

White Papers: Focus on Teaching and Learning

Reinert Center for Teaching Excellence
Saint Louis University

May 2008

Published by Saint Louis University's Reinert Center for Teaching Excellence
Saint Louis, Missouri

Copyright remains with the authors of individual chapters, except where noted.

Contact the Center of Teaching Excellence at (314) 977-3944 for more information.

Electronic copies available at <http://cte.slu.edu>

Table of Contents

Taking the Dread Out of Dreaded Courses: The Effects of Humor on Learning	1
The Impact of Direct Writing Assessments in Secondary School on Student's Writing in College	7
Why the Rich Get Richer or Why Are High Achieving Students Much More Likely to Seek Academic Assistance Than Low Achieving Students?	33
Students' Perceptions of Professors in Non-Traditional Teaching Roles	41
Call For Papers	52

Introduction

Welcome to the first issue of the Reinert Center for Teaching Excellence's new peer-reviewed publication, *White Papers: Focus on Teaching and Learning*. The White Papers Series is intended as a showcase for articles written by faculty and graduate students at Saint Louis University that focus on issues related to teaching and learning in higher education. The Center plans to publish White Papers twice a year, during the fall and spring semesters. A call for submissions for the fall 2008 edition can be found at the end of the current publication. Paper copies of the White Papers will be made available to faculty and students at Saint Louis University who prefer that format, however, our primary method of disseminating this publication will be through the Center's website (<http://cte.slu.edu>).

This first issue includes four papers. Psychology graduate student, Jana Hackathorn, reviews the literature on the role of humor in teaching and learning and identifies the need for research on the effects of humor on student learning in "dreaded" courses. Psychology graduate student, Jana Hackathorn, reviews the literature on the role of humor in teaching and learning and identifies the need for research on the effects of humor on student learning in "dreaded" courses. Beth Kania-Gosche, graduate student in educational studies, explores how the direct writing assessment format required in standardized exams, such as SAT and ACT, might impact instruction based on writing as process. Shawn Nordell, Ph.D., assistant professor in the department of biology, reports on a study showing that sessions on study skills strategies for university students can increase student achievement, while identifying the need for more research on why university students who need academic assistance the most are least likely to take advantage of such opportunities. Brien Ashdown and Kristin Kiddoo, graduate students in psychology, report on a study exploring whether students considered professors' personal backgrounds as a factor in students' willingness to learn particular course topics.

The Center staff would like to thank the following members of the Center's advisory board who developed criteria for the peer review process, reviewed articles submitted, and made recommendations for articles to include in this first volume: Eddie Clark, Ph.D. (chair), Dan Finucane, Ph.D., Mary Rose Grant, Ph.D., Brian Till, Ph.D. and Elizabeth Zeibig, Ph.D.

Taking the Dread Out of Dreaded Courses: The Effects of Humor on Learning

Jana Hackathorn,
Experimental (Social) Psychology

It is important to realize that in this day and age, transferring knowledge from your brain into your students' brains is not enough. Some students need, desire, or even demand that their classrooms be entertaining. Inevitably, instructors who have the abilities to fulfill this need are rated higher on their end of the semester evaluations than those who are unable to meet the students' expectations (Fortson & Brown, 1998). Teachers, as a result, have even been compared to stage actors. A stage actor must take someone else's words and repetitively, time after time, make them lively and interesting in order to capture the audience's attention (Ziv, 1998). In this same way, a teacher must convey virtually the same information, semester after semester, and sometimes year after year. So, one must wonder if in order to become an effective teacher, if one must not only look at what is taught, but how it is taught.

Effectiveness in teaching and presentation styles has traditionally been a topic of research in the educational and communication domains. Conversely, humor in education, while not a new construct, is a relatively new topic in research, and is also fairly controversial (Downs, Javidi, & Nussbaum, 1988). Many teaching based papers have been written admiring and encouraging the use of humor in the classroom. Normative based research conducted by Downs and colleagues (1988) found that regardless of empirical support, teachers are utilizing relevant course and material related humor as a means of communicating with their students or in the hopes of clarifying concepts. Furthermore, most teachers would agree that humor is a wonderful way to create an informal classroom

environment, to help put students at ease, to increase students' attention, to encourage the students to see the teacher as a human being that they can connect with on a more personal level, and make learning more fun (Schmitz, 2002).

Although, numerous studies have been conducted in support of utilizing humor, the number of true experimental studies conducted regarding the actual academic related effectiveness of humor in the classroom is limited. The restricted amount of research that has addressed the issues involved in utilizing humor has examined students' and teachers' preferences for humor in the classroom (Ziv, 1998). In a study conducted by Fortson and Brown (1998), graduate students were asked to identify one classroom variable that they felt most represented their 'best' and 'worst' professors. Good sense of humor (30%) was listed as one of the top three variables reported by the students as indicative of their 'best' professor. Conversely, lack of sense of humor (11%) was one of the top three descriptions of their 'worst' professor. The findings of this study provide evidence that students want instructors who can inject humor into their lectures.

The studies conducted by Downs and colleagues (1988) provided support that using humor in a lecture might be an effective tool for showing the significance of the material being conveyed. Based on their results, the authors suggest that if the material becomes salient through the use of humor, the students pay closer attention to the material. It is no surprise that humor in the classroom has emerged as an important aspect of teaching. It can serve a dual purpose by reducing conflict and enhancing human relations, between students and teacher, as well as from student to student (Schmitz, 2002). Humor can be nurtured and integrated into the classroom such that it fosters a sense of openness and respect between students and teachers. It has been suggested that humor can add variety to a class, providing a sometimes well needed change in pace, and can even contribute to reducing tension that

may impede the learning process (Schmitz, 2002). Additionally, some research has shown that appropriate humor can be useful in keeping students interested, establishing an environment that is conducive to learning, or fostering relationships between students and teachers (Fortson & Brown, 1998; Gorham & Christophel, 1990; Neelam, Molstad, & Donahue, 1999). A common assumption is that humor enhances teacher-student relationships and thus enhances learning (Gorham & Christophel, 1990). When students feel safe, they can enjoy the learning process and each other (Neelam, et al., 1999). All of the aforementioned consequences to the use of humor in the class can be seen as beneficial to the students.

However, it is important to remember that while humor might be useful in keeping students interested and engaged, whether it is a necessity for the delivery of academic information is still unknown. In the limited research that has been conducted in the area of humor, there is very little consensus on whether or not humor actually has a positive effect, or any effect, on learning objectives. Humor tends to involve multiple factors in order to be injected naturally and effectively. Typically, it takes an experienced instructor, or an experienced comedian, to incorporate effective humor into their teaching style (Fortson & Brown, 1998). It should go without saying that destructive humor, such as humor that ridicules or demeans an ethnic group, is completely inappropriate. Therefore, the type of humor that is used should be selected with care (Schmitz, 2002). If the humor cannot be interjected into the material in an appropriate and helpful manner, it might prove to be more detrimental than beneficial (Fortson & Brown, 1998).

Inappropriate use of humor notwithstanding, it is not an illogical leap to assume that thoughtful, and appropriate, use of humor by instructors can contribute to cognitive related and academic achievement as well. It has been posited that humor might have an arousal

effect, which could produce cognitive based changes in attention and retention, but most research has not been able to support this notion (Gorham & Christophel, 1990).

Nonetheless, some believe that any form of humor, especially humor relevant to the actual presented class information, can break up the monotonous tone of an environment that consists mostly of lecture, which in turn helps in aiding in attention and retention (Forsyth, 2003). In a study and replication, conducted by Ziv (1998), evidence demonstrated that teachers' use of humor did contribute to an increase in student learning. However, further findings revealed that creating humorous stimuli relevant to the course material was absolutely essential to its effectiveness.

Of particular interest and usefulness is the role of humor in class that students usually would not enjoy. Researchers have not yet linked the impact of humor on learning objectives in those classes that are identified as 'dreaded courses'. 'Dreaded courses' are the courses that many students try to avoid. The students may avoid these classes because they lack self-confidence in the content area, they perceive incredible difficulty or foresee profound boredom with the material in the future, or they have had some negative experiences in that content area in their past academic endeavors (Neelam, et al., 1999). A purely hypothetical example of a dreaded course might be Research Methods and Statistics, Physics, Organic Chemistry, or Macro-Economics. Schmitz (2002) reports that humor, if used appropriately, could be especially helpful in reducing students' tension in upper level foreign language and translation classes. While some students may look forward or immensely enjoy these classes, a large portion of students may try to delay or procrastinate taking these courses. Regardless, it is reasonable to expect that humor will have similar, if not exaggerated, results for dreaded classes, whatever they may be, as the results found for the classes in the Ziv (1998) study. Since students perceive learning barriers to be innate in 'dreaded courses,' the

effects of humor on learning objectives could even be greater than in the average courses (Neelam, et al., 1999).

