

STUDENT PERFORMANCE WITH AND ATTITUDES TOWARD ELECTRONIC
DISTRIBUTED ASSESSMENT IN FIRST-YEAR COMPOSITION CLASSES

ANNETTE CHRISTINE ARRIGUCCI

Department of English

APPROVED:

Beth Brunk-Chavez, Ph.D., chair

Kate Mangelsdorf, Ph.D.

Kerrie Kephart, Ph.D.

Patricia D. Witherspoon, Ph.D.
Dean of the Graduate School

PREVIEW

STUDENT PERFORMANCE WITH AND ATTITUDES TOWARD ELECTRONIC
DISTRIBUTED ASSESSMENT IN FIRST-YEAR COMPOSITION CLASSES

By

ANNETTE CHRISTINE ARRIGUCCI, B.S., M.S.

THESIS

Presented to the Faculty of the Graduate School of the

University of Texas at El Paso

in Partial Fulfillment

of the Requirements

for the Degree of

MASTER OF ARTS

Department of English

THE UNIVERSITY OF TEXAS AT EL PASO

December 2008

UMI Number: FI Î FFI €Ä

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

PREVIEW

UMI[®]

UMI Microform FI Î FFI €Ä

Copyright 2009 by ProQuest LLC

All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

TABLE OF CONTENTS

ACKNOWLEDGMENTS	iv
ABSTRACT	v
1. INTRODUCTION.....	1
2. REVIEW OF LITERATURE	10
3. PROCEDURES.....	30
4. RESULTS	36
5. CONCLUSION.....	64
LIST OF REFERENCES	75
APPENDIX A: SURVEYS AND FOCUS GROUP QUESTIONS	79
APPENDIX B: INTRODUCTION TO REDESIGNED ENGLISH 1312 COURSE	84
APPENDIX C: EVALUATOR TRAINING SCHEDULE.....	85
APPENDIX D: ENGL 1312 COURSE ASSIGNMENTS OVERVIEW	86
APPENDIX E: EXAMPLE EVALUATOR COMMENTING GUIDELINES	87
APPENDIX F: LITERATURE REVIEW AND RESEARCH REPORT ASSIGNMENT.....	88
APPENDIX G: LITERATURE REVIEW AND RESEARCH REPORT RUBRIC.....	92
CURRICULUM VITA.....	94

ACKNOWLEDGMENTS

I would like to thank my mother for all her support during the process of completing this project. I am eternally indebted to Dr. Beth Brunk-Chavez, the greatest advisor a student could ask for. Thank you for being a positive and helpful mentor all the way through this process. I would like to thank all the instructors who taught in the English 1312 redesign for their help in collecting the data for this project. Finally, I want to thank the students who participated in the surveys and interviews for this project.

PREVIEW

ABSTRACT

In Fall 2008, UTEP's composition department implemented a pilot program to test a redesign of English 1312, their second-semester freshman composition course. In addition to a redesigned curriculum, a system of electronic distributed assessment was implemented in ten sections of English 1312. Instead of the traditional format of a class where instructors grade all student assignments, a group of teaching assistants graded student writing anonymously using standardized grading rubrics. The system, which has been used at Texas Tech University since 2002, was put in place at UTEP in order to enhance efficiency and consistency in the teaching of this course.

Previous research on electronic distributed assessment has focused on instructor and administrator perspectives on the grading system; comparatively little research has been done on student attitudes toward and performance with electronic distributed assessment. To measure this, two electronic surveys about the grading system were administered to students in the redesigned sections of the course. Student dropout rates and final grades in the redesigned sections of the course were compared with data from traditionally taught sections of the course. Finally, data from group interviews was collected to determine student attitudes toward the electronic distributed assessment system used in the Fall 2008 semester.

1. INTRODUCTION

In the fall semester of 2008, a pilot program was started at the University of Texas at El Paso (UTEP) composition department. Ten of 23 sections of UTEP's second semester composition courses, required for all entering students, were introduced as redesigned courses.

