ED 454 286   TM 033 030

AUTHOR   Anderson, John O.
TITLE   Investigating the Evaluation of Student Achievement.
PUB DATE   1997-00-00
NOTE   9p.; In: Connections '97. Proceedings of a Faculty
                  Conference (3rd, Victoria, British Columbia, Canada, May
                  1997); see ED 438 144.
AVAILABLE FROM   For full text: http://www.educ.uvic.ca/connections.
PUB TYPE   Reports - Research (143) -- Speeches/Meeting Papers (150)
EDRS PRICE   MF01/PC01 Plus Postage.
DESCRIPTORS   *Academic Achievement; Elementary Secondary Education;
                  Evaluation Methods; Foreign Countries; Grades (Scholastic);
                  *Grading; *Portfolio Assessment; Simulation; *Student
                  Teachers; Teacher Attitudes; Teacher Education
IDENTIFIERS   Canada

ABSTRACT
          A set of constructed portfolios containing achievement
products and background information for a simulated student was used to
investigate how teachers formulate the evaluation of achievement of their
students. The materials, developed for a simulated student named "Chris,"
were presented to 147 student teachers who graded the 8 product components of
the portfolio over a 12-week period and reported a final grade for "Chris."
Results of regression analyses suggest that the contents of the portfolio
accounted for approximately 63% of the variance for a final grade of "Chris,"
leaning more than one-third of the variance unaccounted for. These findings
indicate that student teachers seemed to use some aspect of "Chris" in
grading that was not actually part of the portfolio. Previous studies of
these data suggested that the student teachers thought that "Chris" improved
through the 12 weeks, even though portfolios were structured in such a way
that only about one-third of the portfolios actually showed improvement,
perhaps reflecting the expectation that a student would improve over the
marking period. Results highlight the complexity of grading practices. (SLD)

Investigating the Evaluation of Student Achievement

John O. Anderson

*The evaluation of student achievement is a significant compo-
nent of classrooms and schools. A set of constructed portfolios
containing achievement products and background information
for a simulated student was used in this study to investigate how
teachers formulate the evaluation of achievement of their stu-
dents.*

Introduction

The evaluation of student achievement is a significant component of classrooms
and schools. Teachers in British Columbia, for example, have to report on student
achievement to parents at least three times each year. The completion of tests,
assignments, projects, journals and portfolios for evaluation purposes are typical
student activities within the classroom. However, the ways in which teachers ex-
amine and translate student products such as essays and test responses into marks
or grades is not well researched.

The study reported in this paper was part of a larger collaborative research project
involving colleagues from both Queen's University and the University of Victoria
that is designed to investigate how teachers formulate the evaluation of achieve-
ment of their students. The dataset for this study was developed by Wilson and
Shulha of Queen's University (Wilson, 1996) who created a set of portfolios con-
taining achievement products (such as written assignments and tests) and back-
ground information for a simulated student called *Chris* in a grade 8 language arts
curriculum. The contents of the portfolios were controlled in terms of achievement
level of products and the background of the student. This manipulation resulted in
a number of different students called *Chris*.

As part of an undergraduate teacher education course in classroom assessment,
147 student teachers graded the components of a portfolio over a 12 week period
and reported a final grade for the *Chris* assigned to them at the end of the term.
These scores and grades generated by the student teachers were the basis for an
investigation of the processes and structures developed by these student teachers
as they evaluated student achievement.

The Data

The data consisted of the scores and grades generated by the 147 student teachers for the components of a student portfolio they were given. Over a 12 week period new components were periodically added. A total of eight different student achievement products were included in the portfolio. Information about the background of the student was also provided. Most of the products were presented in bundles of three and all three were included in the portfolio. One of them was identified as belonging to *Chris*. For example, the same three copies of the writing assignment called *A Trip to the Mall* were included in all of the portfolios, one at a high level of achievement, one at medium and one low level of achievement. One of the copies was identified as belonging to *Chris*, the other two were from other students. *Chris'* assignment was either of a high, medium or low level of achievement depending on how the given portfolio was structured. Two of the products (*Final Examination* and *School Dance*) were identical in all portfolios. The achievement products included in the portfolio were:

1. *A Trip to the Mall*: A written piece that had a maximum score of 25 (three levels).
2. *Salmon for Simon*: A multiple-choice item test of reading comprehension that had a maximum score of 9 (three levels).
3. *Did I Order an Elephant?*: A *Cloze*-format test of reading comprehension with a maximum score of 15 (three levels).
4. *New Kid on the Block*: A short-answer, open-ended format assignment on reading that had a maximum score of 18 (three levels).
5. *Ghost Ship of Mohone Bay*: A multiple-choice format test of reading with a maximum score of 9 (three levels).
6. *Mending Wall*: This was a writing and editing assignment completed on the computer to represent a student's *best work* with a maximum score of 25 (three levels).
7. *School Dance*: A written piece with a maximum score of 25 (single level).
8. *Final Examination*: This was a mixed format test (matching, identification, multiple-choice and short answer) with a maximum score of 130 (single level).