This current paper is merely a call for empirical studies that use control and experimental groups in a school context, to examine the effects of humor in the classroom. An experimental endeavor could provide support for humor as an efficient tool to be utilized in furthering student achievement (Downs, et al., 1988). Once it is decided, perhaps via a pilot study, what the average students believe is a ‘dreaded’ class; there are many questions that could be answered through an empirical investigation of the practical applications of humor. Does humor in the classroom actually influence the cognitive aspects involved in learning? Does humor in ‘dreaded classroom’ lectures positively affect learning objectives? Will moderate use of appropriate humor in dreaded class lectures result in greater students’ attention and retention than dreaded class lectures without the use of humor? Is the amount of humor a feasible variable in the usefulness of humor, indicating a curvilinear relationship?

Future research should explore the effects on students’ relevant learning, achievement, attention and retention. More specifically, it would be informative if a study uses ‘dreaded class’ information in a lecture format, to examine the effect of humor on learning objectives in those classes that students least look forward to taking. It is expected that when students are given information in a humorous way, or when information is accompanied by a related comical or amusing anecdote or clipart, the information presented becomes more salient, than if presented void of humor. It is then posited that the salience provided by the humor, aids the student in retention and learning, while simultaneously creating the benefits of a positive classroom atmosphere that can enhance student-teacher relationships. The appropriate use of humor in a lecture should prove to be beneficial in a multitude of ways in the classroom. Now is the time to empirically support that notion.

References

- Downs, V. C., Javidi, M., & Nussbaum, J. F. (1988). An analysis of teachers' verbal communication within the college classroom: Use of humor, self-disclosure, and narratives. *Communication Education, 37*,127-141.
- Forsyth, D. R. (2003). *The professor's guide to teaching: Psychological principles and practices*. Washington, DC: American Psychological Association.
- Fortson, S. B., & Brown, W. E. (1998). Best and worst university instructors: The opinions of graduate students. *College Student Journal, 32*(4), 572-576.
- Gorham, J. & Christophel, D. M. (1990). The relationship of teachers' use of humor in the classroom to immediacy and student learning. *Communication Education, 39*, 46-62.
- Neelam, K., Molstad, S. & Donahue, R. (1999). Using humor in the college classroom to enhance teaching effectiveness in "dread courses". *College Student Journal, 33*(3), 400-406.
- Pett, M. A. (1997). *Nonparametric statistics for health care research: Statistics for small samples and unusual distributions*. California: Sage Publishing.
- Schmitz, J. R. (2002). Humor as a pedagogical tool in foreign language and translation courses. *Humor, 15*(1), 89-113.
- Ziv, A. (1988). Teaching and learning with humor: Experiment and replication. *Journal of Experimental Education, 57*(1), 5-15.

The Impact of Direct Writing Assessments in Secondary School on Students' Writing in College: An Opinion Paper on Needed Future Research

Beth Kania-Gosche
Educational Studies

Abstract

The content and format of a standardized or state mandated test influences instruction. Throughout high school, students take multiple direct writing assessments, especially now that ACT and SAT have added this section. Proponents of these tests cite the improvement of instruction of writing after implementation, high reliability among raters, and preparation for college writing. However, others regard the product of the tests artificial and not representative of students' true writing abilities, as well as undermining the writing process supported by so many. Instruction focused on creating a polished essay in thirty minutes is completely different than instruction based on writing as a process. More research is needed to examine the long term effects of direct writing assessments through the eyes of students, especially college students who have experienced post-secondary expectations for writing, rather than teachers and policy makers.

Overview

High stakes assessments influence instruction in public schools. The content and format of these tests can have implications for not only classroom assessment but also the method in which certain skills are taught. Writing is one of these skills. Instruction in a classroom where students are expected to write a completed essay in a limited period of time will be completely different from a classroom where students will be assessed with a portfolio, or collection of writings. Students in a classroom where instruction is only focused toward a timed essay will use a different process than those who received time and feedback from instructor and peers.

High school students are taking more standardized tests that require composition under timed conditions, often called “writing on demand” or “direct writing” assessments. Contrary to popular belief, the increase in the use of this format of assessment is not all due to the laws of No Child Left Behind. In February 2005, ACT added an optional Writing Test. The SAT also changed its format to a required essay, called “the most extensive change to that test in a decade, and perhaps the most important change in the test’s history” (NCTE, 2005, p. 3). In the state of Missouri, public high school students take a state test that requires an essay, in addition to requiring essays at earlier grade levels, considered a “high stakes” test because of the potential implications for the school. Advanced Placement exams for history, literature, and writing require an essay. As competition for college scholarships increases due to greater tuition prices, enrollment in Advanced Placement courses also has increased in Missouri by 11% since 1999, according to the state Department of Education (2007). Students could conceivably take six or more standardized writing tests in the four years of high school. In college, students receive much less writing instruction (typically only one or two courses) than they did in high school (usually at least four courses over four

years). The changes in high school instruction due to assessment could have long term impact on students whose main writing instruction occurred in secondary school.

Although the addition of the essay to the ACT and SAT is new, AP and state assessments have long utilized this direct writing format, although the stakes have risen for schools. Do students in college write differently if they experienced these high stakes tests? Do students who came from a public school, where such tests have huge implications for teachers and schools, use a different writing process than those who only took the exams for college credit or admission?

Definitions of Terms

1. Writing on Demand: Used interchangeably but less frequently in the literature than the term “direct writing assessment.” Some state tests use this term rather than direct writing assessment. Students are given an open-ended prompt and expected to respond with a semi-polished essay within a given amount of time, and trained graders score the essays either holistically or analytically using a rubric (Gere, Christenbury, & Sassi, 2005; Jones, Jones, & Hargrove, 2003).
2. Direct Writing Assessment: Used interchangeably but more frequently in the literature than the term “writing on demand,” which has come into use more recently. These tests are more expensive to grade because they require a trained scorer rather than a computer. These types of tests include the optional essay on the ACT, the required essay on the SAT, the essay sections of Advanced Placement exams, and the performance event on the MAP test in Missouri.
3. Indirect Writing Assessment: Multiple choice tests given to students to assess their writing ability. Typically these items focus on grammar, spelling, although some do

address word choice, organization within a larger piece, revision, or addition of words, phrases, or sentences. These items are easy and fast to score because they can be run through a computer. Indirect writing assessments are often paired with direct writing assessments, as on the ACT, SAT, many AP exams, and the MAP test in Missouri.

4. Standardized Testing:

“Standardized tests typically are developed by commercial publishers or government agencies for use in a large number of sites. These tests have several advantages: The items generally are well written, standard conditions of administration and scoring have been established, and tables of norms are provided” (Gall, Gall, & Borg, p. 208).

5. High Stakes Testing:

“Although schools have historically used testing as a measure of student learning, testing has recently moved from being an individual student assessment to a system for ranking and comparing students. . . These high stakes test are also used as a mechanism for public comparisons of teachers, schools, and school systems. . . In some states, low-performing schools can be taken over by the state” (Jones, Jones, & Hargrove, 2003, p. 2).

While ACT or AP tests may have high stakes for the individual student, these scores usually do not have an impact on the schools. High stakes testing most often affects public schools who receive state and federal funding. In this study, the term “high stakes testing” will be used to describe assessments written by individual states to fulfill No Child Left Behind. It does not refer to college entrance exams or college credit exams.

6. State Mandated Testing: While No Child Left Behind requires that states give assessments at specified grade levels and in certain subjects, there is not a national assessment mandated. Therefore, states each write their own (Jones, Jones, & Hargrove, 2003). This term will be used interchangeably with high stakes testing, although not as frequently.

History of Writing Assessment

Why were direct writing assessments implemented? Standardized assessments began as direct writing when Harvard University in 1873 started requiring a written exam for entrance (Greenburg, Wiener, & Donovan, 1986). Multiple choice tests were developed as a more efficient and reliable way to gauge students' content knowledge, and those in the field of writing soon adopted multiple choice tests because of the time and expense of hand scoring essays (Greenburg, Wiener, & Donovan, 1986). Today, indirect assessments of writing are easily and quickly scored by a computer, and thus frequently used. In fact, ACT and SAT still use these measures in addition to the new essay sections. Indirect writing assessments require students to choose the correct spelling, grammar, word choice, transitions, and organizations for sections of passages.

More recently, critiques of multiple choice questions led to tasks that were considered more authentic and representative of the abilities of students. The identification of an error on a multiple choice test did not necessarily mean a student could correct the error in his or her own writing (Spandel & Stiggins, 1997). Actually requiring students to write on the assessment was theorized to promote this practice in classrooms. Higher education has always embraced direct writing rather than indirect, both for determining placement and to determine proficiency. The 1983 National Testing Network in Writing survey responses revealed that "More than two-thirds (69 percent) of the tests given by responding institutions included writing samples: half of these tests (53 percent) are writing samples alone . . . Furthermore, survey respondents also indicated that almost all (97 percent) of their holistically scored writing sample tests are locally developed" (Greenburg, Wiener, & Donova, p. xvi). Therefore, direct writing assessments in high school may prepare students for these same tasks in higher education.

Policy makers wanted to shift focus in the elementary and secondary classrooms away from grammar drills to actual writing (NCTE, 2005). Breland and Gaynor's study comparing the two types of tests "suggests that direct and indirect writing assessments tap similar skills . . . However, indirect measures lack face validity and credibility among members of the English profession" (1979, p. 127). In the journal *Assessing Writing*, Elbow and Yancey debate the value of direct writing assessments, specifically the holistic scoring aspect where readers are asked to agree with each other, an unnatural reaction according to the authors (1994). White, Lutz, and Kumasikiri agree in their book *Assessing Writing: Policies, Practices, and Politics*, which examines issues of validity and reliability in scoring and encourages the use of portfolios (1996).

In an issue of the *English Journal*, influential scholar Arthur Applebee writes "Lessons from the Past" about English language arts assessments. "Once performance has been leveled through the scoring procedure, the scores that resulted proved very useful in comparing groups with one another, but not very useful in specifying how well a particular groups of students was doing in addressing a task" (1994, p. 42). In the article, he also outlines how psychometric properties such as validity and reliability do not always mesh with the curricular concerns of teachers, who may not use the results of the assessments. Mabry writes that rubrics are the issue, not the tests or the prompts. She writes in the *Phi Delta Kappan*, "The standardization of a skill that is fundamentally self-expressive and individualistic obstructs its assessment" (Mabry, 1999, p. 673). The author also writes of the differences between holistic grading on the impression of the essay, analytic grading on different characteristics, and then averaging or scoring. The training teachers receive for this scoring is essential for brainwashing teachers into accepting the test, according to Mabry. Lumley did find in his dissertation that training for holistic scoring was a strong predictor of

attitude towards the tests and use of classroom instruction directly related to skills on the test (2000).

Effects of Standardized Testing

No Child Left Behind legislation has increased the number of mandated standardized tests for students, and many states include writing on demand as a part of these assessments. However, there is a growing backlash to the demands of this legislation and the effects of the testing on the students' motivation. Recently in North Carolina, "a state commission agreed today on a draft report saying "there is too much time spent on testing" and recommended that the "fourth-, seventh-, and 10th grade writing tests . . . be eliminated" (Hui, 2007). The commission did not believe that more testing had led to better instruction and education for the students in North Carolina. Unfortunately, there was no explanation as to why the writing tests were considered the most expendable, although they are the most costly to score, requiring a trained person rather than a computer.

The emphasis on test performance presents a dilemma for high school teachers of writing. Writing on demand trains students to write all at once in a limited time frame. However, writing as a process, the approach supported by the National Council of Teachers of English (NCTE), involves multiple drafts based on feedback from teachers and peers. An essay written in half an hour and an essay written in a week will be very different, even if the prompt is the same. High school teachers want their students to succeed, but success on the test may require different skills than success on classroom assignments. Teachers may have little training in helping students succeed on the writing section of standardized tests, especially since the feedback students receive from scorers is minimal. On the other hand, teachers may be pressured to raise test scores so much that they focus exclusively on the

skills necessary to take writing on demand tests, rather than letting students write as a process.

The Conference on College Composition and Communication, a constituent organization of the National Council of Teachers of English (NCTE), has a position statement on writing assessment (1995). Although this document focuses more on classroom rather than large-scale assessment, the second assumption outlined in this document is that “language is social” and assessment that is isolated conflicts with current research the strategies most writing teachers use. The fourth assumption leads to the following conclusion:

“One piece of writing—even if it is generated under the most desirable conditions can never serve as an indicator of overall literacy. Ideally, such literacy must be assessed by more than one piece of writing, in more than one genre, written on different occasions, for different audiences, and evaluated by multiple readers. This realization has led many institutions and programs across the country to use portfolio assessment” (CCC 1995, p. 432).

Unfortunately, direct writing assessment usually involves one piece of writing that can now have an effect on a student’s admission to a college or university, with the ACT and SAT essay additions. These changes led the NCTE to release a report condemning these additions and questioning the assumption that better writing instruction will result from them (2005).

Yet, there are positive implications to using this format of assessment. The following excerpt, from a book written for teachers titled *Creating Writers: Linking Writing Assessment and Instruction*, demonstrates some of the positive impacts of direct writing tests.

“Large scale writing assessment performs several services. First, it keeps writing in the public eye, reminding everyone of its importance in education. Second, it encourages the development and use of sound scoring criteria, which makes the assessment of writing more valid and consistent. And third, it offers teachers who may work for years in one building or at one grade level an opportunity to gain a broader perspective: seeing several hundred samples of student writing from throughout a state or district is very different from seeing only the writing of your own students.” (Spandel & Stiggins, 1997, p. 28)

Standardized testing of any format, especially those with high stakes attached, is often criticized for the negative effects on students and teachers. Assessments drive instruction in schools where tests matter. Amrein & Berliner list many side effects including increased drop-out rates in high schools and stress on young students (2003). Popham adds cheating scandals to the list of unintended side effects (2000). Little research has been focused on specific formats of assessments and their potential impact. Most literature on standardized testing focuses on the perspectives of teachers and how their instruction has been altered due to the test. Very few recent studies have even examined the perspectives of students toward high stakes testing in general, much less writing assessment.

Validity and Reliability of Direct Writing Assessments

A wide gap exists between teachers of writing and researchers in psychometrics. While direct writing assessments can be reliable with trained graders, teachers argue this does not make them valid. Cindy Johannek in her book *Composing Research: A Contextualist Paradigm for Rhetoric and Composition* believed that this gap is promoted by graduate composition programs that do not require statistics or quantitative research courses. She suggested combining qualitative and quantitative methods by writing in context and using first person while still completing statistical analyses.

Specialists in measurement and psychometrics argue that using test scores for evaluating schools and programs means that scores must be standardized, and reliable (Moss, 1994). With training, an acceptable reliability level can be achieved by multiple scorers of essays in these situations. These reliability levels are much higher than when scorers are asked to rate portfolios, or collections of student writings. Moss notes that “A growing body of evidence indicates that when assessments are visible and have consequences for

individuals or programs, they alter educational practice sending an unequivocal message to teachers and students about what is important to teach and learn” (1994, p. 112). This suggests that students from a private school, where students typically take such tests only for college admissions (ACT or SAT) or college credit (AP), would have different experiences in writing classes than public high school students, where tests typically have effects on the teachers and school more than the individual students. While private school students may take the essay for college entrance or AP, these scores usually do not have an impact on the school directly. Additionally, if writing is emphasized on more tests, teachers should follow by emphasizing more writing instruction. But, what kind of writing is being emphasized? Writing on demand for the tests or writing as a process? What are the consequences of requiring so many of these types of assessments? Lumley found in his dissertation that the state writing assessment in Pennsylvania did focus teachers more on writing, in both secondary and elementary classrooms, although they did not use the test as the foundation of their instruction (2000).

Much literature has been devoted to the topic of inter-rater reliability when scoring essays for standardized assessment. Huot writes about the disconnectedness between teachers of writing, who believe in the process and social aspects of writing, and the measurement community, who prize reliability (1996). Ketter and Pool write a detailed literature review explaining the conceptual framework of direct writing assessments before writing an ethnographic study of three teachers and their classes specifically focused on passing the Maryland Writing Test (2001). This class existed solely to help students who failed the test pass the next time. Can any of the skills learned in this class to pass the test transfer to college or the workplace? This issue was not addressed in the study.

One chapter in the book *Evaluating Writing* outlines the procedures the state of California developed for its state writing assessment, which involved the training and input of thousands of teachers from across the state (Cooper & Odell, 1999). An earlier edition of this book (1977) focuses on persuading the reader of the necessity for moving beyond multiple-choice, standardized assessments to evaluate students' writing abilities. The two completely different collections of articles in these editions demonstrate the change in thinking about writing assessment over the past 25 years.

Teacher Opinions of Direct Writing Assessments

Do teachers think direct writing assessments are helpful? Schneider and Schneider surveyed over 100 fourth grade teachers from a diverse school district with quantitative, Likert scale questions and open-ended, qualitative ones. The authors then focused on nine teachers whose responses about beliefs and instructions did not align, most attributing this difference to standardized state assessment. They found that, generally, although teachers felt a lack of control over the writing curriculum because of the assessments and focused more on the genres asked in the prompts for these tests, teachers felt “that they were modeling writing more often, and their writing instruction had moved beyond the basics despite survey results that revealed that isolated skills, grammar, and workbook instruction were still prevalent” (Schneider & Schneider, 2002, pg. 3344).

Similarly, Hillocks, Jr, another influential researcher in the field of writing assessment, reported the following about writing teachers in Texas.

“61% [of Texas administrators and teachers] believe that the state writing assessment supports the kind of writing program they wanted for their high school; only 27% believe that other kinds of writing should be included in the curriculum, and nearly all of these teachers suggest that creative writing and writing about literary works be included. Forty-seven percent feel that the assessment has helped to improve writing in the schools. Nearly all of those believe that the testing has forced attention to

writing . . . Taken together, these attitudes suggest that teachers and administrators see the testing program as successful in improving writing” (Hillocks, Jr., 2002, p. 86).

These teachers view the writing on demand tests as a positive influence. In addition, the book *Assessment of Writing Policies, Politics, and Practices* notes the important influence that the companies developing these tests have had, claiming that the Educational Testing Service “has produced some of the most important innovations in writing assessment . . . Both holistic scoring of essays and large-scale use of portfolios began at ETS, whose strong domination of the commercial field continues to allow it to seek advances in testing practice” (White, Lutz, & Kamusikiri, p. 19).

However, Zigo paints a darker picture of the effect standardized testing has had on teachers, in her qualitative study of a group of high school teachers in Georgia, where students must pass the state tests to graduate. Many teachers reported that the Five Paragraph Essay was the primary format they taught their students. The author investigated a high school with consistently high passing scores on the writing exam, to analyze the difference in this and schools with similar demographics with lower passing rates. She found three main factors: cooperation among the teachers in the English department, a recognition of the changing nature of reflective practice, and “as teachers became more confident in their professional knowledge, they became more willing to question the tests they administered. The emergence of this critical stance led them to support even their weakest students . . . in recognizing the inherent power issues residing within the very existence of such tests” (Zigo, 2001, p. 221). In addition to a wealth of additional research on this topic, Cooper and Odell also devote an entire chapter to their book to recognizing the cultural differences in language and revising standardized writing test rubrics to reflect this (1999).

Barksdale-Ladd and Thomas examined not only teacher attitudes and perceptions towards testing but also parents'. Teachers shared no positive effects on instruction because of the testing, and parents discussed how stressful the testing experience was. Although a heading under discussion is "Children's response to the tests," the authors did not interview or observe any students directly. Sharing the views of many teachers, Kirby & Kirby write in the chapter on evaluation of their new book on contemporary memoir writing "Students do typically write more often and revise more often in the regular English/language arts classroom if the mandated tests in their district require actual writing . . . but such tests are still often isolated, artificial evaluations" (2007, p. 141). They lament the time wasted preparing for these tests when students could actually be writing for a real audience, not an unseen grader. The dreaded five paragraph essay is often overtaught because it produces an essay that will earn students a passing score (Schwartz, 2004; Penny, 2003). However, this formula can also help struggling writers who lack organization, and it at least requires students to actually write as opposed to answer multiple choice questions to assess their writing ability (Albertson, 2007).

In their book *Writing on Demand*, Gere and Christenby (2005) suggest that direct writing assessments mirrors tasks required in the workplace and the essay tests often required in college. They maintain that quality writing instruction can take place in the context of these assessments. Mielo writes in his article "Rewarding the Slipshod" that writing as a process trains students that they need not contribute much effort to their first draft, since it will be revised multiple times (2007). Yet, in the words of a themed issue of *English Journal*, NCTE's publication for high school English teachers, dedicated exclusively to assessment:

"In effect, standards and high stakes testing have solidified isolated instruction and inauthentic purposes for both reading and writing in classrooms at all grade levels.

These standards and tests have overshadowed decades of research on the most effective and best practices for teaching reading and writing” (Thomas, 2001).

If standardized direct writing assessments are a necessary evil, the voices of the students must be heard to truly produce a clear picture of writing assessment. It may have a farther reaching impact than just the few hours it takes for students to complete the test.

Assessments influence the way students are taught, which may impact the way students write in college and possibly even the rest of their lives.

Scorers’ Perceptions of Direct Writing Assessments

In the book, *Creating Writers*, the authors suggest that many teachers’ attitudes towards direct writing assessment change if they participate in the process as scorers (Spandel & Stiggins, 1997). The researcher has experienced this firsthand as a writer for a state assessment. “Large scale writing assessment looks quite different today from how it looked, say, ten or even five years ago. It is, in many cases, more teacher and student-friendly. It offers teachers a chance to participate as raters and to learn about assessment as they do so.” (Spandel & Stiggins, 1997, p. 24). However, the authors never specify how these new assessments have become more “student-friendly.”

Also at the high school level, Jeffrey Schwartz examines the impact of Advanced Placement English Language and Composition exams in an article in the *English Journal*. Proficient scores on AP exams are accepted by many colleges and universities as course credit, and many high school students take these challenging courses to enhance their transcript. The author also outlines the strict rules and regulations governing scoring of these tests to ensure reliability from reader to reader. As an AP teacher himself, the author reflects on his experience. “The broad goals of the AP exams are lost, however, when teachers teach only to the test or when students take the AP course or exam merely as a form of academic

status. . .The AP English exam preparation is best when it challenges students to think deeply, read closely, and write from a wide range of contexts” (Schwartz, 2004, p. 55). While AP exams have been using direct writing assessments far longer than the ACT and SAT, the students who take them are typically the top of their class, and the scores from these tests do not have an effect on the schools like state mandated tests. The AP exam often drives the content of an Advanced Placement course, proving the point that assessments influence instruction.

Similarly, in an entertaining article in *Assessing Writing*, Penny, a research methodologist, becomes a reader for a company scoring essays for a state assessment (2003). This first person account describes the author’s experience, from handwriting an essay to be qualified as a grader to the rules restricting discussion of the prompt outside of the building. Completely devoid of statistics, this qualitative account reveals the process behind the scoring of writing assessments and who is doing the scoring for the miniscule amount of money offered.

Formulaic Response to Direct Writing Assessment

The scoring guides for writing on demand essays have often been criticized for rewarding formulaic writing, the most popular being the Five Paragraph Essay. Mabry believes that “rubrics have the power to undermine assessment” (1999, p. 673). In a recent article in *Research in the Teaching of English*, Albertson examined over 1,000 essays written in response to a state writing prompt, discovering a relationship between organization and score, although the scores were not higher the more formulaic the structure (2007). Albertson suggests that the five paragraph essay might provide a scaffold for those students who struggle when writing on demand (2007). However, the article cautions that formulaic

writing leads to less elaboration, and thus may hurt better writers. Other noted techniques in high scoring papers were figurative language and humor, more complex strategies than are usually taught for writing on demand prompts.

Penny echoes this in his qualitative article, reporting that no five paragraph essay received the top score, although many received a passing one (2003). Wesley agrees, analyzing her own students' essays on a given prompt in "The Ill Effects of the Five Paragraph Theme" (2000). However, long before high stakes testing was mandated, Emig noted that the Five Paragraph theme should be called the "Fifty Star Theme" because it is so quintessentially American (1971, p. 97). The tests are not to blame for the creation of this organizational structure, although it is still embraced by many teachers as a saving grace for struggling writers.

High School Students' Writing

What types of writing are students doing in high school? Is this writing similar to what is expected in college? Freedman and McLeod (1988) conducted surveys of successful teachers of writing in both the United States and Great Britain, with the samples from the American National Writing Project and a national British organization for teachers of English. The students of these teachers were also surveyed, and the results compared to Applebee's influential 1981 study, which was a random assortment of teachers rather than those designated as "successful." Freedman and McLeod (1988) note that, "The two groups of successful teachers stand apart from Applebee's teachers in that the successful groups place less stress on mechanics and testing and more stress on allowing students to express their feelings" (p. 23). However, it is important to note that this study is over twenty years old, when standardized testing had not yet become as important as it is today.

Prompts can have an effect on students' scores on writing assessments. Johanek reprints Olivier's quantitative article in *Research in the Teaching of English*, interspersed with qualitative data from an interview with the author, an experienced writing teacher (2000). Interestingly, Olivier cites her data as supporting her instincts as a teacher. She found that seventh grade students needed a simple prompt, while ninth graders wrote best to topics of interest to their age group. Eleventh grade essays were actually found, in many cases to be comparable to the college freshmen pieces evaluated, often incorporating self-reflection.

In their textbook, *Teaching Writing in Middle and Secondary School*, Blasingame and Bushman examine the literature on high school students' attitudes toward writing (2005). Several self-report instruments exist for evaluating student attitude toward writing and typing students, although none of these address direct writing assessment although they are still widely used in research (Fagan, Cooper, & Jensen, 1985).

More recently, Scherff and Piazza surveyed high school students in four Florida high schools, with a sample of 1,801 (2005). Students in the sample reported the most frequent writing was response to literature, followed by expository and persuasive essays in those grade levels tested by the state in writing. This is in direct contradiction to Applebee's earlier findings that most writing in high school was informative (1984). The respondents to the survey reacted negatively to the writing focused exclusively on writing test preparation. These authors note that, "It is our strong belief that student opinions and perceptions can contribute relevant and necessary information concerning the status of writing in the English language arts classroom" (Scherff & Piazza, 2005, p. 274). This study is one of the few to analyze the opinions of students rather than merely analyzing their writing.

College Students' Writing

“Much of academic writing in higher education involves the problem of recognizing multiple and conflicting authorities in one’s reading, while working from these to construct in writing a position—a voice—of one’s own” (Dyson & Freedman, 1991, p. 15). ACT and SAT added the essay section of their tests in response to college professors who complained about their students’ writing (NCTE, 2005). However, voice is one of the most difficult items to assess on a scoring guide used by multiple readers. Lavelle and Zuercher (2001) developed an instrument for assessing students’ writing processes but did not equate this with high school experiences.

In another NCTE publication, *Research in the Teaching of English*, Hansen et. al. compared the writing of sophomore students who scored high enough on the AP exam to opt out of first year composition with those who took the course (2006). The researchers had a sample size of 182 students who were given an essay to complete as well as a questionnaire, consisting of self-efficacy items, a writing process questionnaire, questions about high school writing experiences, and the Writing Apprehension scale. In addition, Hansen et al. found that almost 75% of the students felt prepared for college after their high school writing experiences, although barely more than half reported writing a research paper once a year (Hansen et. al., 2006, p. 481). Although memory is usually an inaccurate measure, “Even the students who 6-12 months earlier had taken a FYC [first year composition] course that explicitly aimed to prepare them for future college writing assignments . . . did not perform appreciably better than those who had had an AP experience one, two, or perhaps even three years prior to the study” (Hansen, et. al., 2006, p. 483). This study demonstrates that most students feel prepared for college writing and that high school writing courses may have an even bigger impact than writing courses taken in college.

Beck and Jeffrey believe that the reason so many high school students fail to make the successful transition to college writing is because of the discrepancy between tasks required by direct writing assessments in high school and tasks required of students in college (2007). They examined the writing prompts for three states in not only language arts but other subjects as well. They found inconsistencies internally, from prompt, to scoring guide, to sample papers, especially in Texas (Beck & Jeffrey, 2007). Narratives, often not required beyond first year composition in college, were often rewarded with high scores, even if the prompt seemed to be asking for a different genre of writing.

Recently, several dissertations have focused on the transition of high school writers to the college level or on the differences in perception of teachers and students (West-White, 2001; Maxfield, 1998; Oates, 1999). Most of these are qualitative and involve a case study or ethnographic approach. Through her interviews with eleven secondary students and analysis of their writing, Nielson determined that “Florida’s writing assessment test has a negative influence on students’ perceptions of writing” (2002). However, Nielson’s study only used a small sample and did not focus exclusively on this topic. Willis conducted a similar study in Kentucky, where state assessment involves portfolios in addition to timed writing (2000). However, this dissertation was focused on the creation of the portfolio assessment process in one school. Unfortunately, portfolios are often criticized for lack of reliability in scoring.

Writing Assessment’s Influence on Instruction

In his book *The Testing Trap: How State Writing Assessments Control Learning*, George Hillocks, Jr. examines the standards, prompts, scoring guides, and procedures for several different states and their assessments, including Kentucky, which uses a portfolio

system (2002). Only in Kentucky did Hillocks find a significant difference in teaching style, away from the traditional lecture model to one using a writer's workshop. In addition, only in this state did teachers consistently use each step of the writing process, instead of emphasizing a few, such as prewriting. This state also provided funding for improving writing instruction, especially through its unusually high number of National Writing Project sites.

In addition, Hillocks, Jr. noted that reducing student-to-teacher ratios would do more than mandated tests to improve writing skills, although the assessments have had some positive results. "There is little question that the assessments have made the teaching of writing a more common and extensive activity than Applebee found in 1981" (Hillocks, Jr., 2002, p. 205). This was the reason ACT added the essay in 2005, to improve writing instruction in secondary schools and move away from grammar drills (NCTE).

Needed Research

Although the impact of high stakes testing has been widely discussed in the literature (Amrein & Berliner, 2003; Jones, Jones & Hargrove, 2003), little has been written specifically about the direct writing format and its influence on instruction and students' writing processes. These assessments vary widely from state to state with a few even using portfolios in addition or in place of these tests (Hillocks, Jr. 2002). Even on nationally administered exams, the SAT and ACT have completely different prompts, expectations of genre, and scoring guides (NCTE, 2005). Scoring can be holistic or analytic, and in itself is a source of debate, both among educators and psychometric researchers. The opinions of teachers towards these tests seem to be mixed, and the opinions of students are rarely studied at all.

Currently, 38 states use some type of writing on demand as a standardized test, in accordance with No Child Left Behind, often in combination with multiple choice questions (Ketter & Pool, 2001; Heck & Crislip, 2001). Yet there is little research that investigates the impact of these types of tests on students' future writing skills. In fact, the book *The Unintended Consequences of High-Stakes Testing* devotes only two of its 178 pages to direct writing assessments (Jones, Jones & Hargrove, 2003). Hillocks, Jr. in his book *The Testing Trap: How State Writing Assessments Control Learning* interviews only teachers and examines the relationship between the type of assessment and the type of instruction reported.

High stakes standardized testing in general has long been criticized for its negative impact on students and teachers. Amrein and Berliner examine the impact of standardized testing on factors of student motivation, such as high school dropout rates, by examining several recent studies on this issue (2003). The National Board on Educational Testing and Public Policy completed a 150 page report in March of 2003 examining the effects of standardized testing, finding similar to Amrein and Berliner that high stakes testing in high school increases the drop out rate. This report also noted that in high stakes testing schools, writing teachers were more likely not to use computers in composition because the state tests were handwritten (2003, p. 5). They found that "three quarters of all teachers, regardless of stakes or grade level, found the benefits of the testing program were not worth the time and money involved" (2003, p. 13). However, other researchers have found conflicting evidence (Hillocks, Jr., 2002). Do students agree with this evaluation, especially those who have gone on to attend a university?

The addition of the essay section to the ACT and SAT is relatively new, although state tests have required writing for many years. Most research on the impact of direct writing

assessment has focused on the opinions of teachers, ignoring the voice of the students. What value do students see in writing on demand assessments? Rather than asking high school students, the responses of college students would be more valuable. College students have years of distance from their high school experiences, and they have experienced the postsecondary expectations for writing. Finally, studying college students may demonstrate if there is an impact from these tests at all once students have left secondary school. By taking a sample of a variety of college students, from freshmen to senior, a pattern may emerge as students progress through their college coursework. While some recent research has taken a case study approach or a longitudinal one, following students from high school to college, a quantitative method for a large sample would be a greater addition to the literature.

Conclusion

Janet Emig, in her influential book *The Composing Process of Twelfth Graders* wrote in 1971, “It is an extremely rare situation for writers, particularly student writers, to proceed from initial stimulus to final draft, or revision, without interruption” (40). Now, over thirty years later, the direct writing assessment format asks students to do just that. With more students taking more of these tests than before, this pattern of writing is being emphasized again and again, year after year. If most instruction is geared toward preparation for a direct writing assessment, How does this change students’ writing processes? Most students have only one chance to take a direct writing assessment. While they may take another version of the test at a later date, the ACT for example, the prompt will undoubtedly be different.

Unfortunately, this type of testing is not going away any time soon. Many graduate students must take the GRE, which requires an essay, as does the MCAT for medical students. Education students in the state of Missouri must pass the College Base Academic

Subject Examination or CBASE, which now also requires an essay. Perhaps requiring high school students, and even middle and elementary school students, is just preparing them for the future. Although there are many critics of these tests, there may be unintended consequences, both negative and positive, to teaching students the skill of writing a polished essay in thirty minutes.

Direct writing assessments may be the best compromise between indirect measures of testing writing and portfolios. Students are actually writing instead of filling in bubbles, and yet the raters are reliable enough to satisfy psychometric experts. However, the perspectives of the students are essential to determining if the long term effects of tailoring instruction to fit the test. As Greenburg, Wienever, and Donovan write, “Various tensions exist between the values of composition and the values of assessment, the values of education and the values of policy” (1986, p. 44). Students, however, should be a value shared by all.

References

- Albertson, B.R. (2007). Organization and development features of grade 8 and grade 10 writers: a descriptive study of Delaware students testing program (DSTP) essays. *Research in the Teaching of English* 41(4), 435-64.
- Amrein, A.L. & Berliner, D.C. (2003). The effects of high stakes testing on student motivation and learning. *Educational Leadership* 60(5), 32-8.
- Applebee, A.N. (1994). English language arts assessment: lessons from the past. *The English Journal* 82(4), 40-46.
- Barksdale-Ladd, M.A. & Thomas, K.F. (2000). What’s at stake in high-stakes testing: teachers and parents speak out. *Journal of Teacher Education* 51, 384-397.
- Beck, S.W. & Jeffrey, J.V. (2007) Genres of high stakes writing assessments and the construct of writing competence. *Assessing Writing* 12, 60-79.
- Blasingame, J. & Bushman, J.H. (2005). *Teaching Writing in Middle and Secondary Schools*. Upper Saddle River, NJ: Pearson.
- Breland, H.M. & Gaynor, J.L. (1979). A comparison of direct and indirect assessments of writing skill. *Journal of Education Measurement* 16 (2), 119-128.
- Cooper, C.R. & Odell, L., Eds. (1999). *Evaluating writing: the role of teacher’s knowledge about text, learning, and culture*. Urbana, IL: NCTE.
- Cooper, C.R. & Odell, L. Eds. (1977). *Research on composing: points of departure*. Urbana, IL: NCTE.