The redesigned course was dramatically different from the course taught in previous years. Traditionally, the course, titled Research and Critical Writing, entailed the teaching of the process of writing an 8-12 page research paper. Students would analyze arguments, prepare an annotated bibliography, write a persuasive essay, and at the end of the course complete a long research paper. The course was similar in format to many other universities' second-semester composition courses. The majority of classes were taught in the traditional face-to-face manner, though UTEP did offer several sections of online hybrid courses (courses that met once a week face-to-face, typically in a computer classroom, and online the rest of the time) in the past few years.

While a comprehensive overview of the changes to the 1312 curriculum is beyond the scope of this study, it is important to note the context for changes to the system of assessment. Through a grant funded by the Texas Higher Education Coordinating Board, a team comprised of professors Kate Mangelsdorf, director of the University Writing Center, and Beth Brunk-Chavez, director of First-Year Composition, and six experienced composition lecturers designed the changes to the content and delivery of the 1312 courses. They began the redesign work in January of 2008 and met regularly through the summer to discuss this redesign.

Although the purpose of the grant was to redesign the delivery of the course, the pilot instructors agreed that a curriculum redesign was long overdue to make the course more relevant to students' lives and allow students to develop skills that they could use in other courses and in

the workplace. Anne Beaufort's book *College Writing and Beyond* was influential in instructors' changes to the assignments for the course. Beaufort emphasizes the need to make students aware of genres and of audience in their writing for composition courses, in order to help students adapt their writing to the various discourse communities they will encounter throughout their college education (9-12). In addition, instructors wanted to make the course more helpful to students entering the workforce. Clay Spinuzzi describes how the work structures that characterize the Information Age are very different from those in a past era of "relatively stable, rationalized, modular work structures" ("Technical" 266). He writes that the knowledge work that is required of 21st century workers demands "different sorts of texts, and it also demands different ways of thinking about how those texts are produced, received, and managed" (Spinuzzi, "What" 3).

In view of this, the long research paper began to seem like a relic, considering the emphasis of shorter formats in writing on the Internet and other media. In addition, the traditional English 1312 curriculum offered few chances to incorporate digital media into projects. In previous years, individual instructors had leeway to offer some projects that incorporated more technology; however, it was not a standard part of the curriculum.

To address these issues, instructors changed the required assignments for the course (see Appendix D to view a list of the major assignments in the course). A discourse community map and a genre analysis assignment were added to address issues brought up by Beaufort. Other assignments required students to use digital media to a much greater degree than before. A group documentary project required students to use filmmaking software such as iMovie or GarageBand. In place of the long research paper, the final project was an advocacy website that included assignments completed throughout the semester.

The committee also decided to look at ways to address nagging problems in the way composition courses are delivered. Mangelsdorf was particularly concerned about the “one size fits all” model of composition. In a grant proposal for the pilot program, she writes, “Although instructors and administrators in the First-Year Composition program are dedicated to student progress, the program structure limits its ability to respond effectively to the wide range of skills students bring to the classroom” (4). In particular at UTEP, many students whose second language is English enroll in composition courses. An instructor would typically find students ranging from those highly proficient in written English to those who struggled to produce a short essay (Mangelsdorf 4). Many instructors found they did not have the time or training to adequately help these students.

To help ESL students as well as other students who typically struggle in composition courses, a major component of the redesign was to direct students who needed extra help to tutoring sessions at the Writing Center, where they could receive help for specific issues that could not be provided in class. All students who received less than a C- on an assignment were required to visit the Writing Center at least twice before the next assignment. The Writing Center would allow students not meeting the expectations of the course to get additional, individualized feedback on their writing.