Background information on *Chris* was presented in the form of notes, memos and school reports, and included:

A. *Expectation*: Expectations were to be inferred from information on student scores on the *Canadian Tests of Basic Skills* and parental occupational status. There were three levels of expectation: *low, medium* and *high*.

B. *Growth*: Achievement reports from other school areas such as Mathematics, Science, History and Resource Centre reports suggested that *Chris* was either *falling behind*, performing *steadily* or *improving*.

3

C. *Parental Involvement*: Parental involvement was to be inferred from
school memos and notes related to parental involvement with school
activities such as parent interviews or volunteer work. There were
two levels of involvement: low or high.

D. *Sex*: This could be inferred from an audio tape of *Chris* reading a
passage of text for miscue analysis.

The Analysis

The descriptive statistics for the scores awarded the different assignments and tests,
and the final grade on the report card are summarized in Table 1. Of particular
interest is the range of scores given to the *Final Examination* and the assignment
on the *School Dance*, both of which were identical products common to all 147
portfolios. *Final Examination* scores had a standard deviation of 6.7 with a range
of scores of 79 to 122, and *School Dance* had a standard deviation of 2.1 with a
range of 15 to 24. This suggests that even though the student teachers were all
evaluating the same piece of work, there existed another element that caused
different results to be generated.

The correlations between scores (Table 2) show generally positive, low to moder-
ate linear relationships between scores. There appears to be no single achieve-
ment product dominating the final grade (*Report Card*) for *Chris*, although all but
A Salmon for Simon show moderate, positive correlations with the final grade.
However, the test Salmon *for Simon* has negative correlations with three other
products (*A Trip to the Mall, Ghost Ship of Mohone Bay* and *Mending Wall*) and a
near zero relationship to the final grade on the report card. There is nothing obvi-
ous in the data that suggests an explanation for this puzzling result.

Table 1   Summary Statistics for Portfolio Contents

Score Source	Mean	Standard Deviation	Maximum	Range Minimum
A Trip to the Mall	18.8	2.75	25	13
Salmon for Simon	6.0	2.19	10	1
Did I Order an Elephant?	12.4	1.43	15	8
New Kid on the Block	14.3	2.53	18	7
Ghost Ship of Mohone Bay	5.4	2.15	9	2
Mending Wall	19.7	3.34	25	6
School Dance	20.2	2.09	24	15
Final Examination	100.3	6.70	122	79
Report Card	77.2	4.78	90	65

Table 2   Correlations Between Scores

Score Source	1	2	3	4	5	6	7	8
1. A Trip to the Mall								
2. Salmon for Simon	-.33							
3. Did I Order an Elephant?	.09	.28						
4. New Kid on the Block	.16	.04	.11					
5. Ghost Ship Mohone Bay	.25	-.59	-.03	.03				
6. Mending Wall	.34	-.38	.14	.13	.57			
7. School Dance	.31	.02	.14	.20	-.15	.01		
8. Final Examination	.34	.03	.04	.19	-.03	.14	.28	
9. Report Card	.50	-.07	.26	.38	.37	.49	.36	.57

Table 3   Regression with Achievement Products

REPORT = CONSTANT + TRIPMALL + SALMON + ELEPHANT + DANCE + NEWKID + GHOST + MENDWALL + EXAM

N: 147   Multiple R: 0.787 $R^2 = 0.619$

Effect	Coefficient	Std Error	p (2 tail)
CONSTANT	25.77	4.01	0.000
TRIPMALL	0.55	0.11	0.000
SALMON	0.71	0.16	0.000
ELEPHANT	0.17	0.20	0.375
DANCE	0.58	0.13	0.000
NEWKID	0.42	0.10	0.000
GHOST	0.79	0.17	0.000
MENDWALL	0.36	0.10	0.000
EXAM	0.06	0.02	0.003

