

Article

Learning to Improve: Using Writing to Increase Critical Thinking Performance in General Education Biology

Ian J. Quitadamo* and Martha J. Kurtz†

*Department of Biological Sciences, Central Washington University, Ellensburg, WA 98926-7537; and

†Department of Chemistry, Central Washington University, Ellensburg, WA 98926-7539

Submitted November 27, 2006; Revised February 16, 2007; Accepted February 19, 2007

Monitoring Editor: W. Bradley Kincaid

Increasingly, national stakeholders express concern that U.S. college graduates cannot adequately solve problems and think critically. As a set of cognitive abilities, critical thinking skills provide students with tangible academic, personal, and professional benefits that may ultimately address these concerns. As an instructional method, writing has long been perceived as a way to improve critical thinking. In the current study, the researchers compared critical thinking performance of students who experienced a laboratory writing treatment with those who experienced traditional quiz-based laboratory in a general education biology course. The effects of writing were determined within the context of multiple covariables. Results indicated that the writing group significantly improved critical thinking skills whereas the nonwriting group did not. Specifically, analysis and inference skills increased significantly in the writing group but not the nonwriting group. Writing students also showed greater gains in evaluation skills; however, these were not significant. In addition to writing, prior critical thinking skill and instructor significantly affected critical thinking performance, whereas other covariables such as gender, ethnicity, and age were not significant. With improved critical thinking skill, general education biology students will be better prepared to solve problems as engaged and productive citizens.

INTRODUCTION

A National Call to Improve Critical Thinking in Science

In the past several years, an increasing number of national reports indicate a growing concern over the effectiveness of higher education teaching practices and the decreased science (and math) performance of U.S. students relative to other industrialized countries (Project Kaleidoscope, 2006). A variety of national stakeholders, including business and educational leaders, politicians, parents, and public agencies, have called for long-term transformation of the K–20 educational system to produce graduates who are well trained in science, can engage intelligently in global issues that require local action, and in general are better able to solve problems and think critically. Specifically, business leaders are calling for graduates who possess advanced analysis and communication skills, for instructional meth-

ods that improve lifelong learning, and ultimately for an educational system that builds a nation of innovative and effective thinkers (Business-Higher Education Forum and American Council on Education, 2003). Education leaders are similarly calling for institutions of higher education to produce graduates who think critically, communicate effectively, and who employ lifelong learning skills to address important scientific and civic issues (Association of American Colleges and Universities, [AACU] 2005).

Many college faculty consider critical thinking to be one of the most important indicators of student learning quality. In its 2005 national report, the AACU indicated that 93% of higher education faculty perceived analytical and critical thinking to be an essential learning outcome (AACU, 2005) whereas 87% of undergraduate students indicated that college experiences contributed to their ability to think analytically and creatively. This same AACU report showed that only 6% of undergraduate seniors demonstrated critical thinking proficiency based on Educational Testing Services standardized assessments from 2003 to 2004. During the same time frame, data from the ACT Collegiate Assessment of Academic Proficiency test showed a similar trend, with

DOI: 10.1187/cbe.06–11–0203

Address correspondence to: Ian J. Quitadamo (iq@cwu.edu).

undergraduates improving their critical thinking less than 1 SD from freshman to senior year. Thus, it appears a discrepancy exists between faculty expectations of critical thinking and students' ability to perceive and demonstrate critical thinking proficiency using standardized assessments (AACU, 2005).

Teaching that supports the development of critical thinking skills has become a cornerstone of nearly every major educational objective since the Department of Education released its six goals for the nation's schools in 1990. In particular, goal three of the National Goals for Education stated that more students should be able to reason, solve problems, and apply knowledge. Goal six specifically stated that college graduates must be able to think critically (Office of Educational Research and Improvement, 1991). Since 1990, American education has tried—with some success—to make a fundamental shift from traditional teacher-focused instruction to more student-centered constructivist learning that encourages discovery, reflection, and in general is thought to improve student critical thinking skill. National science organizations have supported this trend with recommendations to improve the advanced thinking skills that support scientific literacy (American Association for Higher Education, 1989; National Research Council, 1995; National Science Foundation, 1996).

More recent reports describe the need for improved biological literacy as well as international competitiveness (Bybee and Fuchs, 2006; Klymkowsky, 2006). Despite the collective call for enhanced problem solving and critical thinking, educators, researchers, and policymakers are discovering a lack of evidence in existing literature for methods that measurably improve critical thinking skills (Tsui, 1998, 2002). As more reports call for improved K–20 student performance, it is essential that research-supported teaching and learning practices be used to better help students develop the cognitive skills that underlie effective science learning (Malcom *et al.*, 2005; Bybee and Fuchs, 2006).