The committee also looked at other inefficiencies in the way these courses were administered. Classes were capped at 25 students, due primarily to the time-intensive nature of grading writing assignments. There was no standardized curriculum for 1312 beyond a few broad requirements. Each 1312 instructor could have widely-varying grading standards and widely-varying styles of providing feedback on assignments. Lack of a standardized grading rubric did not allow for fair comparison of grades across courses. Mangelsdorf reports that grade inflation

had been a problem in previous semesters, with 29 percent of students receiving As in 2006, and 33 percent receiving failing grades or withdrawing from the course (4).

The lack of standardization was made worse by the fact that the instructors who taught 1312 courses were often graduate students and part-time or even full-time lecturers with no background in the field of composition studies. There was little composition specialists could do to monitor the teaching of the course and the evaluation of assignments. Student experiences in the course could vary widely, with some students taught by instructors with years of experience and others taught by instructors with no previous classroom experience.

In order to allow English 1312 courses to be delivered more uniformly and more efficiently, some dramatic changes were made to the redesigned sections. One major shift was to change the format of the classes in the pilot program to online hybrid classes, where students would meet once a week in a classroom for 80 minutes and then “meet online” the other class day. Part of the motivation behind this change was to use classroom space more efficiently; another motivation was to allow students to develop their online literacy skills.

However, a more dramatic change in the way the course was administered was in the way students were assessed, which is the subject of this study. In the redesigned courses, grading was separate from classroom teaching, with outside evaluators grading most student assignments rather than instructors. The Texas Tech University composition program had had much success (from an administrative standpoint) with “electronic distributed assessment” in their Texas Tech Online Print Integrated Curriculum(TOPIC)/ Interactive Composition Online (ICON) program that was adopted in 2002. Seeing this success, the department decided to implement a similar electronic grading system for the redesigned courses.

While I discuss TOPIC/ICON in more detail in Chapter 2, in brief, TOPIC/ICON is a system that utilizes sophisticated software to separate classroom grading from teaching in order to maximize the efficiency of composition courses. In TOPIC/ICON, all student writing assignments in both first- and second-semester composition classes are submitted to a software program called ICON. Rather than instructors grading student assignments, pools of “document instructors” grade and comment on all assignments, and “classroom instructors” provide classroom instruction to students. A database called TOPIC stores student data and grades (Kemp). See Chapter 2 for a detailed discussion of the benefits and drawbacks of TOPIC/ICON.

The system of assessment implemented in Fall 2008 in UTEP’s English 1312 classes was conceptually based on Texas Tech’s and had the following features: A pool of 13 teaching assistants (master’s students studying literature, creative writing, and rhetoric and writing studies as well as one rhetoric and composition Ph.D. student) were trained as evaluators to grade most student writing assignments. (Instructors graded three assignments and had some input into the students’ final project grades. They also were in charge of students’ participation grades). Students submitted assignments to the online classroom management system WebCT. For each assignment, one or two evaluators were assigned to grade and comment on all submitted assignments for each section. All grading was done anonymously. Evaluators commented on drafts of assignments using the comments feature on Microsoft Word. They graded final submissions of assignments using an online grading rubric with preset comments to assign grades. The evaluators could also add a summary comment to the rubric if they wished. The rubrics were saved as PDF files and provided to students through WebCT. .

To try to ensure consistency in grading, before each grading cycle, evaluators met with the director and the associate director of composition for a norming session to review the

assignment and rubric, read several documents, and apply the rubric together. They met for approximately two hours seven times during the semester. See the appendices for examples of grading rubrics used during the semester, the schedule of meetings to norm papers, and the guidelines for commenting on student assignments. The director and associate director would also randomly check the comments and grades before they were released to students. Students could appeal grades to a committee if they thought an evaluator grade was unfair. In the future, additional quality control features may be added to the system (for example, Texas Tech allows students to rate grader feedback).

Due to technological constraints, in the first semester of implementation, only one evaluator graded each final draft (Texas Tech typically has two graders for each final draft, with an average of the two scores taken as the final grade, to make sure each grade is fair). Also, the process of assigning evaluators to assignments was done manually. To avoid an evaluator getting too familiar with students' writing in one section of the course, administrators assigned evaluators to different sections for every assignment, although the same TA did comment on drafts and grade the final version.