Analysis of Variance

Source	DF	Mean-Square	F-Ratio	p
Regression	8	254.9	28.0	0.000
Residual	138	9.1		

The *Report* was the final grade assigned *Chris* by the student teacher. The student teachers were not given explicit instructions as to how to derive the final grade. It was assumed that this grade would be based upon the contents of the portfolio, that is, the marks the student teachers assigned to each achievement product. However, it was also thought that the background information could influence the decisions of the student teachers in regard to *Chris'* final grade. To investigate the relationship between the final grade (*Report*) and achievement products and background information, regression analyses were conducted. One analysis used the achievement products as the predictors of *Report*, another analysis used background information as the predictor, and a final analysis used all data as predictors of *Report*.

The regression of achievement products on *Report* (Table 3) was significant and accounted for 62% of the variance in *Report*. All achievement products with the exception of *Did I Order an Elephant?* had significant regression weights.

The regression of background information on *Report* (Table 4) was significant, but accounted for only 11% of the variance in *Report*. Only *Expectations* and *Growth* had significant regression weights.

Table 4   Regression with Background Variables

REPORT = CONSTANT+EXPECT+GROWTH+INTEREST+SEX

N: 147   Multiple R: 0.336 $R^2 = 0.113$

Effect	Coefficient	Std Error	p(2 tail)
CONSTANT	73.35	2.02	0.000
EXPECT	1.28	0.46	0.006
GROWTH	1.41	0.46	0.002
INTEREST	0.09	0.75	0.905
SEX	-0.36	0.75	0.634

Analysis of Variance

Source	DF	Mean-Square	F-Ratio	p
Regression	4	93.2	4.5	0.002
Residual	142	20.1		

The regression of both achievement products and the background information on Report (Table 5) accounted for 63% of the variance in *Report*. In this part of the analysis, none of the background variables had significant regression weights and the achievement product. *Did I Order and Elephant?* had a non-significant regression weight which was a similar situation when achievement products alone were regressed on *Report*.

Discussion

The results of the regression analyses suggest that the contents of the portfolio account for approximately 63% of the variance of a *Chris'* final grade *(Report)*, leaving over a third of the variance unaccounted for. This finding raises the question: What were other sources that student teachers used to develop a final grade for *Chris* ? The numerical data do not provide an answer to this important question. Both Wilson (1996) and Shulha (1996) have also studied these data and their work suggests some possible answers to this question. As part of the study, all student teachers were asked if they thought their *Chris* was showing improvement over the 12 week period of the study. Wilson observed that all student teachers rated their Chris as improving, whereas the portfolios were structured such that only about one third of the Chris' showed improving performance in terms of the information that was included in the portfolio. About one third of the portfolios were structured to have declining performance, and about third had steady performance. This suggests that indeed the student teachers were observing something about their *Chris* that was not actually part of the portfolio. Perhaps it was simply due to the expectations that educators bring to the classroom that all children will improve as a result of schooling, and this expectation was realized in their perceptions of *Chris'* performance over the 12 week period of the study. Wilson also found some interaction between the sex of the student teacher and the sex of *Chris*. These findings suggest that the student teachers were bringing some expectations, preconceptions or some such other predetermined perceptions into the decision processes used to derive a final grade for *Chris*.

Shulha (1996) analyzed qualitative data that was generated through questionnaires and interviews with the student teachers during the 12 weeks of the study. Her findings indicate that student teachers did not simply aggregate scores on the achievement products to calculate a final grade. Most student teachers commented that grading is a complex task, and many noted that they did not have enough information in the portfolio to determine a final achievement status for Chris. Many noted that student *effort*, and some indication of change and improvement are essential elements in evaluating students. Without access to some knowledge of student effort, the student teachers felt unable to come up with a final grade. One student teacher comment that seems to capture the frustration some student teachers felt and that would be quite disconcerting for most educational measurement

specialists is :

> *I hate to admit it but the mark really came from the grades earned on the assignments.*
> *This did not show development. This showed how Chris averaged out.*

Table 5   Regression with Achievement Products and Background Varables

REPORT = CONSTANT + EXPECT + GROWTH + INTEREST + SEX + TRIPMALL + SALMON + ELEPHANT + DANCE + NEWKID + GHOST + MENDWALL + EXAM