Critical Thinking

Although they are not always transparent to many college students, the academic and personal benefits of critical thinking are well established; students who can think critically tend to get better grades, are often better able to use reasoning in daily decisions (U.S. Department of Education, 1990), and are generally more employable (Carnevale and American Society for Training and Development, 1990; Holmes and Clizbe, 1997; National Academy of Sciences, 2005). By focusing on instructional efforts that develop critical thinking skills, it may be possible to increase student performance while satisfying national stakeholder calls for educational improvement and increased ability to solve problems as engaged and productive citizens.

Although academics and business professionals consider critical thinking skill to be a crucial outcome of higher education, many would have difficulty defining exactly what critical thinking is. Historically, there has been little agreement on how to conceptualize critical thinking. Of the literally dozens of definitions that exist, one of the most organized efforts to define (and measure) critical thinking emerged from research done by Peter Facione and others in the early 1990s. Their consensus work, referred to as the

Delphi report, was accomplished by a group of 46 leading theorists, teachers, and critical thinking assessment specialists from a variety of academic and business disciplines (Facione and American Philosophical Association, 1990). Initial results from the Delphi report were later confirmed in a national survey and replication study (Jones *et al.*, 1995). In short, the Delphi panel expert consensus describes critical thinking as a “process of purposeful self-regulatory judgment that drives problem-solving and decision-making” (Facione and American Philosophical Association, 1990). This definition implies that critical thinking is an intentional, self-regulated process that provides a mechanism for solving problems and making decisions based on reasoning and logic, which is particularly useful when dealing with issues of national and global significance.

The Delphi conceptualization of critical thinking encompasses several cognitive skills that include: 1) *analysis* (the ability to break a concept or idea into component pieces in order to understand its structure and inherent relationships), 2) *inference* (the skills used to arrive at a conclusion by reconciling what is known with what is unknown), and 3) *evaluation* (the ability to weigh and consider evidence and make reasoned judgments within a given context). Other critical thinking skills that are similarly relevant to science include interpretation, explanation, and self-regulation (Facione and American Philosophical Association, 1990). The concept of critical thinking includes behavioral tendencies or dispositions as well as cognitive skills (Ennis, 1985); these include the tendency to seek truth, to be open-minded, to be analytical, to be orderly and systematic, and to be inquisitive (Facione and American Philosophical Association, 1990). These behavioral tendencies also align closely with behaviors considered to be important in science. Thus, an increased focus on teaching critical thinking may directly benefit students who are engaged in science.

Prior research on critical thinking indicates that students' behavioral dispositions do not change in the short term (Giancarlo and Facione, 2001), but cognitive skills can be developed over a relatively short period of time (Quitadamo, Brahler, and Crouch, unpublished results). In their longitudinal study of behavioral disposition toward critical thinking, Giancarlo and Facione (2001) discovered that undergraduate critical thinking disposition changed significantly after two years. Specifically, significant changes in student tendency to seek truth and confidence in thinking critically occurred during the junior and senior years. Also, females tended to be more open-minded and have more mature judgment than males (Giancarlo and Facione, 2001). Although additional studies are necessary to confirm results from the Giancarlo study, existing research seems to indicate that changes in undergraduate critical thinking disposition are measured in years, not weeks.

In contrast to behavioral disposition, prior research indicates that critical thinking skills can be measurably changed in weeks. In their study of undergraduate critical thinking skill in university science and math courses, Quitadamo, Brahler, and Crouch (unpublished results) showed that critical thinking skills changed within 15 wk in response to Peer Led Team Learning (a national best practice for small group learning). This preliminary study provided some evidence that undergraduate critical thinking skills could be measurably improved within an academic semester, but provided

no information about whether critical thinking skills could be changed during a shorter academic quarter. It was also unclear whether the development of critical thinking skills was a function of chronological time or whether it was related to instructional time.

Numerous studies provide anecdotal evidence for pedagogies that improve critical thinking, but much of existing research relies on student self-report, which limits the scope of interpretation. From the literature it is clear that, although critical thinking skills are some of the most valued outcomes of a quality education, additional research investigating the effects of instructional factors on critical thinking performance is necessary (Tsui, 1998, 2002).