Six instructors taught the ten sections of 1312 in the pilot program. Four instructors were experienced full-time lecturers who each had several years of experience teaching hybrid courses and had helped develop the new 1312 curriculum. One instructor who was expected to teach two sections of the program unexpectedly quit at the beginning of the semester, so two rhetoric and composition Ph.D. students each volunteered to teach one of the sections assigned to this instructor although these two assistant instructors had neither the experience of teaching hybrid courses nor the ability to prepare to teach the new curriculum in advance, the director determined—because of their disciplinary knowledge—that they would be the ideal

replacements. While more experienced instructors would have been preferable to teach the two sections, the sudden switch of instructors was necessary due to the circumstances..

In comparison, one teaching assistant, one assistant instructor, and four part-time instructors taught the thirteen traditional sections of 1312 in Fall 2008.

At the beginning of the Fall 2008 semester, a total of 338 students enrolled in the pilot sections of English 1312 at UTEP (versus 282 who enrolled in the traditionally taught sections). The sections were identified as hybrid courses in UTEP's Fall 2008 class schedule, though students were not aware of the other changes to the course, including the new grading system, until the semester started. However, instructors did make students aware of the changes to the course through a letter from the director of composition, which can be viewed in Appendix B. While not encouraged, students did have the option to switch to traditional sections of the course, though no data is available on how many students did this.

Until a new software system planned for future semesters could be designed and launched, many aspects of the grading process were handled manually in UTEP's program. This is in contrast to Texas Tech's TOPIC/ICON, where most of the grading process is handled by software. However, the system was similar enough to Texas Tech's that it would allow administrators and instructors to determine whether the distributed assessment system was viable to implement in future semesters.

Pilot instructors who planned UTEP's redesigned course saw the following benefits to using the system: Class capacity in the sections could be raised from 25 to 35 due to reduced instructor grading load. Instructors would also be more able to focus on content and pedagogy rather than assessment. From a work distribution perspective, the electronic distributed assessment system would allow administrators to give first-year TAs (who typically could not

teach until earning 18 hours in graduate-level English classes) time to gradually adjust to the composition program as evaluators while being trained to teach students in the classroom.

A standardized set of assignments would allow the course to be taught in a consistent way that was more connected to the goals of the composition program than in previous 1312 classes. In addition, pilot instructors believed uniform procedures for grading carried many benefits. Students would be evaluated based on pre-determined rubrics developed by the composition program rather than those developed by individual instructors. Evaluators would be given extensive training in providing student feedback, reducing instructor feedback that was limited or not useful to students. Grading would also be more consistent across class sections, likely resulting in less grade inflation. The more standardized grading would allow administrators to more easily compare grade data across class sections, which would help them make changes to the program.

In addition, instructors would be able to give students more assignments throughout the semester, and students could receive feedback on those assignments more quickly. With the new system, evaluators could provide feedback in days rather than in a week or more, which is typical grading time for many instructors.

Though adopting electronic distributed assessment clearly has many potential benefits from an administrator's perspective, little has been written about the effects of the system from a pedagogical perspective. Though some research has been published about electronic distributed assessment since TOPIC/ICON's inception in 2002 (which I will discuss in Chapter 2), most of it has been written from the perspective of administrators and faculty. None, to my knowledge, have focused on student response to the system. It is still unclear whether the system improves student writing performance or makes students more successful in composition courses. In

addition, no studies have discussed in detail student satisfaction with electronic distributed assessment.

In this study I will report on student attitudes and performance using electronic distributed assessment. To study this, I will analyze student surveys and interviews as well as dropout rates and final grades.

Here is an overview of the chapters that follow:

In Chapter 2, I will review previous studies on electronic distributed assessment as well as previous research on assessment of writing.

In Chapter 3, I will discuss the methodology used in this study.

In Chapter 4, I will provide a review of the results of the study.