N: 147   Multiple R: 0.792 $R^2 = 0.628$

Effect	Coefficient	Std Error	p(2 tail)
CONSTANT	27.77	4.42	0.000
EXPECT	0.32	0.33	0.339
GROWTH	-0.57	0.83	0.496
INTEREST	0.17	0.51	0.732
SEX	-0.67	0.51	0.191
TRIPMALL	0.56	0.11	0.000
SALMON	0.66	0.19	0.001
ELEPHANT	0.09	0.20	0.663
DANCE	0.53	0.15	0.000
NEWKID	0.42	0.11	0.000
GHOST	0.90	0.25	0.000
MENDWALL	0.40	0.10	0.000
EXAM	0.06	0.02	0.002

Analysis of Variance

Source	DF	Mean-Square	F-Ratio	p
Regression	14	172.3	18.8	0.000
Residual	134	9.2		

Conclusion

These initial findings are not completely unexpected. Most educators and students are aware of the subjective, idiosyncratic components of evaluation and grading. Evaluation of student achievement within the classroom is not a science to the extent practiced within the context of large-scale assessments such as final examinations or selection testing. However, classroom evaluation is the most frequent and pervasive assessment students are exposed to over their years of formal education. Investigating the ways in which this assessment is conducted should and can serve to improve practice over the long term.

This study has provided some interesting and useful insight into how the process of student evaluation works and has developed a useful approach to the investigation. The development of structured portfolios can provide an excellent base upon which to study how teachers evaluate students in the classroom. This study is but one part of a larger collaborative project that will continue this investigation of the processes involved in the evaluation of student achievement.

References

Shulha, L.M., (1996). *Understanding pre-service teachers' thinking about assessment: A window into the world of classroom assessment practice.* A paper presented at the annual meeting of the Canadian Educational Researchers' Association, Brock University, June, 1996.

Wilson, R.J. (1996). *Factors affecting the assessment of student achievement.* A paper presented at the annual meeting of the Canadian Educational Researchers' Association, Brock University, June, 1996.

X-Sender: voelkels@aelvis.ael.org
Date: Tue, 18 Jan 2000 16:15:10 -0500
To: anderson@uvic.ca
From: voelkels@ael.org (Susan Voelkel)
Subject: Connections papers

TM033030

RE: Investigating the Evaluation of Student Achievement (Connections '97)

Recently, the ERIC Clearinghouse on Rural Education and Small Schools
(ERIC/CRESS) was given permission by the University of Victoria's Faculty
of Education to include proceedings of the Connections conferences in the
ERIC database. When processing proceedings, we usually select some of the
papers to go into the database individually. Normally, the proceedings
release would also cover the inclusion of such analytics. But in this
case, each individual piece is also copyrighted to its contributor. We
have selected the above paper. If you would like this paper to have an
individual entry in ERIC, please sign the release attached below and return
it as soon as possible by mail or fax. If you do not want to sign the
release, please let me know and I'll remove your paper from processing.
Thanks for your help.

Susan Voelkel
Processing Coordinator
ERIC/CRESS
P.O. Box 1348
Charleston, WV 25325-1348
USA
Fax: 304-347-0467

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)
ERIC REPRODUCTION RELEASE

I. Document Identification:

   Title: Investigating the Evaluation of Student Achievement

   Author: John O. Anderson

   Corporate Source:

   Publication Date:

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 the
release form.

   Level 1 - Permitting reproduction and dissemination in microfiche or
other ERIC archival media (e.g. electronic) and paper copy.

   Level 2A - Permitting reproduction and dissemination in microfiche and
in electronic media for ERIC archival collection subscribers only.

   Level 2B - 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.

Sign Here: "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."
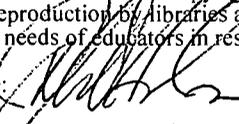
Signature: _____ Position:

Printed Name: _John Anderson_ Organization: _University of Victoria_

Address: _Faculty of Education, University of Victoria, Victoria BC V8W 3N4_ Telephone No: _250-721/7798_

Date: _Jan. 18/00_

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 per copy: Quantity 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 complete the following:

Name:

Address:

V. Attach this form to the document being submitted and send both to:

ERIC Clearinghouse on Rural Education and Small Schools
P.O. Box 1348
1031 Quarrier Street
Charleston, WV 25325-1348

Phone and electronic mail numbers:

800/624-9120 (Clearinghouse toll-free number)
304/347-0467 (Clearinghouse FAX number)
mitchelv@ael.org