- CCCC Committee on Assessment. (1995). Writing assessment: a position statement. *College Composition and Communication* 46 (3), 430-437.
- Dyson, A.H. & Freedman, S.W. (1991). Critical challenges for research on writing and literacy: 1990-1995. Final report to the Office of Educational Research and Improvement. Center for Study of Writing.
- Elbow, P. & Yancey, K.B. (1994). On the nature of holistic scoring: an inquiry composed on email. *Assessing Writing* 1(1), 91-107.
- Emig, J. (1971). *The composing process of twelfth graders*. Urbana, IL: NCTE.
- Fagan, W.T., Cooper, C.R., & Jenson, J.M. (1985). *Measures for Research and Evaluation in the English Language Arts* (2nd ed.). Urbana, IL: ERIC Clearinghouse.
- Freedman, S.W. & McLeod, A. (1988). National surveys of successful teachers of writing and their students: the United Kingdom and the United States. Final report to the Office of Educational Research and Improvement. Center for Study of Writing.
- Gall, M.D., Gall, J.P. & Borg, W.R. (2007). *Educational Research*. 8th Ed. New York: Pearson.
- Gere, A.R., Christenbury, L. & Sassi, K. (2005). *Writing on demand: best practices and strategies for success*. Portsmouth, NH: Heinemann.
- Greenburg, K.L, Wiener, H.S., Donovan, R.A. (1986). *Writing assessment: issues and strategies*. New York: Longman.
- Hansen, K. et. al. (2006). Are advanced placement English and first year college composition equivalent? A comparison of outcomes in the writing of three groups of sophomore college students. *Research in the Teaching of English* 40(4), 461-501.
- Heck, R.H. & Crislip, M. (2001). Direct and indirect writing assessments: examining issues of equity and utility. *Educational Evaluation and Policy Analysis* 23(1), 19- 36.
- Hillocks Jr., G. (2002). *The testing trap: how state writing assessments control learning*. New York: Teachers College Press.
- Hui, T.K. (2007). Too much school testing, panel says. *News & Observer* Nov. 19.
- Huot, B. (1996). Toward a new theory of writing assessment. *College Composition and Communication* 47 (4), 549-566.
- Johanek, C. (2000). *Composing research: a contextualist paradigm for rhetoric and composition*. Logan, UT: Utah State University Press.
- Jones, M.G., Jones, B.D. & Hargrove, T.Y. (2003). *The unintended consequences of high-stakes testing*. Boulder, CO: Rowman & Littlefield.
- Ketter, J. & Pool, J. (2001). Exploring the impact of a high-stakes direct writing assessment in two high school classrooms. *Research in the Teaching of English* 35 (3), 344-393.
- Kirby, D.L. & Kirby, D. (2007). *New directions in teaching memoir: a studio workshop approach*. Portsmouth, NH: Heinemann.
- Lavelle, E. & Zeurcher, N. (2001). The writing approaches of university students [electronic version]. *Higher Education* 42 (3), 373-391.
- Lumley, D.R. (2000). Teacher's perceptions of the impact of the Pennsylvania Writing assessment on classroom writing instruction. [electronic version] *Dissertation Abstracts International* 61,(4), 1241.
- Mabry, L. (1999). Writing to the rubric: lingering effects of traditional standardized testing on direct writing assessment.[electronic version] *Phi Delta Kappan* 80, (9), 673-9.
- Maxfield, B.A.M. (1998). Rhetorical competence and performance: a case study of highschool and college writers. [electronic version]. *Dissertation Abstract International* 59 (8), 2961A.

- Mielo, G. (2007). Writing as a process: rewarding the slipshod? *Education Week*, 27, (3), 26.
- Missouri Department of Elementary and Secondary Education (2007). News release: Missouri's SAT scores improve; popularity of "AP" classes is still growing rapidly. www.dese.mo.gov/news/2007/sat.htm
- Moss, P.A. (1994). Validity in high stakes writing assessment: problems and possibilities. *Assessing Writing* 1, (1), 109-128.
- National Board on Educational Testing and Public Policy (2003). The perceived effects of state-mandated testing programs on teaching and learning: findings from a national survey of teachers. Boston College.
- NCTE (2005). The Impact of the ACT and SAT Timed Writing Tests. http://www.ncte.org/library/files/About_NCTE/Press_Center/SAT/SAT-ACT-tf-report.pdf?source=gs
- Nielson, S.M. (2002). The will to write: students' and teachers' perceptions of writing. [electronic version]. *Dissertation Abstract International* 64 (1), 59A.
- Oates, S.F. (1999). Literacy as an everyday practice: case studies of students and literacy instruction in high school, community college, and university writing classes. [electronic version]. *Dissertation Abstract International* 60 (7), 2474A.
- Penny, J.A. (2003). Reading high stakes writing samples: my life as a reader. *Assessing Writing* 8, 192-215.
- Pipho, C. (2000). The sting of high-stakes testing and accountability. *Phi Delta Kappan*, May 2000, 645-646.
- Schneider, R. & Schneider, J.J. (2002). Writing instruction or destruction: lessons to be learned from fourth grade teachers' perspectives on teaching writing. *Journal of Teacher Education* 53, (4), 328-341.
- Schwartz, J. (2004). Reflections of an AP reader. *English Journal*, March, 53-57.
- Scherff, L. & Piazza, C. (2005). The more things change, the more things stay the same. *Research in the Teaching of English* 39 (3), 271-301.
- Spandel, V. & Stiggins, R.J. (1997). *Creating writers: linking writing assessment and instruction*. New York: Longman.
- Thomas, P.L. (2001). Standards, standards everywhere, not a spot to think. *English Journal*, Sept. 2001, 63-67.
- West-White, C.N. (2001). The writing assignment: the impact teachers' assignments and instruction have on students' preferences, perceptions, and class procedures. [electronic version]. *Dissertation Abstract International* 62 (7), 2391A.
- White, E.M., Lutz, W.D., & Kamusikiri, S., Eds. (1996). *Assessment of writing politics, policies, practices*. New York: The Modern Language Association of America.
- Willis, D.J. (2000). Students' perceptions of their experiences with Kentucky's mandated portfolio. [electronic version]. *Dissertation Abstract International* 61 (8), 3045A.
- Zigo, D. (2001). Constructing firebreaks against high stakes testing. *English Education* 33(3), 214-232.

Why the rich get richer or Why are high achieving students much more likely to seek academic assistance than low achieving students?

Shawn E. Nordell, Ph.D.
Biology

Introduction

Instructors sometimes observe that the students who attend their office hours or send them email questions about course material tend to be higher achieving students. Their perception is that higher achieving students are most likely to approach an instructor before, after or outside of class and ask for clarification of course material. Instructors express concern that they rarely see lower achieving students during their office hours or receive questions from these students via email. The low achieving students can benefit from academic assistance and yet these are exactly the students who rarely seem to seek this assistance. This is highly frustrating for most instructors as most are readily willing to offer assistance for these students to help them become academically successful.

This conundrum has been documented in secondary education (e.g. Karabenick & Knapp 1988; Ryan et al.1998) and more recently in higher education (Stavrianopoulos 2007). These studies corroborate the observation that high achieving students appear to be more likely to seek academic self help than low achieving students. This is despite the fact that academic help seeking behavior has been shown to be an important strategy that contributes to the academic success of students (e.g. Nelson-LeGall 1981).

One aspect that has not been studied is whether or not high achieving students are also most likely to seek academic assistance in learning strategies. Effective learning strategies allow students to adapt and regulate their learning protocols to best deal with their level of prior information and the current information and tasks (McWhorter 1988). Freshmen in particular, often have difficulties making the academic transition from high school to college. The higher education model of learning is often quite different from the high school model of learning. The quantity of knowledge learned, the pace of academic learning and the ability to synthesize and utilize knowledge at the college level is usually at a much higher pace or level than students have previously experienced. Therefore, it can be very challenging for new college students to understand and be successful using this different model of higher education learning. In order to be academically successful, students often need to revise their learning strategies for their college courses.

I examined which students attended voluntary study skills workshops associated with an introductory level course in order to examine whether or not high achieving students were most likely to seek assistance in learning strategies.

Methods

In the Fall 2006, I offered four identical one hour study skills workshops shortly after the first lecture exam in a large introductory biology majors level course (n =348). The workshops were held in the late afternoon or early evening for students in a majors level introductory biology course shortly after their first lecture exam. These workshops were announced in all lecture sections and students were encouraged to attend these workshops by the instructors and their advisors. Students were asked to sign in for each workshop but were not given credit for attending a workshop.

Results

Out of a total lecture enrollment of 343 students, 68 (almost 20% of the class) students attended the workshops. The students in the workshops represented 10 different majors with Biology, Biochemistry, and Exercise Science Physical Therapy comprising the majority of these majors (n= 42) (Table 1).

The majority of the students who attended a workshop received a B or better on their first exam (46 of 68) (Table 2). Students who scored a B or higher on Exam 1 had the highest participation rates while students who scored a B- or lower had the very lowest participation rates (22 of 68). The attendance rate of students who scored a B or higher was 20-25% while the attendance rate of students who scored a B- or lower ranged from 0-15%. Only three students with the very lowest scores on the first exam (D or F) attended a workshop (less than 9% of total students with this score).

To examine the effectiveness of attending a study skills workshop I evaluated the performance of students on the second lecture exam for two groups of students: those that did attend a study skills workshop and those that did not attend a study skills workshop.

The mean of Exam 1 was higher than the mean of Exam 2, (mean exam 1= 83.3% +/- S.E. .61; mean exam 2 = 68.7% +/- S.E. = .008) therefore the difference between the two exams was examined (e.g. Exam 1% - 2%). Students who did not attend a study skills workshop had a larger decrease (mean decrease = 15.5% +/- S.E. .7) than students who did attend a study skills workshop (mean decrease = 10.6% +/- S.E. 1.2) from Exam 1 to Exam 2 (t-test, $p < .0008$). This represents a five percent difference in the performance between these two groups, equal to half of a letter grade.

