: