AN INVESTIGATION INTO THE RELATIONSHIP BETWEEN PARTICIPATION IN COMPETITIVE FORENSICS AND STANDARDIZED TEST SCORES

by Tammie L. Peters

Abstract: Sponsors of competitive speech programs must prove scientifically how forensics improves student achievement, as defined by No Child Left Behind (NCLB; 2001). While many studies have shown a connection between debate experience and improved critical thinking skills, few studies have linked competitive speaking specifically to the standardized tests required by NCLB. This researcher examined the state and national test scores of similarly motivated honors English students in a single high school, over the course of 4 years, and compared the scores of forensic students against the scores of non-forensic students. It was found that students with experience in competitive speech scored significantly higher ($\alpha = 0.03$) on state administered writing tests and significantly higher ($\alpha = 0.07$) on a nationally normed reading test. Additionally, this study revealed no significant difference in test scores between students who competed in the debate events vs. those students who competed in the non-debate events.

Those of us who coach forensics know that this is an extraordinarily worthwhile and valuable activity. However, in these times of financial belt-tightening and the requirements of No Child Left Behind, forensics programs are in jeopardy. School administrators and teachers feel pressured by national and state requirements to raise test scores or be forced to endure official sanctions and punishments. Supporters of programs, especially those outside of the core classes of basic Language Arts (i.e., reading and writing), Mathematics, Social Studies, Science, and Foreign Languages, find themselves forced to demonstrate how student participation in these non-core subjects will improve test scores or else risk reduction or elimination of support, including funding and teacher time.

STATEMENT OF THE PROBLEM

Researchers (Collier, 2004; Rogers, 2002) have observed that no studies have been done on the impact of forensics on standardized test scores. Instead, a number of authors (Allen, Berkowitz & Louden, 1995; Carroll, 2007; Crawford, 2003; Hier, 1997; Massey, 1999; McCrady, 2004; Minch, 2006; Parcher, 1998; Re, 2002; Rogers, 2002; Sellnow, 1994; Tumposky, 2004; Warner & Bruschke, 2001) have commented on the logical effects that participation in forensics should have on student test scores; after all: (a) debaters and extemporaneous speakers must research and evaluate evidence while they organize arguments quickly, (b) orators must do a great deal of research and compose carefully worded speeches, and (c) interpreters of literature must study it carefully in order to understand the best way to orally present the material to make an emotional impact on the audience. Certainly, the acquisition of these skills should be expected to contribute to higher scores in reading and writing. Credible support for the connection to standardized test scores is essential if sponsors of forensics are to be able to defend their programs over the next few years.

Other researchers (Allan et al., 1999; Fine, 2001; Greenstreet, 1993; Vaughn & Winner, 2000; Whalen, 1991) have noted that any possible connection between participation in forensics and higher test scores is the result of the higher motivation and intellectual levels of forensics students in comparison to the rest of the student population. Greenstreet described this problem as the "chicken/egg" (p. 18) quandary: if forensics participants have better test scores, is it because of something the students learn in forensics, or is it because they are smarter and more motivated students? Any researcher will have to consider this problem in order for the results of his or her study to be considered credible.

Findings from credible studies in regard to the connection between participation in forensics and any intellectual and educational benefits are critical if forensics programs are to survive. Anderson (1974, as cited in Greenstreet, 1993) warned, "In an age of educational accountability, the forensics community is and will increasingly be called upon to tell what it seeks to do, how well it accomplishes its goals, and what other effects it has" (p. 24). Without solid research findings that connect forensics participation to increased test scores, this activity will be lost in "budgetary shuffles and the panic to improve the basics" (Warner & Bruschke, 2001, p. 2). These writers were harbingers of the research necessary to defend any academic program under NCLB (2001): under this law, only those programs "that have been demonstrated to be effective through rigorous scientific research" (U.S. Department of Education, 2003, ¶1) justify inclusion in schools. McCrady (2004) observed that forensics programs have been cut already in various schools because administrators see them as expendable programs that do not contribute to the basic education mandated by law. As long as the U.S. has a culture that values standardized test scores as a measurement of school, teacher, and student success, forensics educators will have to establish a definite link between competitive speech and higher test scores in order to survive.

Background on CSAP

Mandatory student testing began in Colorado before the U.S. Congress passed

NCLB (2001). Members of the Colorado State Legislature established the Colorado Student Assessment Program (CSAP; 1997) in 1997. It is a series of criterion referenced tests based on curriculum performance standards. In 1997, state officials mandated that two tests in Reading and Writing be administered to students in fourth grade (Colorado Department of Education, CSAP Summary Data section, 2007); by 2006, the number of tests had increased to 31 tests across eight grades, plus students in the eleventh grade were required to take the ACT (1989). Currently, all students in Grades 3-10 are tested in Reading, Writing, and Mathematics; additionally, students in Grades 5, 8 and 10 are tested in Science.

Since students in forensics learn about how to read information and manipulate language to create an argument, this researcher is interested in the Reading and Writing tests, especially the tests administered to high school students. The Reading tests include short passages of fiction and nonfiction accompanied by both multiple choice questions and paragraph length constructed responses. In the Writing tests, students are required to: (a) edit texts, (b) evaluate sentences, (c) write paragraph length constructed responses, and (d) write one essay length extended constructed response. More information about the CSAP and how it will be used in this study is provided below.

REVIEW OF LITERATURE

In light of No Child Left Behind (2001) requirements, findings of scientific research are essential for supporters of high school forensics programs to defend their programs. To date, no specific research has been done to study the connection between state mandated standardized test scores and participation in competitive speech. Before this researcher examined the test score data, it was important to understand the context of this research. Participation in speech and debate has been important to a variety of cultures, and the development of competitive speaking has been notable. The impact that participation in competitive speech has on critical thought and other skills has been studied for many years. While there is a strong connection between participation in forensics and the development of academic abilities, there are some deficiencies in past studies.

Historical Perspective of Competitive Speech

The history of forensic speaking is long and rich. Lewis (2004) noted that, in numerous cultures, quality speaking skills have been valued since the very earliest days. In ancient Greece, books were rare. Trained performers would present poems, both original and by other authors. Contests occurred in which the talents of the interpreters were pitted against each other. The importance of such performers continued through the time of the ancient Hebrews, the Roman Empire, and the Middle Ages of Europe. Stories, histories, myths, legends, and other important ideas were recited by talented storytellers termed *bards* in France, *scops* in England, and *fili* in Ireland. Similarly, Crawford (2003) explained that orators were important in ancient Rome, and debates in the Senate were a critical part of the government. When Demosthenes used pebbles to practice clear speech, he was involved in a long established tradition in which public speaking was valued. During the Enlightenment, debate skills were essential to the salons of France (Carroll, 2007). In the United States, the value of public speakers was evident in the growth of the Chautauqua movement of the late 19th century, which brought speakers and musicians to towns across the U.S. (Canning, 2000). According to Canning, Theodore Roosevelt called the Chautauqua movement "the most American thing in America" (¶1). The most popular speakers were lecturers, like William Jennings Bryan, and elocutionists, who created one person shows from pieces of literature. Clearly, public speaking, including debate and interpretation of literature, has been valued throughout time and across cultures.

In the U.S., the historical respect for excellent speaking skills led to the development of interscholastic competitive speech (Barfield, 1989). Intercollegiate competitions began in 1872, and Southwestern College created the tournament format for multiple teams in 1923 (Barfield). The oldest high school debate society in the U. S. is at Phillips Academy in Andover, Massachusetts, where debate was an established student activity as early as 1825 (Phillips Academy, 2007). "From 1855 to 1890, debate presented one of the more popular forms of intellectual entertainment in many schools" (Borchers and Wagner, 1954, as cited in Barfield, p. 49). In 1895, teachers in a group of high schools in Wisconsin formed the High School Lyceum Association, which was "devoted to promoting interscholastic debate" (Barfield, p. 51). By 1925, a group of high school teachers began to organize a national honor society for interscholastic speech competitors (National Forensic League [NFL], 2007). This group established the NFL as a national honor society for speech. Initially, only six events were offered: debate, original oratory, extemporaneous speaking, interpretation of dramatic literature, interpretation of humorous literature, and oratorical declamation. Over the next 70 years, other events were added. In 1945, members of the National Association of High School Principals placed NFL competitions on their list of approved competitions and activities. In 2007, over 1.2 million students had become members of the NFL, and over 2,000 schools had earned NFL charters.

Forensic competition continues to be valued worldwide. In 1999, members of the Open Society Institute created the International Debate Education Association (IDEA; 2007). The IDEA was designed to promote debate and discussion in "those societies where democracy is in its infancy and where negotiated resolution to conflicts and cross-community dialogue are little-established concepts" (¶2). Currently, IDEA events take place in 27 countries.

Throughout time, members of many cultures have valued speech and the benefits it provides, especially to young people. Excellent speaking skills have been respected from the earliest times through today. Competitive speaking has been appreciated since ancient Greece and continues to be important today.