Discussion

Most of the students who attended the workshops were already high achieving students (received a B or better on Exam 1). Very few of the very low achieving students attended a workshop. Interestingly enough, even the performance of the high achieving students increased after attending a study skills workshop. Students attending a study skills workshop performed significantly better on the second lecture exam than students who did not attend a study skills workshop indicating that these workshops can enhance student academic performance. Study skills strategies, such as the ones taught in this workshop, are rarely taught within a course setting and are also even rarely taught in freshmen seminar classes. These results indicate that the addition of study skills strategies may be an important component to student achievement.

Why do the students who need academic assistance the most usually not seek help? Why are low achieving students not effective at seeking help? There are numerous possible explanations for this behavior. One possible explanation could be related to students academic self-efficacy. Low achieving students may perceive that seeking assistance will create the perception that they are lacking in ability (Pintrich & Schunk 1996). Some have proposed that high achieving students do not seem to create this perception or be concerned about the ramifications of this perception if they seek help (Ryan & Pintrich 1997). It is unclear why low and high achieving students might have different perceptions of their academic self-efficacy and how this might originate. It would appear that high achieving students would have the most to lose in terms of the instructor perceiving their lack of ability given that they are higher achieving students. Also, in this example, the students did not know before the workshop that they would be asked to sign in so attending the workshop could be perceived as being an anonymous event. These study skills workshops were offered

in conjunction with a large lecture class early in the semester. Therefore, the instructor had little opportunity to identify individual students.

Another possible explanation could lie in the social climate of the classroom. There is the potential that low achieving students are not comfortable asking questions due to their perception of an unfamiliar and perhaps not welcoming social climate of the classroom (Ryan et al. 1998). However, social climate could have the same impression on all students and not just low achieving students.

A new explanation may lie in the perception of students that it is not possible for them to alter their learning strategies or increase their efficiency or learning abilities. Carol Dweck has found that students may believe that intelligence is a fixed quantity or it may be acquirable (Mangels et al. 2006). If students believe that their learning strategies or capabilities are fixed they may perceive that they can not change their learning strategies or their academic outcomes. Therefore, low achieving students with this belief would not seek academic assistance. This is an important consideration for this problem that may help instructors better understand this critical problem. Dweck (2006) has begun to explore this hypothesis across a broad spectrum disciplines. Clearly, academic self help seeking is an important issue in academic excellence and future research on this topic is greatly needed.

Literature Cited

- Bransford, J.D., Brown A.L. & Cocking R.R. (editors) 2000. *How People Learn: Brain, Mind, Experience, and School*. National Academy Press, Washington. 374 pp.
- Dweck, C.S. 2006. *Mindset*. New York: Random House
- Karabenick, Stuart A. & John R. Knapp. 1988. Help seeking and the need for academic assistance. *Journal of Educational Psychology*. 80(3):406+408.

- McWhorter, K.T. 1988. *Study and Thinking Skills in College*. Glenview, IL: Scott, Foresman and Company.
- Nelson-Le Gall, S. 1981. Help seeking: an understudied problem-solving skill in children. *Developmental Review*. 1:224-246.
- Pintrich, P.L. & Schunk, D.H. 1996. *Motivation in education: Theory, research and applications*. Englewood Cliffs, NJ: Prentice Hall
- Ryan, A. M. Gheen, M. H. & C. Medgley. 1998. Why do some students avoid asking for help? An examination of the interplay among students' academic efficacy, teachers' social-emotional role, and the classroom goal structure. *Journal of Educational Psychology*. 90(3):528-535.
- Ryan, A.M. & Pintrich, P.L. 1997. Achievement and social motivational influences on help seeking in the classroom. In S. Karabenick (ed.), *Strategic help seeking: Implications for knowledge acquisition* (pp. 103-123), Hillsdale, NJ: Erlbaum
- Stavrianopoulos, Katherine. 2007. Adolescents' metacognitive knowledge monitoring and academic help seeking: The role of motivation orientation. *College Student Journal*. 41(2):444-453.

Table 1. The majors of the students attending the study skills workshop (n = 68 students)

Major	Number of Students Attending Workshop
Biochemistry	6
Biology	21
Biomedical Engineering	2
Business Administration	1
Chemistry	4
Deciding	9
Exercise Science Physical Therapy	15
Nutrition and Dietetics	3
Nuclear Medicine Technology	1
Occupational Sciences	2
Psychology	4

Table 2. The exam scores of students who attended a study skills workshop.

Grade on Exam 1	Number of students attending workshop (% of total in class with that grade)
A (94-100%)	13 (25%)
A- (90-93%)	12 (19.6%)
B+ (87-89%)	8 (27.6%)
B (84-86%)	13 (24.5%)
B- (80-83%)	4 (14.3%)
C+ (77-79%)	6 (13.9%)
C (74-76%)	6 (26.1%)
C- (70-73%)	3 (15.7%)
D (57-69%)	3 (9.4%)
F (0-56%)	0 (0%)

Students' Perceptions of Professors in Non-traditional Teaching Roles

Brien K. Ashdown and Kristin L. Kiddoo
Psychology

Originally published in *The New School Psychology Bulletin*, 2007, Vol. 5, No. 2, pp. 37 – 40.
Reprinted with Permission

Abstract

University students participated in 4 focus groups (Caucasian males, Caucasian females, ethnic minority males, and ethnic minority females). Participants were asked to discuss 3 classroom scenarios: (a) a White male teaching a race relations course, (b) a male teaching a women's studies course, and (c) a young, childless female teaching a child psychology course. Most participants preferred an ethnic minority professor for the race relations course, a female professor for the women's studies course, and showed a slight preference for a married female professor for the child psychology course.

Previous research has addressed the hypothesis that professors who are members of minority groups—specifically, African-Americans, Latinos, and women—are evaluated differently than members of majority groups (Anderson & Smith, 2005; Hendrix, 1998;

Kierstead, D'Agostino, & Dill, 1989). Basow (1998) reported that female professors are frequently evaluated differently by students than are male professors, and that these evaluative differences include differences in teaching style and perceptual biases.

Latina professors are often evaluated differently than European-American female professors with similar teaching styles. In one study, Latina women were seen as warmer when they had lenient teaching styles and less warm when they had stricter teaching styles than European-American female professors with comparable teaching styles (Anderson & Smith, 2005). In other research, student evaluations were impacted by three factors: professor's sex, facial expression and contact with students. Female professors who were perceived as more friendly were given higher student evaluations, but similar perceived friendliness did not impact male professors' evaluations. In addition, male professors were seen as more effective than female professors regardless of facial expression or contact with students (Kierstead, D'Agostino, & Dill, 1989).

Hendrix (1998) conducted one-on-one interviews with students in classes taught by European-American professors and African-American professors. Students reported perceiving African-American professors as usually having to work harder to establish credibility, but having an easier time doing so when they are teaching courses that could be linked to their "Blackness." Thus, students conferred more credibility to African-American teachers who taught "ethnic" classes. Interestingly, students also conferred more credibility to African-American professors who taught English, stating that with a host of prevalent "Black" dialects, an African-American who could teach standard English was perceived as "an extra plus" (p. 749). At least one interviewee stated that it would be difficult for a European-American professor to establish credibility in a class that addressed ethnic studies

because students may desire a “minority view” (p. 749), even if the European-American professor had experience living with minorities.

The Present Research

The objective of the current research was to determine whether there were certain professors that students would or would not be willing to learn from due to the professors’ personal background and the particular course topics. For example, how would students feel about taking a race relations course from a White male professor? How would students feel about taking a women’s studies course taught by a male professor? We wanted to address specific questions about professors in what might be called ‘non-traditional roles’. How would the students perceive the professors’ credibility and how difficult would it be for the professors to establish credibility with the students? Would students feel that a professor in a non-traditional role would evaluate students fairly, or would certain groups (i.e., ethnic minority students, women) raise the possibility of being subjected to different grading standards? We wanted to know whom these students would choose as the ideal candidates for certain courses.

Method

Participants

Participants were undergraduate psychology students at a private Midwestern university who participated in exchange for credit toward a psychology course requirement. Four focus groups were held. Since race and gender were critical components of the research questions, focus group sessions were demographically divided as follows: Caucasian Males, Caucasian Females, Ethnic Minority Males, and Ethnic Minority Females. When signing up,

participants were asked only to participate in the session that applied to them. The 4 focus groups had varying numbers of participants. There were 3 participants in the ethnic minority male group, 5 in the ethnic minority female group, and 10 each in the white female and white male groups.

Procedure and Materials

Participants signed up for the study via the psychology department's participant pool website (Experimetrix). The moderator for each focus group was of the same gender as the members of the group. A note-taker of the opposite gender was also present during each session.

After reading the recruitment statement and obtaining informed consent, which included a statement explaining that the session would be audiotaped to allow for transcription of the conversation, the group was given the first hypothetical scenario and asked to provide their personal opinions (see Appendix A). After answering all the questions, participants were thanked for their participation, debriefed about the study, and dismissed.

Analysis

After audiotapes from each of the 4 focus group sessions were transcribed verbatim, open-coding following the method provided by Strauss and Corbin (1998) was conducted. The statements of participants were organized into 'segments', which were then clustered together into categories, and finally coded into broad themes. Segments were simply key terms and phrases that were deemed important and relevant to the focus group conversation by the researchers. After segments were pulled from the transcribed conversations, segments that were similar to one another were grouped together into categories. Finally, similar categories were grouped together into general themes. These broad themes were then used to analyze the focus group conversations (Strauss & Corbin, 1998).