In Chapter 5, I will discuss the implications of the results.

2. REVIEW OF LITERATURE

Composition departments have been dealing with a challenging set of problems for years, in terms of curriculum, delivery, and assessment in first-year composition classes. In this chapter, I will discuss a system I will term “electronic distributed assessment” (EDA) as a measure to address some of these problems. I will begin by addressing issues of course delivery. I will then give a brief overview of writing assessment theory as it relates to this study. Finally, I will discuss research on specific issues relating to how EDA affects students.

Kathleen Blake Yancey has described the model of delivery for composition courses at colleges and universities as a “laboratory model” (“Delivering” 8) or a “one-to-one tutorial model” (“Made” 310). Yancey characterizes the courses taught under this model as a single teacher in a small class with many assignments. Under the tutorial model, compositionists have sought to “reduce class size...to conference with students, to respond vociferously to each student paper, and to understand that in our students’ eyes we are the respondent who matters” (Yancey, “Made” 310).

However, composition programs across the United States, including at the University of Texas at El Paso (UTEP), have struggled with this model. At UTEP, as at a majority of universities across the United States, two semesters of composition courses are required for all entering freshmen, unless they test out of the course or take it at a community college. A great number of sections are required to accommodate students, since maximum class capacities are typically set at between 20 and 25 students.

Due to the number of sections required, schools often struggle with hiring qualified staff to teach these students. Composition departments at large public universities such as UTEP rely heavily on adjunct faculty and teaching assistants, many of whom have little or no background in

composition studies. Adjuncts and teaching assistants typically receive only a semester or two of training in teaching composition before entering the classroom. Often there is not much oversight into how individual teachers develop syllabi and grade assignments.

Recently, institutions have begun to look for systematic reforms in the way composition classes are delivered. Some schools have reformed the requirements for composition. At Purdue University, the traditional two-semester course has been condensed into a four-credit one-semester course (Weiser 36). Many states, including Texas, are allowing students to fulfill composition requirements in high school through growing dual credit programs.

Some schools are tinkering with the tutorial model of composition. Joyce Mangano Neff proposes restructuring composition classes to a “lecture-and-lab” model common in other departments. Large lecture sections of 200 to 300 students, taught by specialists in rhetoric and composition, would be accompanied by lab sections with 20 to 25 students, headed by teaching assistants or instructors (Neff 55). However, models like Neff’s have yet to be widely adopted.

Technology has provided more opportunities for systematic reform of composition programs. At UTEP and many other institutions, composition courses have been offered as online hybrid courses. Of course, one motivation for the move toward hybrid classes is to utilize classroom space more efficiently. However, Brunk-Chavez and Miller note the benefits of the hybrid format, particularly in encouraging collaboration, and describe how the hybrid format is particularly compatible with the type of learning done in first-year composition courses: “[the course’s] process-oriented pedagogy, its emphasis on collaboration, and the kind of learning students generally do in first-year writing; these aspects coincided with current scholarship on teaching online” (“Hybrid” 3).

TOPIC/ICON: A NEW WAY TO DELIVER COMPOSITION

However, even with the use of technology, efforts at changing the laboratory model of teaching composition have been stymied by the time-intensive nature of assessing student texts. Raising course capacities puts an unfair burden on instructors who are expected to grade students' writing assignments. It also reduces the time instructors can spend responding to each student's work. In one of the most radical efforts to make composition classes more efficient, writing program administrators at Texas Tech's composition department attempted to reduce the bottleneck of assessment by separating teaching from assessment. In 2002, the university implemented Texas Tech Online-Print Integrated Curriculum (TOPIC), a sophisticated online software program and database, as part of the Interactive Composition Online (ICON) program to facilitate assessment of student texts in first-year composition courses. By making assessment more efficient, Texas Tech was able to raise class capacities in its composition classes from 25 to 35 students (Rickly 192).