Various Benefits of Participation in Forensics

McCrady (2004) commented, "All veteran and even novice coaches know in our hearts that our programs have immeasurable educational value" (p. 41). A variety of benefits have been ascribed to participation in competitive speech. For example, competitors in forensics develop better academic skills and succeed more than their peers in school. Barfield (1989, as cited in Bellon, 2000) "found that participation in competitive debate among high school students positively correlates with significant gains in cumulative GPA" (p. 166). Collier (2004) found a similar effect in her study of high school debater students in inner city schools. She concluded, "Two results are clear – debaters achieve significantly higher grades and intend to attend college at a substantially greater rate than their non-debating peers" (p. 28). Warner and Bruschke (2001) concurred: high school debate can lead to improvement in student grades in other academic courses. In his study, Rogers (2002) found collegiate debaters "were able to maintain slightly better GPAs than their non-debate peers. They were significantly stronger academically" (p. 21),

as determined through a variety of indicators. Fine (1999, as quoted in Bellon) hypothesized that this positive effect of competitive speech was because forensics, particularly debate, "appears to strengthen students' ability to persevere, remain focused, and work toward challenging goals" (p. 166).

Another important benefit attributed to participation in forensics is the increase in civic awareness and the empowerment of students to be productive members of a democracy. Re (2002) argued that forensics, especially debate and extemporaneous speaking, "expose young people to global and international perspectives" (p. 4). The study of current events and the experience of public speaking lead students to participate actively in civic activities. Warner and Bruschke (2001) found, "Students who can face and overcome those challenges and those fears [of competitive speaking] are seldom afraid of public dialogue in any other context, be it a political rally, city board meeting, electoral campaign, legal proceeding, or town hall meeting" (p. 7). Rogers (2002) concluded even more strongly: "Debaters were significantly more likely to participate in the democratic process through voting, volunteering their time and resources to political campaigns, and participating in social activism" (p. 21).

Also, participation in forensics may decrease adolescent violence. Bellon (2000) explained that increased verbal skills and argumentation skills could provide youths with alternatives to violence. Collier (2004) suggested that participating in debates provided students with the requisite tools to resist negative peer pressure. Warner and Bruschke (2001) explained that debaters "are actually more empathetic, less ego-centric, and better at taking the perspective of others" (p. 15). Rogers (2002) found similar results in his study of college student attitudes. Collier found the same effect, and she hypothesized that "debate gives these students a reason to achieve – a reason to reject risky behaviors" (p. 27). Students with forensic experience may learn how to use words instead of violence to solve problems.

Student participation in forensics, especially debate, may lead to these benefits because of the teaching methods used by speech coaches. In particular, forensics is a type of experiential education, in which students analyze real issues and then defend their analyses outside of the classroom (Sellnow, 1994). Hier (1997) suggested that forensics is an excellent delivery system for education because forensic educators use "hands-on methods that produce more retention" (p. 7). Bellon (2000) explained that the use of constructivist teaching methods, where students are actively engaged in the construction of knowledge, are powerful tools to increase student achievement; participation in competitive forensics provides these types of constructivist opportunities.

Overall, the list of benefits attributed to participation in competitive speech and debate is impressive. Forensic competitors tend to achieve higher grades, be better citizens, and accept others' views and fight less. These perceived benefits may be why many college admission officers prefer forensic competitors, especially captains of debate teams, when they accept applicants for their schools (Luong, 2000). Also, it may explain why many employers tend to prefer to hire former debaters over other applicants (Parcher, 1998). Colbert and Biggers (1985, as quoted in Bellon, 2000) explained, "In a time when many of our students ask us how educational activities will help them get a job, the answer seems to be unequivocal. Debate experience is highly valued by the business world" (p. 167).

Critical Thinking Skills and Forensics

Historically, one reason public speaking has been valued is due to the association between it and critical thinking skills. Critical thinking skills are vital to society. As Dressel and Mayhew (1954, as cited in Korcok, 1997) noted, "The good life in a democratic society. . . seems to rest fundamentally on one's ability to think critically about those problems with which he (or she) is confronted" (¶7). Massey (1999) wrote about the importance of critical thinking to the Postindustrial Era since "those with a diversity of knowledge (i.e., those with training in critical thinking skills) are the ones who seem to have the best ability to attain success" (p. 24). Members of the Partnership for 21st Century Skills (2004), an education advocacy group made up of representatives from major businesses, defined "critical thinking and problem solving skills" as "essential to prepare students for the future" (¶1). The former Governor of California even issued an executive order about the importance of critical thinking skills for students (Korcok, 1997). Katula and Martin (1984, as cited in Whalen, 1991) "identified critical thinking as an essential element of our society's ability to develop literacy" (p. 391). Also, critical thinking skills are valued in the State of Colorado as identified in the goals of the Colorado Student Assessment Program (CSAP; 1997; Colorado Department of Education, 2007).

There is little agreement on the exact definition of critical thinking. However, most of the definitions share similarities. Watson and Glaser (1939, as quoted in Brembeck, 1949) explained:

> Critical thinking involves (a) a persistent effort to examine any belief or supposed form of knowledge in the light of the evidence that supports it and the further conclusions to which it tends, as well as the ability (b) to recognize problems, (c) to weigh evidence, (d) to comprehend and use language with accuracy and discrimination, (e) to interpret data, (f) to recognize the existence (or non-existence) of logical relationships between propositions, (g) to draw warranted conclusions and generalizations and (h) to test the conclusions by applying them to new situations to which they seem pertinent. (p. 177)

Dressel and Mayhew (1954, as quoted in Whalen, 1991) maintained that critical thinking involves five characteristics, the:

> (a) ability to define a problem, (b) ability to select the appropriate information for the solution, (c) ability to recognize both stated and unstated assumptions, (d) ability to select relevant hypotheses, and (e) ability to draw valid conclusions and inferences. (p. 391)

Garside (1996; as quoted in Allen et al., 1999) concluded that:

> the literature suggests at least four defining aspects of thinking that make it *critical*: (a) thinking that is clear, precise, accurate, relevant, logical, and consistent; (b) thinking that reflects a controlled sense of skepticism or disbelief of any assertion, claim or conclusion until sufficient evidence and reasoning is provided to conclusively support it; (c) thinking that takes stock of existing information and identifies holes and weaknesses, thereby certifying what we know or don't know; and (d) thinking that is free from bias, prejudice, and one-sidedness of thought. (p. 18)

Finally, the members of the Partnership for 21st Century Skills (2004) define critical thinking as:

> (a) exercising sound reasoning in understanding; (b) making complex choices and decisions; (c) understanding interconnections

among systems; (d) identifying and asking significant questions that clarify various points of view and lead to better solutions; and (e) framing, analyzing and synthesizing information in order to solve problems and answer questions. (¶1)

These definitions share certain commonalities; in particular, critical thinking seems to include the ability to (a) gather and carefully evaluate evidence to solve a problem, (b) avoid preconceived notions and biases, (c) remain open to new ideas, and (d) apply information to a variety of situations.

Logical Connections to Critical Thinking

In terms of academic and life skills, students who participate in forensics are exposed to critical thinking techniques. The connection between forensics participation and critical thinking skills is logical. Hunt (1994, as quoted in Parcher, 1998) commented, "Forensics helps you learn to be able to think clearly and adapt to rapid change" (¶5). Parcher wrote that the "creation of an argument is one of the most complex cognitive acts that a person can engage in" (¶6); since students in all forensics events must create arguments, typically forensics students engage in such complex thinking, regardless of the specific type of competitive event in which they are engaged.. The development of these arguments requires (a) research, (b) organization and arrangement of information, (c) anticipation of what others might think about the same subject, and (d) evaluation of how to best use materials (Minch, 2006; Parcher, Tumposky, 2004); these requirements are the elements of critical thinking. Freeley (1990, as cited in Korcok, 1997) explained that the fundamental elements in the creation of an argument are the essence of critical thinking. Carroll (2007) wrote that when students participate in forensics, especially the public speaking events, they are introduced to formal logic and argumentation, which "build critical thinking skills" (p. 34).

Studies about Critical Thinking and Forensics

Investigations into the connection between the ability to think critically and participation in forensics have been conducted for more than 60 years (Korcok, 1997). The first study was conducted by Howell in 1943 (Korcok).

Howell (1943)

Howell (1943) studied the impact of participation in high school debate on the scores on the Watson-Glaser Critical Thinking Appraisal (WGCTA). He asked 218 debate students from 28 Wisconsin schools to participate. In order to develop a control group, he matched each debate student with a similar student who did not participate in debate, and he matched the participants by age, academic record, gender, and I.Q. scores. Howell administered four of the Watson-Glaser tests to each student as both a pretest and a posttest. Overall, Howell found "debaters outgained non-debaters in critical thinking scores over the experimental period of six months" (p. 100). However, the difference between the scores of the debaters vs. those of the control group was not statistically significant. In the comparison between the scores of all debate students and the scores of all control students, Howell found an 85% chance that the improvements in critical thinking skills were not due to chance. When he compared the scores of debate students to the scores of non-debate students with matched I.Q. scores, "the debaters again outgained the non-debaters" (p. 100), but there was only an 89% chance that this difference was not due to chance. Howell suggested that the reason his quasi-experiment did not attain statistical significance was due to the variety of teaching methods and program emphases in the 25 different schools. He noted, "Great differences in mean gains of debaters over non-debaters were found among the participating schools" (p. 100-101). Similarly, Colbert (1995) wrote that the "findings implied instructional techniques, methods, and/or content probably influenced the acquisition of critical thinking skills" (p. 60). Even though Howell's findings did not demonstrate a definitive connection between participation in debate and increased critical thinking scores, Korcok (1997) observed this study "was sufficiently suggestive of a relationship to motivate further research" (¶21).

Also, Howell (1943) established the design for such studies: (a) establish a control group and a test group, (b) administer the WGCTA as a pretest, (c) wait while students engage in forensics events for a specified period of time, and (d) administer the WGCTA as a posttest. Allen et al. (1999) pointed out that, in 14 later studies, this basic protocol was followed.

Brembeck (1949)

The next major study was conducted by Brembeck (1949). Brembeck was interested in how participation in argumentation courses might affect critical thinking abilities of college students. He examined courses in argumentation at 11 different universities, and a total of 202 students were in his experimental group. His control group consisted of the same number of students from each of the schools. "The two groups were equated as carefully as possible according to age, sex, educational background, debating experience, and number" (p. 178). Like Howell (1943), Brembeck administered four of the Watson-Glaser Tests of Critical Thinking (WGTCT), which had been revised since Howell's study. Brembeck concluded, "The argumentation students, as a whole, significantly outgained the control students in critical thinking scores . . . There is approximately one time in a hundred that this difference could occur by chance" (p. 187). Also, Brembeck noted, "Argumentation students with high school and/or college debate training made significantly higher pretest scores than those without debating experience" (p. 188). Brembeck's study is important to this project in two ways: (a) forensics programs are one type of argumentation course offered in high schools, and (b) students with high school debate experience seem to be better prepared for critical thinking requirements in collegiate courses than students without debate experience.

Cross (1971)

Another important study was conducted by Cross (1971, as cited in Colbert, 1995). Cross administered the WGCTA to 136 students from nine high schools. The participants "were novice debaters participating in their first semester of debate" (Cross, as cited in Korcok, 1997, ¶30). In addition, Cross noted the amount of participation by each student over the course of the year and assigned them to groups of high participation and low participation. After a year, Cross administered the WGCTA again and found, "Those who are drawn to competitive debate, low and high participants, and continue for one academic year have greater thinking facilities than those who are not attracted to debate" (as quoted in Colbert, p. 56). He found that "high participation in competitive debate accelerates debaters' capacity in critical thinking while low participation may not enhance critical thinking beyond the normal improvement in an academic year" (as quoted in Korcok, ¶31).

Allen, Berkowitz and Louden (1995)

Allen et al. (1995) compared the gains in

critical thinking skills among: (a) students in introductory communications courses, (b) students in argumentation courses, and (c) students in competitive debate. They administered the WGCTA test, as revised in 1961, to 138 undergraduate students at 5 universities. They tested 34 students in introduction to interpersonal communication courses, 37 students in public speaking courses, 32 students in argumentation courses, and 35 students involved in some form of competitive speech including debate and non-debate events. After a semester, they readministered the tests to the same students. They found, "Both argumentation classes and forensic participation increased the ability in critical thinking compared to introductory interpersonal communication and public speaking classes" (p. 6). Among the four types of experiences, they found "participation in competitive forensics demonstrates the largest gain in critical thinking skills" (p. 6).

Allen, Berkowitz, Hunt and Louden (1999)

Allen et al. (1999) conducted a meta-analysis of studies in which the connection between communication instruction, including competitive forensics, and critical thinking skills was examined. First, they critiqued the design of the Watson-Glaser tests, in all of the forms; "the methodological issue is whether one can measure critical thinking using an objective test and whether an objective test completely captures the domain of critical thinking" (p. 20). However, since most of the researchers examined used various editions of the Watson-Glaser tests, Allen et al. recognized that they were limited in their study, and additional research needed to be done to determine the validity of these tests. Then, Allen et al. established the methodology of their meta-analysis. They limited their study to manuscripts, both published and unpublished, that contained quantitative data; examined some type of communication skill improvement exercise, such as a course or participation in competitive speech; and included some method to assess critical thinking skill improvement. They examined both longitudinal studies and cross-sectional studies. "The data were analyzed using the variance-centered form of meta-analysis developed by Hunter and Schmidt (1990)" (p. 23). They found that both longitudinal designs and cross-sectional designs showed "communication skill exercises improve critical thinking" (p. 24). Participants in competitive forensics "demonstrated the largest improvement in critical thinking scores" (p. 27). As for the deficiencies in the Watson-Glaser tests, Allen et al. found that, "when compared to other instruments, the Watson-Glaser measurement for critical thinking reported smaller not larger gains for communication skills training" (p. 25). Thus, in any studies in which the Watson-Glaser tests were used, the researchers may have underestimated the effect of communication instruction, such as competitive speech, on critical thinking skills. The Allen et al. conclusion means the connection between forensic participation and critical thinking may be greater than previous researchers had suspected.

Participation in Forensics and Standardized Test Scores

While the studies about the effects of participation in forensics on critical thinking are intriguing, because of the NCLB (2001) and CSAP (1997) requirements, students must improve specifically their scores in reading, writing and mathematics. Even though CSAP includes questions that evaluate critical thinking skills, it evaluates other skills and knowledge as well. Thus, any studies that examined the relationship between forensics participation and standardized test scores are especially important to this researcher.

Barfield (1989)

The first major study to use nationally normed standardized tests was conducted by Barfield (1989). He used the Stanford Achievement Test, seventh edition (also known as the SAT-7), to evaluate claims about critical thinking skills in high school debate students. Barfield identified a total of 300 students from three different private schools in the southeastern U.S. Half of the students had been involved in highly competitive debate programs for at least two years; the other half were non-debate students who were carefully paired to the debate students on the basis of class rank and course loads. Barfield compared the percentile scores of the SAT-7 prior to the debate students "engaging in academic debating" (p. 152) to the percentile scores of the SAT-7 after two years of competitive debate; the percentile scores for the non-debate students were compared for a comparative time period. Barfield also compared the grade point averages (GPAs) of both sets of students. He found a "statistically significant increase" (p.153) in reading comprehension scores. He also found a "definite correlation between active participation in a highly-competitive interscholastic debate program and gains in student GPAs" (p. 158).

Collier (2004)

The second important study in this era of standardized test scores was the study conducted by Collier (2004) on the impact of participation in high school debate on reading scores. Collier administered the Scholastic Reading Inventory (SRI) as a pretest to students, who participated in competitive debate, as well as students, who did not participate in competitive speaking, a total of 421 students, from 22 high schools in five cities. Teachers at each of the schools recommended debate students for the study, as well as students who had not participated in debate for the control group. Collier identified Honors students in both groups. After the debate season was completed, again, Collier administered the SRI to all students. Based on the test scores, Collier concluded that participants in debate scored 25% higher on the reading test than those in the control group and 18% higher than the control subgroup of honors students, which was significant ($p < 0.10$>. Collier suggested that the research requirements of debate motivated students to read and comprehend a wider variety of materials than other students. Collier's findings are especially important because she assessed the reading scores of high school students, as opposed to college students. Additionally, while critical thinking skills are important to society, the focus of state required tests is on reading, writing, mathematics, and science skills.

Vaughn and Winner (2000)

The only other study this researcher found, which was related to the connection between forensics participation by high school students and test scores, was conducted by Vaughn and Winner (2000). They examined the connection between acting and the Scholastic Aptitude Test (SAT) scores. The findings from this study are relevant to this project because acting is very similar to oral interpretation in forensics. Vaughn and Winner based their study on survey responses from students on SAT tests over 12 years and found that the highest SAT scores were achieved by "students taking acting/play production courses" (p. 83). When they examined the component SAT scores of Verbal and Mathematics, the relationship between acting and high test scores was even more evident. While they did not claim a causal relationship, they did find a correlation between participation in acting and higher test scores.

Overall, the findings of many studies (Allen et al., 1995, Allen et al., 1999, Barfield 1989, Brembeck, 1949; Collier, 2004;

Cross, 1971; Howell, 1943; Vaughn & Winner, 2000) have indicated a positive relationship between participation in forensics and academic skills. Most studies have been conducted to investigate the relationship between forensic participation, especially debate, and critical thinking skills. More recently, researchers have begun to study the impact of participation in competitive speech and similar events on standardized test scores.

Criticisms of Recent Research

While the findings from the above studies appeared to demonstrate the positive effects of forensics participation on academic abilities, there were weaknesses in these studies. The greatest weakness found was the chicken/egg dilemma posed by Greenstreet (1993). McGlone (1974, as cited in Greenstreet) wrote, "There is a rather large number of investigations which demonstrate that debate improves certain cognitive abilities and a large body of criticism of these studies which point out that people who have these abilities are simply attracted to debate" (p. 18). Many of the authors of these studies acknowledged this very problem; for example, Whalen (1991) noted, "those who are drawn to debate simply have a tendency to be better critical thinkers" (p. 393). Allen et al. (1999) concurred when they stated,

> Forensic participants are self-selected, and the choice to participate in competitive forensics might be related to higher levels of existing critical thinking. Basically, the claim is that comparisons of forensic participants to nonforensic samples are not a fair comparison because of the bias in self-selection. (p. 20)

In her study of reading scores, Collier (2004) wrote, "more research is warranted. . . to remove the myth of self-selection" (p. 29). Vaughn and Winner (2000) acknowledged the same problem when they wrote, "Alternative explanations include the possibility that students who choose to study the arts are high achievers to begin with" (p. 87). In order for new research projects to be regarded as credible, such projects will have to be designed to avoid the self-selection problem. Another problem with past research on participation in forensics and increased academic skills is that most of these studies were based on data collected from college students. Collier (2004) pointed out that these studies were conducted with college level subjects, who are notably different from high school students. Collier observed, "15 year-olds in urban public high schools can't be compared with college students, particularly those at some of the more elite institutions involved in the debate studies" (p. 7). In her review of literature, this author found only six studies in which the test scores and survey responses of high school students were examined: Howell (1943); Cross (1971, as cited in Korcok); Huseman, Ware, and Gruner (1972, as cited in Greenstreet, 1993); Barfield (1989); Vaughn and Winner (2000); and Collier. As Collier astutely commented, high school students, who are required by law to attend school and take particular courses, are different from college students, who have self-selected both college attendance and particular coursework. In order to meet the requirements of the NCLB (2001), future researchers will have to examine how forensics participation affects the academic skills of high school students if they are to provide evidence to secondary school administrators of the value of competitive speech programs.

Barfield (1989) criticized past studies regarding the positive benefits from debate experience because researchers compared the test scores of students from schools with unequal forensics programs. "In fact, no study has yet collected data which specifically address the quality of instruction received in the debate and non-debate groups" (p. 14). Barfield specifically pointed to Howell's work; Howell found greater improvement in critical thinking skills among students at some schools than students at other schools. Barfield asked, "Could this imply that training in debate can either be 'good' or 'bad' and that the quality of the instruction might bias the outcome of the research?" (p. 14). In order to conduct truly meaningful research about the academic benefits of debate, researchers will need to compare data from students in schools with similar instruction methods and academic priorities for forensic participants.

The final weakness of past research is that the focus has been mostly on the effects of participation in debate. This author found no empirical studies in which the effects of participation in original oratory, extemporaneous speaking, or interpretation of literature were examined. Only a few authors (Carroll, 2007; Crawford, 2003; Hier, 1997; McCrady, 2004; Minch, 2006; Re, 2002; Sellnow, 1994) even mentioned the non-debate events, and those references were limited to the logical connections these events should have on academic skills. Hier, for example, discussed how "speech and debate are almost completely discovery activities. Students select their poetry readings or their prose readings in speech. They select their arguments in debate" (p. 8). McCrady argued, "It's obvious that kids who probe deeply into literature are developing higher order thinking skills" (p. 41), and "logic is taught in extemp, persuasive oratory, and debate" (p. 44). Re mentioned, in passing, that extemporaneous speaking and student congress are events that require knowledge acquisition. Sellnow included oral interpretation as an example of an activity that provides "different ways of knowing for participants" (p. 7). Minch cited a survey of college students, who had participated in individual events, in which they perceived that this experience helped to develop their critical thinking and reading comprehension skills. The problem with such limited research on the non-debate events is that supporters of comprehensive high school forensic programs must be able to justify their entire programs to critical administrators. Also, educators, who need financial support for programs that include the non-debate events, must have empirical findings about the effects of these other events.

Literature Summary

Researchers have explored the positive effects of competitive speech on academic skills since 1943. Since public speaking has been valued for centuries in a variety of cultures, it makes sense that it would have a positive impact on thinking and comprehension skills. Logically, student participation in forensics should increase academic skills, especially critical thinking skills. In light of the current testing requirements, Barfield's (1989) research on how participation in debate leads to higher scores on the SAT-7 and higher GPAs and Collier's (2004) work on how participation in debate leads to higher reading scores are very exciting. However, often, studies about the effects of forensic participation are flawed in terms of the chicken/egg effect (Greenstreet, 1993): the positive results of these studies may be due to the higher abilities and motivation of students who are involved in competitive speech. Additionally, most of the studies have been focused on college students and may not apply to high school students. Finally, the focus of most of the quantitative studies has been exclusively on debate and has ignored the possible benefits of other forensic events. It is hoped that the design of this project will avoid some of these criticisms and add to the credible research on this activity.

METHODS

Currently, the NCLB (2001) requires that only those programs "that have been

demonstrated to be effective through rigorous scientific research" (U.S. Department of Education, 2004, ¶ 1) receive administrative support. While participation in forensics has been linked to increased critical thinking skills, it is essential that studies be designed that use scientific methods to establish the value of this activity in terms of standardized tests, especially those tests required by law. Also, studies need to be designed to avoid the chicken/egg (Greenstreet, 1993) dilemma; in new studies, the researchers must design methods that take into account student motivation and intellectual levels. Finally, in order to meet the NCLB research expectations, new studies must be designed to evaluate the impact of forensics participation on high school students. While studies about college students provide useful information, current laws require studies be conducted that apply to secondary students. This researcher hoped to meet those requirements in this project.

Procedures

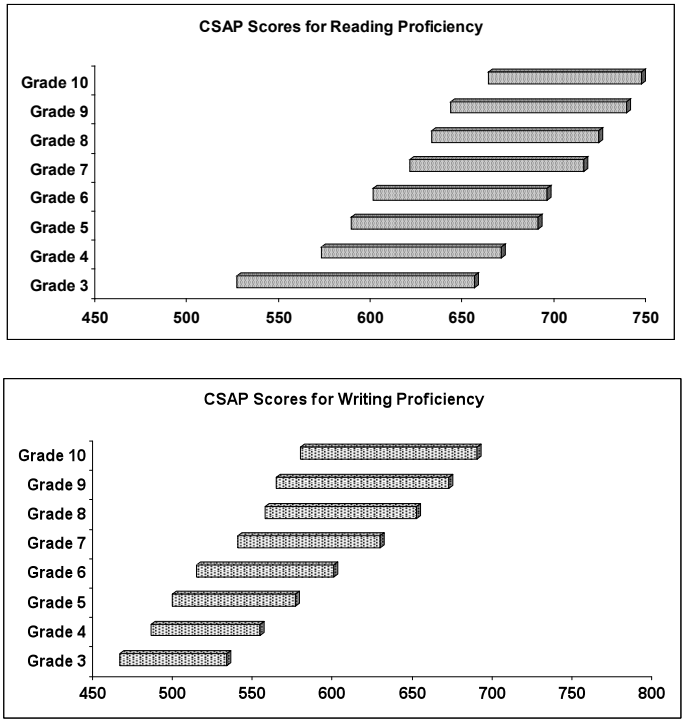
In order to study the possible effects of participation in forensics on standardized test scores, this researcher designed an experiment, based on quantitative data, in order to avoid the deficiencies in other studies. However, since this researcher examined the test scores of students who had chosen, individually, to participate in forensics, or not, as opposed to random assignment of students to the test group and the control group, it was a quasi-experiment, as defined by Korcok (1997). Part of this study was a longitudinal study to examine pretest and posttest scores of state level tests; part of this study was a cross-sectional study to examine the test scores for one national level test.

Instrumentation

The Colorado Student Achievement Program (CSAP; 1997; Colorado Department of Education, 2007) is the required series of tests for students in this state. While there have been challenges to both NCLB (2001) and CSAP, this researcher did not evaluate or justify the use of the CSAP. Since the CSAP was developed by the staff of McGraw-Hill (CTB/McGraw Hill, 2006) to align with the Colorado State Content Standards, the results of this study should be similar to studies conducted in other states that use instruments from the McGraw-Hill for state content standard tests, such as California, District of Columbia, Florida, Mississippi, Missouri, New Jersey, and Oklahoma. CTB/McGraw-Hill tests are used in 23 states and are given to 35 percent of the nation's students (Toch, 2006).

The CSAP (1997) is required of all Colorado students, Grades 3-10. At each grade level, it is comprised of 3 tests each in Reading, Writing, and Mathematics and 2 tests in Science. The CSAP is administered in all Colorado high schools during March each year. This researcher was most interested in the effect of forensics on the Reading and Writing scores. The CSAP provides a unique opportunity to collect pretest and posttest data since the CSAP uses a graduated scoring system: all students, Grades 3-10, receive scores based on a scale of 0-999, and all students are expected to increase their scores each year in order to be deemed proficient, as shown in Figure 1 (CDE, 2007). This continuous scale allows researchers to examine the data as pretests and posttests.

Figure 1. CSAP scores for proficiency rating



Source: Colorado Department of Education (CDE). (2007). CSAP summary data 1997-2005. Denver, CO: Author.

Since officials at the national level have begun to consider the implementation of a growth model of student scores (ED to test NCLB, 2005), whereby schools would be evaluated based on whether individual student test scores increased from year to year, CSAP is a good instrument to use. Additionally, the analysis of gains in student scores, as opposed to a single score, is better aligned with the best practices identified by the National Education Association (NEA, 2005).

In order to evaluate the effects of forensics participation on a nationally normed test, this researcher will examine ACT scores. The ACT is another element of the CSAP (1997) and all eleventh grade students in Colorado are required to take the ACT as the final element of student testing (CDE, 2007). Also, the examination of a nationally normed test will make this study useful to educators and administrators in parts of the country that do not administer tests designed by McGraw-Hill. However, this researcher was not able to determine if participation improves ACT scores since there is no pretest available; thus, this portion of the project is a cross sectional study.

Sample Population

This researcher will examine the test scores of students from Golden High School, a suburban high school in the Denver metropolitan area. Demographically, the school is 87.1% Anglo, 7.9% Hispanic, and 1.3% African American (Jefferson County Public Schools, 2007).

Also, 17.6% of students receive free or reduced lunch.

Participation in the forensics class and program at this school was self-selected. No specific recruiting of particular students was done. The program was open to students in grades 9 through 12 of all academic levels. On average, 35 to 40 students participated in forensics in any given school year; of those students, approximately one-third to one-half were honors students, who took honors level courses in other subjects, and the remainder were average students. All students in the program were expected to take a formal course in forensics for at least one semester; after that semester, students could continue to participate in the forensics class during the regular school day or they could participate in forensics through an independent study program for academic credit. In the regular forensics class, students were introduced to all forms of forensic speaking: debate, original oratory, student congress, and oral interpretation. After this introduction to all the speaking events, students were allowed to choose one event as their focus of study and competition.

In order to avoid the problem of self-selection as defined above, the data to be studied was from students who participated in Honors English 11 during a 4 year period. First, the honors and Advanced Placement (AP) courses at Golden High School were self-selected; students did not test into these classes nor did teachers assign them to these classes. Also, in this author's observations, all of these students participated in extracurricular and cocurricular activities, such as peer counseling/mentoring, student government, student publications, athletic teams, play productions, music performance groups, and other similar extra curricular and cocurricular activities. Additionally, all of these students were involved in multiple honors and AP courses. Thus, the data for all of the students involved in this study represented the test scores of motivated students who self-selected to participate in a variety of school activities and intellectual challenges. In terms of the self-selection, this was a homogenous group of motivated, intellectual students, some of whom chose to participate in forensics and some of whom did not.

The examination of students over several years from a single program should have eliminated several variables that could negate or weaken the conclusions. Variables, such as teaching styles, departmental priorities, and cocurricular vs. extracurricular status of forensics, did not have an impact on this study. All students examined had experienced the same curriculum, the same expectations, and the same grading requirements. Furthermore, the students from the first 3 years that the data were collected had the same English 9 Honors teacher and the same English 10 Honors teacher. During the fourth year studied, faculty changes occurred due to retirements; however, the curriculum and expectations remained the same. Finally, all the participants with forensics experience had the same coach and learned about competitive speech under the same conditions.

Of this group of honors students, the test scores of those students, who did not participate in forensics, were designated as the control group. The test scores of those students, who did participate in forensics, were designated as the test group. The 4 years studied produced CSAP test score data for 205 English 11 Honors students without forensics experience, and 32 English 11 Honors students with forensics experience. Since the scores for the 2007-2008 class of English 11 Honors students were not available until after the due date for this project, only 3 years of data were available for that part of the study: 160 English 11 Honors students without forensics experience, and 24 English 11 Honors students with forensics experience. Since the test group for the analysis of ACT scores was not 30, it was less reliable than the study of CSAP scores.

Within the test group, all students who participated in forensics were considered as a single group when compared to the control group. These students competed in debate, public speaking and oral interpretation events. This researcher did not limit student participation to only debate students. Also, no distinctions were made in regard to the length of participation. As described below, in this project, the researcher compared Grade 8 test scores to Grade 10 test scores. Forensics at this high school was a semester long course; some of the participants were involved for only 1 semester while others were involved for the entire 2 years covered by the testing framework. A complete description of the test group is included in an Appendix. In the future, additional studies can be done to examine the value of the different events or the effect of participation time on test scores.

Data Collection

Since the advent of NCLB (2001), administrators of Jefferson County Public Schools have made the CSAP (1997) scores of each teacher's students available to that teacher. Teachers have been required to use these data, especially CSAP test data, to adjust teaching methods as an element of their evaluations. Thus, all CSAP scores were available to this teacher researcher. Since 11th grade students in Colorado are required to take the ACT (1989), those scores were available to teachers as well.

This researcher examined the scores in both Reading and Writing in Grade 8 and Grade 10. Grade 9 test scores were not used in this analysis for specific reasons. This researcher felt that use of the Grade 9 tests would limit the data to only students who chose to participate in forensics in Grade 9 and eliminate the data of those students who opted to begin forensics in Grade 10, and Grade 9 tests scores could be skewed by the turmoil of freshmen as they adapt to the high school environment.

Only the test scores of students, who have taken all four tests, were included: Reading Grade 8, Reading Grade 10, Writing Grade 8, and Writing Grade 10. The test scores of any student who missed one or more of these tests were not included in this study. Also, only the test scores of students who attended this school for their entire ninth and tenth grade years were included.

Since this researcher was concerned with whether forensics participation has a positive effect on test scores, standard statistical analysis were used on two types of data: test scores and changes in scores. This researcher examined mean scores and standard deviations, and the test of differences of means at the $\alpha < 0.10$ level of significance.

Also, this researcher examined the ACT (1989) scores for each student involved in this study. Since every Colorado student is required to take the ACT in Grade 11, the data was easily accessible. Standard statistical analysis was conducted on the composite scores, as well as scores for the English and Reading portions. Unfortunately, the examination of ACT scores could not include pretesting and posttesting. This researcher examined: (a) mean scores and standard deviations, and (b) the test of differences of means at the $\alpha < 0.10$ level of significance.

Anticipated Results

At the end of this project, it was believed that the positive effects of participation in forensics would translate into increased reading and writing test scores on both the CSAP (1997) and the ACT (formerly known at the American College Test, 1989). To that end, this researcher posited several null hypotheses to be tested.

H_1: There shall be no significant difference in the gains of writing skills of high school students in honors English courses who participated in forensics when compared with the gains of writing skills of high school students in honors English courses as measured by CSAP (1997) scores in 8th grade and 10th grade.

H_2: There shall be no significant difference in the gains of reading skills of high school students in honors English courses who participated in forensics when compared with the gains of reading skills of high school students in honors English courses as measured by CSAP (1997) scores in 8th grade and 10th grade.

H_3: There shall be no significant difference in the scores for reading of high school students in honors English courses who participated in forensics when compared with the scores for reading of high school students in honors English courses as measured by ACT (1989) Reading scores.

H_4: There shall be no significant difference in the scores for English usage and editing of high school students in honors English courses who participated in forensics when compared to the scores for English usage and editing of high school students in honors English courses as measured by ACT (1989) English scores.

H_5: There shall be no significant difference in the gains in writing skills of high school students in honors English courses who participated in debate events when compared with the gains of writing skills of high school students in non-debate events as measured by CSAP (1997) scores in 8th grade and 10th grade.

H_6: There shall be no significant difference in the gains in reading skills of high school students in honors English courses who participated in debate events when compared with the gains of writing skills of high school students in non-debate events as measured by CSAP (1997) scores in 8th grade and 10th grade.

H_7: There shall be no significant difference in the scores for reading of high school students in honors English courses who participated in debate events when compared with the scores for reading of high school students who participated in non-debate events as measured by ACT (1989) Reading scores.

H_8: There shall be no significant difference in the scores for English usage and editing of high school students in honors English courses who participated in debate events when compared with the scores for English usage and editing of high school students who participated in non-debate events as measured by ACT (1989) English scores.

Results

In order to determine the appropriate level of significance for each test, this researcher examined the literature in this discipline. Barfield (1989), Collier (2004), and Howell (1943) established a significance level of $\alpha = 0.10$; thus, this researcher used this established threshold to determine the significance of results. Additionally, this researcher used a one-tail test since previous research indicated that students with forensics experience should have higher scores than students without forensics experience.

H_1: CSAP Writing Scores

The CSAP (1997) scores for writing would indicate that students who participated in forensics had greater gains in writing skills than the students who did not participate in forensics (after applying a trim for extremes), as displayed in Table 1.

Table 1

Summary of CSAP Writing Data with a 5% Trim

	Non-Forensics Participants	Forensics Participants
Number of student scores in study	185	28
Grade 8 CSAP Writing – mean	640.50	634.29
Grade 10 CSAP Writing – mean	663.70	671.82
Change in CSAP Writing – mean	23.19	37.54
Change in CSAP Writing – stand. Dev.	36.47	41.37

This amount of data required the use of the test of differences of means for small samples, which uses the Student's *t* distribution for critical values.

The scores for students who did not participate in forensics are identified as Group 1 and the scores for students who did participate in forensics are Group 2. The calculations for this test revealed a Student's *t* score of $t = 1.906$. This number met the requirement for $\alpha = 0.10$. In fact, this number revealed a significance of $\alpha = 0.030$ for a one-tailed test. Thus, after a 5% trim to reduce the effects of extreme cases, participation in forensics increased CSAP (1997) writing scores at a significant level, and the null hypothesis was rejected.

H_2: CSAP Reading Scores

Displayed in Table 2 are the data for CSAP (1997) reading scores.

Table 2

Summary of CSAP Reading Data with a 5% Trim

	Non-Forensics Participants	Forensics Participants
Number of student scores in study	185	28
Grade 8 CSAP Reading – mean7	16.82	713.18
Grade 10 CSAP Reading – mean	736.89	738.21
Change in CSAP Reading – mean	20.07	25.04
Change in CSAP Reading – stand. dev.	20.58	18.11

The calculations for the test of differences of means for small samples revealed a Student's *t* score of $t = 1.209$. This number did not meet the requirement for $\alpha < 0.10$. This number revealed a significance of $\alpha = 0.11$ for a one-tailed test, which approached the desired significance level but did not achieve it. Thus, participation in forensics did not increase CSAP reading scores at a significant level, and the null hypothesis was accepted, although the reading scores approached the desired significance level.

H_3: ACT Reading Scores

Displayed in Table 3 are the data for ACT (1989) reading scores. The total number of scores studied was less for this test because the class of 2009 had not yet taken the ACT scores; the data were based on 3 years of test scores instead of 4 years.

Table 3

Summary of ACT Reading Data with a 5% Trim

	Non-Forensics Participants	Forensics Participants
Number of student scores in study	150	22
ACT Reading – mean	26.13	27.59
ACT Reading – standard deviation	4.14	4.69

The calculations for the test of differences of means for small samples revealed a Student's t score of $t = 1.517$. This number met the requirement for $\alpha < 0.10$. This number revealed a significance of $\alpha = 0.07$ for a one-tailed test. Thus, after a 5% trim to reduce the effects of extreme cases, participation in forensics did increase ACT reading scores at a significant level, and the null hypothesis was rejected.

H_4: ACT English Scores

Displayed in Table 4 are the data for ACT (1989) English scores.

Table 4

Summary of ACT English Data with a 5% Trim

	Non-Forensics Participants	Forensics Participants
Number of student scores in study	150	22
ACT English – mean	25.60	26.36
ACT English – standard deviation	4.28	5.02

The calculations for the test of differences of means for small samples revealed a Student's t score of 0.7137. This number did not meet the requirement for $\alpha < 0.10$. Thus, after a 5% trim to reduce the effects of extreme cases, participation in forensics did not increase ACT English scores at a significant level and the null hypothesis was accepted.

H5: Debate Students vs. Non-Debate Students and CSAP Writing Scores

Since the test scores examined in this study included scores by students who had debate experience as well as students who participated in only non-debate events, this project provided the author an opportunity to examine whether the differences, or lack of differences, of the various test scores were related to whether the students had debate experience or participated only in the non-debate events. Provided in the Appendix is a description of each student participant in the forensics group. Since all of the current research available attributed gains in critical thinking and reading scores to debate experience, student scores in this study are divided into two categories: (a) students with any debate experience, regardless of the type of debate, alone or in conjunction with participation in other events; and (b) students with no debate experience. Based on this criterion, 20 students were defined as debate students, and 12 students were defined as non-debate students. Because of the small numbers of test scores, the test of differences of means for small samples, which uses the Student's t distribution for critical values, was used. Also, since the number of scores was so small, no trim was used. This small sample examined indicates that this statistical analysis is less reliable than a larger sample.

Table 5

Summary of CSAP Writing Scores for Forensics Participants

	Debate Participants	Non-debate Participants
Number of student scores in study	20	12
Grade 8 CSAP Writing – mean	651.60	642.92
Grade 10 CSAP Writing – mean	690.80	673.08
Change in CSAP Writing – mean	39.20	30.17
Change in CSAP Writing – stand. Dev.	65.56	62.10

The calculations for the test of differences of means for small samples revealed a Student's t score of 0.385. This number did not meet the requirement for $\alpha < 0.10$. Thus, there was no statistically significant difference between the gains in the writing abilities of debate students and the gains in writing abilities of non-debate students, and the null hypothesis was accepted.

H_6: Debate Students vs. Non-Debate Students and CSAP Writing Scores

Displayed in Table 6 are the data for CSAP (1997) Reading scores.

Table 6

Summary of CSAP Reading Scores for Forensics Participants

	Debate Participants	Non-debate Participants
Number of student scores in study	20	12
Grade 8 CSAP Reading – mean	722.50	704.50
Grade 10 CSAP Reading – mean	742.45	734.58
Change in CSAP Reading – mean	19.95	30.08
Change in CSAP Reading – stand. Dev.	21.41	23.05

The calculations for the test of differences of means for small samples revealed a Student's t score of 1.261. This number did not meet the requirement for $\alpha < 0.10$.

This number revealed a significance of $\alpha < 0.11$ for a one-tailed test, which approached the desired significance level but did not achieve it. Thus, there was no statistically significant difference between the improvement in reading between debate students and non-debate students, and the null hypothesis was accepted, although the reading scores approached the desired significance level.

H_7: ACT Reading Scores

Displayed in Table 7 are the data for ACT (1989) Reading scores.

Table 7

Summary of ACT Reading Data for Forensics Participants

	Debate Participants	Non-debate Participants
Number of student scores in study	12	12
ACT Reading – mean	27.50	27.58
ACT Reading – standard deviation	5.21	5.43

It is obvious from the data that there was no significant difference between the ACT Reading scores of debate students and non-debate students. No statistical analysis was needed to accept the null hypothesis.

H_8: ACT English Scores

Displayed in Table 8 are the data for ACT (1989) English scores.

Table 8

Summary of ACT English Data for Forensics Participants

	Debate Participants	Non-debate Participants
Number of student scores in study	12	12
ACT English – mean	25.58	26.92
ACT English – standard deviation	5.95	5.84

The calculations for the test of differences of means for small samples revealed a Student's *t* score of 0.5569. This number did not meet the requirement for $\alpha < 0.10$. Thus, there was no statistically significant difference between the ACT English scores of debate students and the ACT English scores of non-debate students, and the null hypothesis was accepted.

A statistical analysis of the data revealed the following:

1. there was a significant relationship at $\alpha < 0.10$ between students' participation in forensics and greater gains in CSAP (1997) writing scores; in fact, the level of significance was $\alpha = 0.03$;
2. there was no significant relationship at $\alpha < 0.10$ between students' participation in forensics and greater gains in CSAP reading scores, although the results approached significance at the $\alpha = 0.11$ level;
3. there was a significant relationship at $\alpha < 0.10$ between students' participation in forensics and higher ACT (1989) reading scores, as $\alpha = 0.07$;
4. there was no significant relationship at $\alpha < 0.10$ between students' participation in forensics and higher ACT English scores;
5. there was no significant relationship at $\alpha < 0.10$ between students' participation in debate and students' participation in non-debate events in terms of gains in CSAP writing scores;
6. there was no significant relationship at $\alpha < 0.10$ between students' participation in debate and students' participation in non-debate events in terms of gains in CSAP reading scores, although the results approached significance at the $\alpha = 0.11$;
7. there was no significant relationship at $\alpha < 0.10$ between students' participation in debate and students' participation in non-debate events in terms of ACT reading scores, and
8. there was no significant relationship at $\alpha < 0.10$ between students' participation in debate and students' participation in non-debate events in terms of ACT English scores.

Thus, hypotheses H_1 and H_3 demonstrated a statistically significant ($\alpha < 0.10$) relationship between participation in forensics and higher test scores, specifically the CSAP Writing test and the ACT Reading test. Hypothesis H_2 was rejected, and participation in forensics was not significantly linked to higher CSAP Reading scores, although the results approached significance and suggested a relationship. Hypothesis H_4 was accepted; participation in forensics did not significantly affect ACT English scores. In terms of the relationship between forensics students who had debate experience vs. forensics students who had no debate experience, all four hypotheses, H_5, H_6, H_7, and H_8, were accepted; there were no significant differences between the test scores of debate students and students in the non-debate events.

DISCUSSION

The results from this study seemed to confirm the logical association between forensics participation and higher academic achievement, particularly higher standardized test scores. The greatest improvements in test scores were in the CSAP (1997) writing, CSAP reading and ACT (1989) reading tests. These findings seemed logical in light of past research and conjecture. On the other hand, students in forensics did not significantly outscore the control group in terms of the ACT English test. Furthermore, there was no significant difference between the gains by forensics students who studied debate and the forensics students who focused on the non-debate events.

Numerous researchers (Carroll, 2007; Freeley 1990, as cited in Korcok, 1997; Hunt 1994, as cited in Parcher, 1998; Minch, 2006; Parcher; Tumposky, 2004) hypothesized that participation in forensics should lead to greater critical thinking skills. Researchers, such as Allen, Berkowitz, and Louden (1995), Brembeck (1949), Cross (1971, as cited in Colbert, 1995), and Howell (1943), found statistical evidence to suggest that participation in forensics increased critical thinking abilities. Within this framework of previous research, it makes sense that, in this study, forensics students improved their scores on the CSAP (1997) writing test more than non-forensics students. Officials at the Colorado Department of Education (CDE; 2007) explained that at least half of the writing test involves critical thinking abilities; students must demonstrate they can reason, plan, use evidence, defend a hypothesis, and explain their thinking. By writing extended constructed responses and short constructed responses, students have the opportunity to demonstrate their critical thinking skills. In light of the previous research, it makes sense that the greatest gains of students who participated in forensics vs. students who did not participate in forensics would be in the improvement of writing scores. The statistical analysis showed the strongest relationship between participation in forensics and improvement in writing scores; this relationship was found at the $\alpha = 0.03$ level of significance.

When one considers Barfield's (1989) and Collier's (2004) studies, in which a connection was found between debate students and improved reading skills, it is not surprising that a strong connection was found between forensics participants and reading test scores, especially the ACT (1989) reading test. Students in forensics have to read a variety of information carefully. Debate students and oratory students must evaluate pieces of nonfiction for evidence that may help support an argument. Extemporaneous speakers must read a variety of news sources in order to synthesize information into speeches. Interpreters must do intensive literary analysis of their performance pieces in order to understand and portray all the nuances. Reading is a key element of all forensics events; thus, it makes sense that students who participate in forensics would have higher ACT reading scores. This author found a significant relationship at the $\alpha = 0.07$ level. It is a bit puzzling as to why the connection between participation in forensics and improvement in CSAP (1989) reading scores was not as strong. The connection did not meet the requirement for significance, although it approached significance at the $\alpha = 0.11$ level. It is possible that the reading selections on the CSAP were too simplistic to

challenge students and reveal their improvements in reading ability. After all, the CSAP reading questions were designed to test students' comprehension of grade level texts. The reading selections for the ACT were designed to emulate college level texts (ACT, 2007). Since forensics students are accustomed to reading complex texts and must defend intricate interpretations, the type of reading selections on the ACT are closer to the types of reading that forensics participants do.

There was no significant relationship between participation in forensics and ACT (1989) English scores. However, the types of questions asked on the ACT English test do not align well with the types of skills practiced in most forensics courses, particularly in the program studied. According to the ACT Technical Manual (2007), the majority of questions on the ACT English test are designed to assess mechanics: 13% of the questions assess punctuation, 16% of the questions assess grammar and usage, 24% of the questions assess sentence structure, 16% of the questions assess revision strategies, 15% of the questions assess organization of sentences within paragraphs, and 16% of the questions assess style and tone. While most students in forensics must think carefully about organization and style issues, the oral nature of forensics means that students do not have to practice, necessarily, editing skills on a written text; certainly, sentence structures meant to be heard can be different from sentence structures meant to be read. The focus of the ACT English test is on editing skills of written texts. In the forensics course studied in this project, mechanics were never explicitly taught or discussed, especially in terms of editing written texts.

When one considers the previous research in regard to the connection between debate and critical thinking skills, reading skills and academic success, it may surprise some readers that there was no statistical difference between the test scores of forensics students with debate experience and forensics students with no debate experience. All the elements of critical thinking, such as evaluation and organization of information, seem more applicable to the debate events. However, the results from this study suggested that the reading and writing skills used in the non-debate events are as beneficial as the reading and writing skills used in the debate events. Students who compete in Original Oratory or Extemporaneous Speaking must have a thorough understanding of their topics in order to write effective speeches. Students who participate in the interpretation events must use critical thinking skills as well. They have to analyze carefully literary pieces, such as plays or novels, evaluate which parts of the literary work should be included in their performance and which parts should be cut, and evaluate the most effective ways to present the information so that the audience understands and appreciates the nuances of the characters' situations. The results from this study suggested that participation in all the forensics events is equally valuable to student achievement.

The greatest weakness of other studies, noted by scholars (Allen et al, 1999; Collier, 2004; Greenstreet, 1993; Vaughn & Winner, 2000; Whalen, 1991), is the problem of self-selection. Participation in forensics may seem to influence test scores because brighter, more motivated students elect to join forensics teams. However, this author attempted to design a study to minimize the chicken/egg effect. All the test scores examined were from students who were intelligent, motivated, and active in school activities. All the students had plans to attend universities. Moreover, all the students had self-selected into Honors English courses, which emphasized language skills, like reading and writing, and presented challenging materials. All these students were confident enough of their reading and writing abilities to

enroll in advanced English courses. This researcher designed a study that examined two like groups, and this study should mitigate any concerns about the chicken/egg dilemma.

Limitations to This Project and Suggestions for Further Research

One possible weakness of this study is that the examination of test scores by honors students may not be applicable to other students. Because they are more interested in language, they may be more susceptible to the educational effects of forensics. Their interest in complex reading and writing may increase their gains in critical thinking skills while they participate in competitive speech. On the other hand, it is possible that the gains in reading and writing abilities by forensics students could be understated since the scores examined were by honors students instead of grade level students. Allen et al. (1999) discussed the ceiling effect, where students who begin with high test scores cannot improve greatly because their scores are already near the top, or ceiling. If participation in forensics increases reading and writing skills, the increases may have been muted because the honors students did not have much room to improve their scores. Additional studies, which examine the impact of participation in forensics on grade level students, need to be designed to evaluate the impact of the ceiling effect.

Another weakness of this study is the limited number of test scores examined. Without at least 30 test scores for the honors students with forensics experience, the results were less reliable than a larger sample (Brase and Brase, 1999). Furthermore, the examination of test scores by forensics students with debate experience vs. forensics students with no debate experience was limited because there were only a total of 32 forensics students studied: 20 in debate events and 12 in non-debate events. While the results from the comparison between test scores of honors students vs. the test scores of non-forensics students were supported by previous research, the comparison between the test scores of forensics students who had debate experience vs. the test scores of forensics students who competed only in the non-debate events was unique. The statistical connection between the test scores of these two groups was large enough to suggest only a relationship; additional research needs to be done to confirm the academic benefits of the non-debate events.

Conclusion

The value of public speaking, specifically in a competitive setting, has long been recognized. However, in this era of No Child Left Behind (2001) and high stakes testing, it is important that competitive speaking be connected scientifically to higher standardized test scores. Without such research, the future of forensics programs will be doubtful. This researcher found a statistically significant ($\alpha = 0.03$) relationship between participation in forensics and higher writing scores as well as a statistically significant ($\alpha = 0.07$) relationship between participation in forensics and higher reading scores. This connection seems to be equally true for students who choose to compete in the debate events as well as for students who choose to compete in the non-debate events.

During the writing of this project, the political climate began to change. Members of the Colorado House of Representatives proposed bill HB08-1357 (Fender, 2008), which would eliminate writing tests at all grades and eliminate CSAP (1997) tests in Grades 9 and 10. The ACT (1989) would be retained for high school juniors. While the bill passed the House, it was postponed indefinitely in the Senate Education Committee; thus, it was effectively killed (Colorado General Assembly, 2008). However, this event seems to point to the idea that the evaluation of participation in forensics in terms of CSAP scores may be irrelevant. However, two major points need to be made. First, even though the Colorado legislators may be considering revising their testing program, there is no evidence that this is happening in other states. The NCLB (2001) is still in force. CTB/McGraw Hill still produces tests for 23 states (Toch, 2006). While tests of writing skills are not required under NCLB, legislators in 33 states have added a writing component to their testing laws (Toch). Even if the Colorado testing law is altered, the results of this study still would be relevant to other parts of the nation. Second, regardless of political decisions about mandatory testing, this researcher has demonstrated that participation in forensics has a positive effect on reading and writing abilities. Whether they are tested or not, reading and writing have been the cornerstones of education and literacy. To disregard the implications of this project because of political decisions about the testing instruments would be a mistake.

Overall, the value of forensics programs in terms of specific student achievement cannot be denied. While forensics participation is not the only method that school staffs can use to improve their standardized test scores, especially in reading and writing, it is obviously one tool that should be retained in their arsenal of instructional methods to ensure student success.

Tammie Peters has received her Second Diamond for coaching at Golden High School in Colorado. She wrote this paper as a culminating work for her Masters of Arts in Language and Communication from Regis University, Denver.

REFERENCES

ACT. (2007). *About ACT: History of ACT.* Retrieved December 20, 2007, from http: //www.act.org/aboutact/history.html

ACT. (2007). *Technical Manual.* Retrieved April 7, 2008, from http:// www.act.org/aap/pdf/ ACT_Technical_Manual.pdf

Allen, M., Berkowitz, S., Hunt, S., & Louden, A. (1999). A meta-analysis of the impact of forensics and communication education on critical thinking. Communication Education, 48(1), 18-30.

Allen, M., Berkowitz, S., & Louden, A. (1995). A study comparing the impact of communication classes and competitive forensic experience on critical thinking improvement. *The Forensic of Pi Kappa Delta, 81*(1), 1-8.

Barfield, K. D. (1989). *A study of the relationship between active participation in interscholastic debating and the development of critical thinking skills with implications for school administrators and instructional leaders.* Unpublished doctoral dissertation, University of Alabama, Tuscaloosa.

Bellon, J. (2000). A research-based justification for debate across the curriculum. *Argumentation and Advocacy, 36*(3), 162-175.

Brase, C. H. & Brase, C. P. (1999). *Understandable statistics.* New York: Houghton Mifflin.

Brembeck, W. L. (1949). The effects of a course in argumentation on critical thinking ability. *Speech Monographs, 16*(2), 177-189.

Canning, C. (2000). *What was Chautauqua?* Retrieved December 30, 2007, from http: //sdrc.lib.uiowa.edu/traveling-culture/essay.htm

Carroll, R. C. (2007, February). Forensics participation as gifted and talented education. *Rostrum,* 31, 34-36.

Colbert, K. R. (1995). Enhancing critical thinking ability through academic debate. *Contemporary Argumentation and Debate, 16,* 52-72.

Collier, L. M. (2004, November). *Argument for success: A study of academic debate in the urban high schools of Chicago, Kansas City, New York, St. Louis and Seattle.* Paper presented at the National Communications Association Convention, Chicago, IL. Retrieved June 5, 2007, from http://www.bostondebate.org/UMKC %20Study.pdf

Colorado Department of Education (CDE). (2007). *CSAP summary data 1997-2005.* Denver, CO: Author. Retrieved December 1, 2007, from http://www.ce.state.co.us/ cdeassess/documents/csap/csap_summary.html

Colorado General Assemby. (2008). *History for bill number HB08-1357.* Denver, CO: Author. Retrieved June 13, 2008 from http://www.leg.state.co.us/Clics /CLICS2008A/csl.nsf/BillFoldersHouse?openFrameset

Colorado High School Activities Association. (CHSAA). (2007). *Official speech and debate handbook 2007-2008.* Aurora, CO: Author. Retrieved December 18, 2007, from http://chsaa.org/activities/speech/Speech%20Handbook%2008-FINAL.pdf

Colorado Student Assessment Program (CSAP). (1997). Colorado Revised Statutes 22-7-409.

Crawford, R. (2003, November). In defense of competitive speech. *Rostrum,* 19, 21.

CTB/McGraw-Hill (2006). *Colorado Student Assessment Program: Technical report 2006.* Monterey, CA: McGraw-Hill. Retrieved August 15, 2007, from http: //www.cde.state.co.us/cdeassess/documents/reports/2006/Complete_CSAP_2006 _Technical_Report_2006_FINAL.pdf

ED to test NCLB 'growth model.' (2005, November 21). *eSchool News Online,* Article 5980. Retrieved November 22, 2007, from http://www.eschoolnews. com/news /topnews/index.cfm?i=36675&CFID=354222&CFTOKEN=84223472

Fender, J. (2008, March 16). Bill would swap out CSAP for ACT. *Denver Post Online.* Retrieved June 13, 2008 from http://www.denverpost.com/breakingnews /ci_8597240

Fine, G. A. (2001). *Gifted tongues: High school debate and adolescent culture.* Princeton, NJ: Princeton University Press.

Greenstreet, R. (1993). Academic debate and critical thinking: A look at the evidence. *National Forensic Journal, 11*(1), 13-28.

Hier, C. G. (1997, December). Competitive high school speech and debate: An efficient delivery system for education. *Rostrum,* 7-8.

Howell, W. D. S. (1943). The effects of high school debating on critical thinking. *Speech Monographs, 10*(1), 96-103.

International Debate Education Association (IDEA). (2007). *About IDEA.* Retrieved December 1, 2007, from http://www.idebate.org/about/index.php

Jefferson County Public Schools. (2007, July 10). *Golden High School school profile: 2005-2006.* Retrieved November 22, 2007, from http://jeffcoweb.jeffco.k12. co.us /profiles/demographics/high/golden.pdf

Korcok, M. (1997). The effect of intercollegiate debating on critical thinking ability. Retrieved June 5, 2007, from http://mailer.fsu.edu/~ewotring/com5312 /critical.html

Lewis, T. V. (2004). *Communicating literature.* Dubuque, IA: Kendall/Hunt.

Luh, W. (1999, Spring). Developing trimmed mean test statistics for two-way fixed-effects ANOVA models under variance heterogeneity and nonnormality. *Journal of Experimental Education, 67,* 243-265.

Luong, M. A. (2000, November). Forensics and college admissions. *Rostrum,* 5-6.

Massey, M. (1999, November). The role and necessity of equal education and competitive debate in post capitalist/industrial society. *Rostrum,* 24-26.

McCrady, R. (2004, November). Forensics, debate and the SAT. *Rostrum,* 41, 44.

McSpadden, K. (2007, July 28). Fruit is fruit, right? Wrong: No Child Left Behind gives incomplete measure of progress. *The Charlotte Observer* [Electronic version]. Retrieved July 28, 2007, from http://www.charlotte. com/409/story/214691.html

Minch, K. (2006, December). The value of speech, debate, and theater activities. *Rostrum,* 8-13.

National Education Association (NEA). (2005, November). *NCLB 'growth model' plan called step in right direction.* Retrieved November 22, 2007, from http://www .nea.org/esea/growthmodel.html

National Institute of Standards and Technology (NIST). (2003). *Dataplot reference manual.* Retrieved May 4, 2008, from http://www.itl.nist.gov/div898/software /dataplot/refman1/auxillar/trimmecl.htm

National Forensic League. (NFL). (2007). *About the NFL: Training youth for leadership.* Retrieved December 1, 2007, from http://www.nflonline.org/AboutNFL .AboutNFL

No Child Left Behind Act of 2001, 20 U.S.C. §6301 (2001).

Parcher, J. (1998). The co-curricular value of debate. *In Evaluation report: Philodemic Debate Society of Georgetown University.* Retrieved March 1, 2006, from http: //groups.wfu.edu/NDT/Articles/gtreport.html

Partnership for 21st Century Skills. (2004). *Learning and innovation skills.* Retrieved December 1, 2007, from http://www.21stcenturyskills.org

Phillips Academy. (2007). *Answers.com.* Retrieved December 17, 2007 from http://www .answers.com/topic/phillips-academy

Re, R. M. (2002, April). Unique intellectual environment. *Rostrum,* 4, 80.

Rogers, J. E. (2002). Longitudinal outcome assessment for forensics: Does participation in intercollegiate, competitive forensics contribute to measurable differences in positive student outcomes? *Contemporary Argumentation and Debate, 23,* 1-27.

Sellnow, D. D. (1994). Justifying forensic programs to administrators: An experiential education opportunity. *National Forensic Journal, 11*(2), 1-14.

Spino, C., & Pagano, M. (1991, September). Efficient calculation of the permutation distribution of trimmed means. *Journal of the American Statistical Association, 86,* 729-738.

Toch, T. (2006). *Margins of error: The education testing industry in the No Child Left Behind era.* Washington, DC: Education Sector. Retrived June 17, 2008 from http://www.educationsector.org/ research/research_show.htm?doc_id=346734

Trank, D. M. (1977). *Back to the basics: A case for teaching forensics.* Iowa City, IA: University of Iowa. (ERIC Document Reproduction Service No. ED 170 824)

Tumposky, N. R. (2004). The debate debate. *The Clearing House, 78*(2), 52-55.

U.S. Department of Education. (2003). *Questions and answers on No Child Left Behind.* Washington, DC: Author. Retrieved August 15, 2007 from http://www.ed.gov/nclb/methods/whatworks/doing.html

Vaughn, K., & Winner, E. (2000). SAT scores of students who study the arts: What we can and cannot conclude about the association. *Journal of Aesthetic Education, 34*(3/4), 77-89.

Warner, E., & Bruschke, J. (2001). 'Gone on debating:' Competitive academic debate as a tool of empowerment. *Contemporary Argumentation and Debate, 22,* 1-21.

Whalen, S. (1991). Intercollegiate debate as a co-curricular activity: Effects on critical thinking. In D. W. Parson (Ed.), *Argument in controversy: Proceedings of the Seventh SCA/AFA Conference on Argumentation* (pp. 391-397). Annandale, VA: Speech Communication Association.

APPENDIX

Description of Forensics Students Involved in Study

Student	Grad Year	Years of Participation	Description of Particpation (Event and NFL points earned during each year of participation)
1	2006	9th Grade, 2nd Semester	Poetry Interp – 26
		10th Grade – year	Orig Oratory – 71
			Duo Interp – 30
2	2006	10th Grade – year	Poetry Interp – 44
			Duo Interp – 28
3	2006	10th Grade – year	Orig Oratory – 101
			Extemp – 41
4	2006	10th Grade – year	Duo Interp – 139
5	2006	9th Grade – 2nd Semester	Orig Oratory – 30
		10th Grade – year	Duo Interp – 139
6	2007	10th Grade – year	Public Forum – 156
			Extemp – 21
7	2007	10th Grade – year	Policy Deb – 144
8	2007	10th Grade – year	Public Forum – 156
9	2007	9th Grade – year	LD Debate– 216
			Impromptu – 10
			Congress - 64
		10th Grade – year	LD Debate – 87
			Congress – 72
			Humor Interp – 39
10	2007	10th Grade – 1st semester	Duo Interp – 25
			Drama Interp – 7
11	2007	10th Grade – year	Poetry Interp – 31
			Duo Interp – 7
12	2008	9th Grade – year	LD Debate – 21
			Extemp – 11
			Humor Interp - 13
		10th Grade – 1st semester	Humor Interp – 10
13	2008	9th Grade – year	Duo Interp – 25
			Drama Interp – 37
14	2008	10th Grade – year	Duo Interp – 47
			Humor Interp – 20
15	2008	9th Grade – 1st semester	Public Forum – 45
			Extemp – 12
16	2008	9th Grade – 1st semester	Duo Interp – 14
17	2008	9th Grade – 1st semester	Public Forum – 36
			Extemp – 11
18	2008	9th Grade – 1st semester	Public Forum – 69
			Extemp – 9
		10th Grade – year	Duo Interp – 34
			Drama Interp – 10
			Poetry Interp – 21
19	2008	9th Grade – year	Public Forum – 45
			Duo Interp - 40
		10th Grade – 1st semester	Public Forum – 12
			Duo Interp – 49
20	2008	9th Grade – 1st semester	Duo Interp – 51
21	2008	9th Grade – 1st semester	Duo Interp – 25
			Orig Oratory – 9
22	2008	9th Grade – year	Duo Interp – 59
			Humor Interp – 7
		10th Grade – year	Duo Interp – 11
			Congress – 40
23	2008	9th Grade – year	Public Forum - 122
		10th Grade – year	Public Forum – 262
24	2008	9th Grade – year	LD Debate – 45
			Duo Interp - 53
		10th Grade – 1st semester	Duo Interp – 17
25	2009	9th Grade – 1st semester	Policy Deb – 111
		10th Grade – 1st semester	Policy Deb – 79
26	2009	9th Grade – 1st semester	Policy Deb – 36
27	2009	9th Grade – 1st semester	Policy Deb – 66
28	2009	9th Grade – year	Policy Deb – 100
29	2009	10th Grade – 2nd semester	Policy Deb – 49
30	2009	9th Grade – 1st semester	Policy Deb – 27
31	2009	9th Grade – year	Policy Deb – 160
		10th Grade – year	Policy Deb – 94
32	2009	9th Grade – year	Public Forum – 21
			Extemp - 70
		10th Grade – 1st semester	Extemp – 9
			Congress – 7