Results

Initial Credibility

When given the scenario of a White male teaching a course on race relations, most participants indicated that they would view the instructor as being highly credible to teach, regardless of their own race or gender. One participant stated that, “Just because he’s White doesn’t mean he doesn’t have knowledge or some kind of experience-base; this doesn’t mean he’s not qualified.”

Only the minority males said that they would not be initially surprised that a White male would be teaching this course. All other groups said they would be surprised to find a White male as the instructor for the course. After this initial surprise, though, all of the groups said they would be satisfied with a White male instructor as long as he was qualified. At least 1 participant, however, cited the instructor’s age as being a factor in whether or not he would be viewed as credible, indicating that they “would be very skeptical of him, especially if he was, like, anywhere under 50 [years old].”

When asked about a male teaching a course on women’s studies, females (especially White females) were against the idea, and rated his credibility as very low. Their initial response to the suggestion was often a simple “No.” Male students felt that while he could be well schooled, the instructor would still lack the personal experience of being a female. When asked what this instructor could do to gain more credibility, the most common reaction was along the lines of showing respect for and valuing women. One participant felt that he could gain credibility “if he was gay.” All groups, other than the White females, eventually said that they would give the instructor “the benefit of the doubt,” unless he did something like “mak(e) sexist remarks.”

Asked about a young, childless female teaching child psychology, most participants assumed high initial credibility, often citing the research and academic experience that would be required to teach this type of course. Many also pointed to the fact that she was a female, and that “just the fact that she is a woman and not a male gives her more credibility.”

Fairness in Evaluation

Participants were next asked to express their opinions as to whether the instructors in the given scenarios would grade all groups of students fairly. For example, in the race relations course, participants were asked whether they felt that the White male professor would grade White students and minority students without bias. In general, White participants thought that he would be fair in grading, while some minority students thought that he might favor White students, citing unavoidable biases. For example, one minority student said, “I think it would be difficult for him to be fair to everyone.”

Evaluation fairness by the male teaching a women’s studies course elicited less response, especially from the males. Some females, however, thought that he might “cut [males] a little slack,” and expect more from the female students.

When asked whether they felt that the young, childless female teaching child psychology might expect more from students with children than from the more traditional, young undergraduates, 1 participant agreed, saying that the instructor might “expect a little more out of older students,” especially if a portion of the grade was for participation, when parents would have more stories to share.

Ideal Candidate

The final question posed to participants was whom they would select as the ‘ideal candidate’ to teach each of the courses given in the three scenarios. After being given the

example of two résumés with equivalent qualifications but different demographic characteristics, most participants said that their ideal candidate would be the stereotypically appropriate one. Most believed that “the minority would get the job” to teach the race relations course, a woman would be the ideal candidate to teach the women’s studies course, and a parent (and more specifically a mother) would be the best selection for a course in child psychology.

Discussion

Participants were hesitant to address the issue of a White male teaching a race relations course. This could be due to social desirability or a desire to appear politically correct, especially when speaking to a group of their peers where anonymity could not be guaranteed. Nearly all of the students, regardless of ethnicity or gender, stated that there would be no bias in teachers’ grading of students, and that any person with the appropriate education and training could have credibility teaching the race relations course. Only when directly asked to select a professor did the students suggest that an ethnic minority professor would be a better choice for the class. Nevertheless, White students were defensive of a White male teaching the race relations class, stating, “White is a race, too.”

Regarding the question of a male professor teaching a Women’s Studies course, differences arose among some of the groups. The White females were adamantly against a male professor teaching women’s studies. When initially asked the question, most responded with a simple but definitive “No.” It took more questioning to get them to discuss the topic any further. Finally, most of them based their decision on the idea that all of the necessary information for this class could not be taken from a textbook. They claimed that first-hand experience as a woman was important, even imperative.

On the other hand, the group of minority females seemed most open to the idea of a male teaching women's studies. They claimed that a male might bring an interesting outside perspective to the class. Though he would have to be careful and conscious about the way he taught the class and presented the information, an educated and trained male could do a good job. As a minority group, these women may have had experience trying to understand things from the majority group's perspective, and might be willing to admit that an outsider could indeed understand another group.

As for the scenario of the young, childless female teaching the child psychology course, some participants preferred a woman, a younger teacher, or a parent. However, the consensus was that anyone could teach the course. Most of the participants stated that all the information for the course could be gathered from a textbook or professional research.

An interesting contradiction in the way the participants spoke about the qualifications of the professors was the idea of a textbook versus personal experience. For the women's studies course and race relations course, nearly all participants said that in order to teach the course effectively and credibly, personal experience was required. On the other hand, for the child psychology course, participants said that experience would be helpful, but that textbook information would be sufficient. It appears that courses which participants saw as less "scientific" or more personal could not be confined to the information available from a textbook.

While an attempt was made to form focus groups of balanced numbers of European-American and ethnic minority students, it was difficult to form such comparable groups because of the overall student composition of the university from which the sample was taken. In order to form the focus groups of ethnic minority students, we were forced to assign all ethnic minorities together to the same group, consisting of 3 participants for the ethnic

minority male group and 5 for the ethnic minority female group. Indeed, this is one of the limitations of the present study. If it had been possible to obtain groups of specific ethnicities, the study might have produced more varying and interesting results. For example, if it had been possible to construct a group of Latino males, a group of Latina females, a group of Middle Eastern males, a group of Middle Eastern females and so on, more detailed and ethnic-specific data might have been collected. Instead, the current research was limited to dividing groups simply by gender, ethnic majority and ethnic minority. It would be beneficial if future research on this topic could further delineate the focus groups by ethnicity in order to provide more ethnic-specific information.

The current research was conducted with focus groups of 3 to 10 participants. While at times this was helpful in creating a discussion in which participants fed off of each other's comments, it also seemed to cause some problems. Most participants quickly and readily agreed with the first comment made by a fellow participant. Perhaps because they were discussing potentially controversial subjects in a group of peers where anonymity could not be guaranteed, participants were sometimes hesitant to make their personal views known. Perhaps one-on-one interviews, where both anonymity and confidentiality could be guaranteed, would produce an environment where participants would be more willing to share their personal, and possibly controversial, viewpoints.

While originally willing to grant most instructors "the benefit of the doubt," it is clear that students preferred what might be termed 'stereotypically appropriate professors'. This was especially true for courses that were regarded (perhaps incorrectly) as less scientific, such as women's studies and race relations. Many such courses are generally offered as diversity classes by universities. Yet students' expectations may subtly influence universities to employ only stereotypically appropriate instructors for such courses, thus passively

eliminating other viewpoints on the topics. Ironically, this directly opposes the goal and purpose of diversity in the classroom. This might lead students to believe that conversations about topics such as race or gender are only open to stereotypically appropriate participants.

Steps should be taken to make students aware that this is not the case. Guest lecturers, non-traditional instructors, and new teaching methods should be utilized to ensure that students understand that these topics of learning and conversation are open to everyone, and that true understanding and knowledge take place only when everyone is participating, regardless of ethnicity or gender.

Appendix A

Questions Asked of Each Focus Group

1. As a student, how would you view a professor's credibility as an instructor in the following scenarios:
 - a. A White male teaching a course on Race Relations?
 - i. How could the professor change that initial credibility perception?
 - b. A male teaching a course on Women's Studies
 - i. How could the professor change that initial credibility perception?
 - c. A young, single female with no children teaching a course on Child Psychology.
 - i. How could the professor change that initial credibility perception?

2. As a student, how would you view non-traditional professors' fairness in evaluating students?
 - a. Do you think that the White male teaching Race Relations would grade ethnic minority students differently than White students?
 - i. Do you think that students from a different race than you may have a different view?
 - b. Do you think that the male teaching a course on Women's Studies would grade female students differently than male students?
 - i. Do you think that students of the opposite gender to yours may have a different view?
 - c. Do you think that the young, childless female would grade older students with children differently than younger students without children?
 - i. Do you think that students that differ from you in age and parenting experience may have a different view?

3. In your view, who would best be qualified to teach the following courses:
 - a. Race Relations
 - i. Why?
 - b. Women's Studies
 - i. Why?
 - c. Child Psychology
 - i. Why?

References

- Anderson, K. J., & Smith, G. (2005). Students' preconceptions of professors: Benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Sciences*, 27, 184-201.
- Basow, S. A. (1998). Student evaluations: The role of gender bias and teaching styles. In L. H. Collins, J. C. Chrisler, & K. Quina (Eds.), *Career strategies for women in academe: Arming Athena*. (pp. 135-156). Thousand Oaks, CA: Sage Publications.
- Hendrix, K. G. (1998). Student perceptions of the influence of race on professor credibility. *Journal of Black Studies*, 28, 738-763.
- Kierstead, D., D'Agostino, P., & Dill, H. (1989). Sex role stereotyping of college professors: Bias in students' ratings of instructors. *Journal of Educational Psychology*, 80, 342-344.
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage Publications.