As I described in Chapter 1, the basic premise behind Texas Tech's EDA model is that rather than instructors grading student assignments, as is done under the laboratory model of teaching, pools of "document instructors" grade and comment on all assignments anonymously using standardized rubrics, and "classroom instructors" provide classroom instruction to students using a common curriculum. TOPIC stores student data and grades. In addition, first-year composition courses are taught as hybrid courses, meeting face-to-face once a week for 80 minutes. As mentioned earlier, this reduces the amount of classroom space required for composition classes (Kemp, "TOPIC").

Rebecca Rickly, a writing program administrator at Texas Tech at the time of the switch to EDA, describes the benefits of the change in an essay for *Delivering College Composition*.

Rickly describes how WPAs collaborated to discuss how TOPIC was developed as a means to achieve programmatic goals given the constraints of the writing program. The system has allowed first-year composition to be administered in a way that is in line with departmental goals, such as “more and varied writing experiences,” “timely, helpful feedback,” and “frequent peer and self-critique” (191). Rickly notes that first-year composition is now “fiscally responsible” and more “criteria-driven” (193). She also describes how students now receive more of a standardized education. Criteria for each draft or critique is presented in class and then used by document instructors to assess each draft or critique. Through technology, composition specialists can now easily monitor the quality of assessment of document instructors through the use of technology (Rickly 192).

As I discussed in Chapter 1, UTEP’s composition program has decided to adopt EDA for an overall redesign of its English 1312 course, with the additional component of an enhanced Writing Center, for many of the same reasons as Texas Tech. The model is intended to improve the efficiency and consistency of the program. In a grant proposal to the Texas Higher Education Coordinating Board, Mangelsdorf cites four major reasons for changing to EDA:

- Fewer classroom instructors will be needed because of larger class size.
- Teaching assistants without 18 hours of coursework in the field will be able to fully participate in instruction by grading papers and tutoring.
- Half as many classrooms will be needed.
- Fewer students will need to repeat the course. (5-6)

Like at Texas Tech, it seems a large motivation for the change to EDA is a more efficient use of labor and resources.

However, while the success of electronic distributed assessment in increasing efficiency from an administrative standpoint is clear, pedagogically, EDA remains to be studied in depth. To my knowledge, Texas Tech is the only institution to have adopted EDA for first-year composition courses. While several Texas Tech faculty and students have published research on TOPIC since its inception in 2002, most of the research has focused on administrator and faculty reaction to the system; none to my knowledge has focused on student performance and attitudes under this system of assessment.

STUDENT WRITING ASSESSMENT BACKGROUND

The lack of research about student performance and attitudes under EDA highlights a general lack of research on student writing assessment, particularly classroom assessment. Only recently has student writing assessment been a topic of scholarly research, and much of that research has been focused on developing tests to place students in composition classes.

Recently, some compositionists have called for administrators to carefully consider the effects of assessment on students. Clearly, assessment is an integral part of teaching writing to students and of their progress in the academy. Performance on tests often determines where students are placed, and performance on important class assignments determines whether students pass or fail classes.

In her history of assessment in composition, Kathleen Blake Yancey states that we need to ask the question, “Whose needs does this writing assessment serve?” (“Looking” 498). Yancey sees the need to balance the needs of individuals and of programs in looking at assessment, both programmatic and classroom. Yancey focuses on three major issues: 1) the role that the self should play in any assessment, 2) programmatic assessment, specifically, how assessments can be used to help students, and 3) research, or assessment as “knowledge-making

endeavor” (“Looking” 484-85). She states the need to examine how others understand themselves as a result of “our interpretations, our representations, our assessments” (Yancey, “Looking” 498).

Brian Huot emphasizes the importance of gauging the effects of assessment on stakeholders – administrators as well as teachers and students: “Any use of any writing assessment should be accompanied with a validity argument that addresses technical documentation important to those who work in educational measurement, honors political considerations important to administrative and governmental agencies, and most importantly considers the impact on the educational environment and the consequences for individual students and teachers.” (“(Re)Articulating” 55). While Huot is referring to programmatic assessment in this statement, one can easily see how this type of consideration of stakeholders applies just as well to classroom assessment, particularly in the case of EDA.

While not as widely theorized about as program assessment, classroom writing assessment clearly also has a powerful effect on students. Yancey describes how classroom assessments have the power to shape the students themselves. “[W]riting assessment, because it wields so much power, plays a crucial role in what self, or selves, will be permitted” (“Looking” 498). She describes how agency of students is limited by those who create assignments and the criteria by which it is scored: “[T]he authorship of such a text is likely to be a static, single-voiced self who can only anticipate and fulfill the expert’s expectations” (Yancey, “Looking” 499). Huot notes the power of grades as “evaluative judgment(s) that connect our judgments with specific cultural beliefs and assumptions that affect both group and personal identity beyond our classrooms” (“Toward” 167). De-emphasizing grades, Huot discusses how assessment should be

directly linked to teaching, with students coming away from the classroom being able to assess their own texts, whether graded or not (“Toward” 165).

However, despite the importance of assessment, both Huot and Yancey note writing teachers’ lack of interest in and even opposition to writing assessment in the past half-century. Yancey notes that for a discipline grounded in humanism, assessment is often “still associated with number-crunching and reductionism” (“Looking” 495).

HISTORY OF WRITING ASSESSMENT

At this point I will give a brief history of writing assessment in the past 50 years. While much of the history relates to standardized placement testing procedures, I believe the history is very relevant to the development of recent classroom grading procedures, particularly in the case of EDA, where administrators are seeking to develop standards of grading for classroom writing.

Early assessments of writing were developed by test-makers, not instructors, and were designed more for reliability in placing students in courses than for their validity as measures of writing ability. In the first half of the 20th century, assessment and classroom teaching were generally considered separate spheres. According to Yancey, “in the 1950s, 90 percent of a survey of 100 schools used a placement test, and 44 percent used an exit exam” that was factored into the course grade (“Looking” 489). Many of the exams included no writing at all and were multiple choice reading comprehension tests (Yancey, “Looking” 488). Yancey refers to these assessments, with their emphasis on reliability, as the “first wave” of assessment.

Yancey characterizes teachers as instrumental in urging the development of more valid direct measures of student writing ability. Yancey states that “the last 50 years of writing assessment can be narrativized as the teacher-layperson (often successfully) challenging the psychometric expert, developing and then applying both expertise and theory located not in

psychometrics, but in rhetoric, in reading and hermeneutics, and increasingly, in writing practice” (“Looking” 484).

This resulted in what Yancey calls the “second wave” of assessment that started about 1970, when holistic writing assessment came to be seen as a reliable measure of student writing (“Looking” 484). The Advanced Placement (AP) assessment at was a major milestone in this second wave. The AP assessment included a standardized course curriculums, and the course culminated in a final essay test. The test was notable in that test-makers were able to develop standardized procedures to grade the essays in a reliable way. The test used carefully developed writing prompts, scoring guides, and methods of calculating acceptable rater agreement (Yancey 490). Edward M. White, who helped develop methods for holistic assessment in the 1970s, states that by the early 1980s, a survey showed almost 90 percent used holistic scoring for writing assessment (qtd. in “Holistic” 83). While not as reliable or as inexpensive for administrators as a multiple choice test, Yancey characterizes the holistic assessment of student writing as a victory for classroom instructors who saw the assessment as more relevant to classroom teaching than multiple choice tests.

It was in the second wave of assessment when control of grades generally shifted to classroom teachers, and the holistic scoring model was replicated as the model for classroom grading: “When these same writing teachers returned to their classrooms, they found that their teaching had changed. With newfound confidence in their ability to give consistent and fair grades, they were able to use evaluation as part of their teaching...Some teachers brought scoring guides, sample papers, and peer evaluation directly into their teaching of writing” (“Holistic” 89).