Writing and Critical Thinking

Writing has been widely used as a tool for communicating ideas, but less is known about how writing can improve the thinking process itself (Rivard, 1994; Klein, 2004). Writing is thought to be a vehicle for improving student learning (Champagne and Kouba, 1999; Kelly and Chen, 1999; Keys, 1999; Hand and Prain, 2002), but too often is used as a means to regurgitate content knowledge and derive prescribed outcomes (Keys, 1999; Keys *et al.*, 1999). Historically, writing is thought to contribute to the development of critical thinking skills (Kurfiss, and Association for the Study of Higher Education, 1988). Applebee (1984) suggested that writing improves thinking because it requires an individual to make his or her ideas explicit and to evaluate and choose among tools necessary for effective discourse. Resnick (1987) stressed that writing should provide an opportunity to think through arguments and that, if used in such a way, could serve as a “cultivator and an enabler of higher order thinking.” Marzano (1991) suggested that writing used as a means to restructure knowledge improves higher-order thinking. In this context, writing may provide opportunity for students to think through arguments and use higher-order thinking skills to respond to complex problems (Marzano, 1991).

Writing has also been used as a strategy to improve conceptual learning. Initial work focused on how the recursive and reflective nature of the writing process contributes to student learning (Applebee, 1984; Langer and Applebee, 1985, 1987; Ackerman, 1993). However, conclusions from early writing to learn studies were limited by confounding research designs and mismatches between writing activities and measures of student learning (Ackerman, 1993). Subsequent work has focused on how writing within disciplines helps students to learn content and how to think. Specifically, writing within disciplines is thought to require deeper analytical thinking (Langer and Applebee, 1987), which is closely aligned with critical thinking.

The influence of writing on critical thinking is less defined in science. Researchers have repeatedly called for more empirical investigations of writing in science; however, few provide such evidence (Rivard, 1994; Tsui, 1998; Daempfle, 2002; Klein, 2004). In his extensive review of writing research, Rivard (1994) indicated that gaps in writing research limit its inferential scope, particularly within the sciences. Specifically, Rivard and others indicate that, despite the volume of writing students are asked to produce during their education, they are not learning to use writing to improve their awareness of thinking processes (Resnick,

1987; Howard, 1990). Existing studies are limited because writing has been used either in isolation or outside authentic classroom contexts. Factors like gender, ethnicity, and academic ability that are not directly associated with writing but may nonetheless influence its effectiveness have also not been sufficiently accounted for in previous work (Rivard, 1994).

A more recent review by Daempfle (2002) similarly indicates the need for additional research to clarify relationships between writing and critical thinking in science. In his review, Daempfle identified nine empirical studies that generally support the hypothesis that students who experience writing (and other nontraditional teaching methods) have higher reasoning skills than students who experience traditional science instruction. Of the relatively few noninstructional variables identified in those studies, gender and major did not affect critical thinking performance; however, the amount of time spent on and the explicitness of instruction to teach reasoning skills did affect overall critical thinking performance. Furthermore, the use of writing and other nontraditional teaching methods did not appear to negatively affect content knowledge acquisition (Daempfle, 2002). Daempfle justified his conclusions by systematically describing the methodological inconsistencies for each study. Specifically, incomplete sample descriptions, the use of instruments with insufficient validity and reliability, the absence of suitable comparison groups, and the lack of statistical covariate analyses limit the scope and generalizability of existing studies of writing and critical thinking (Daempfle, 2002).

Writing in the Biological Sciences

The conceptual nature and reliance on the scientific method as a means of understanding make the field of biology a natural place to teach critical thinking through writing. Some work has been done in this area, with literature describing various approaches to writing in the biological sciences that range from linked biology and English courses, writing across the biology curriculum, and directed use of writing to improve reasoning in biology courses (Ebert-May *et al.*, 1997; Holyoak, 1998; Taylor and Sobota, 1998; Steglich, 2000; Lawson, 2001; Kokkala and Gessell, 2003; Tessier, 2006). In their work on integrated biology and English, Taylor and Sobota (1998) discussed several problem areas that affected both biology and English students, including anxiety and frustration associated with writing, difficulty expressing thoughts clearly and succinctly, and a tendency to have strong negative responses to writing critique. Although the authors delineate the usefulness of several composition strategies for writing in biology (Taylor and Sobota, 1998), it was unclear whether student data were used to support their recommendations. Kokkala and Gessell (2003) used English students to evaluate articles written by biology students. Biology students first reflected on initial editorial comments made by English students, and then resubmitted their work for an improved grade. In turn, English students had to justify their editorial comments with written work of their own. Qualitative results generated from a list of reflective questions at the end of the writing experience seemed to indicate that both groups of students improved editorial skills and writing logic. However, no formal measures of

student editorial skill were collected before biology-English student collaboration, so no definitive conclusions on the usefulness of this strategy could be made.

Taking a slightly different tack, Steglich (2000) informally assessed student attitudes in nonmajors biology courses, and noted that writing produced positive changes in student attitudes toward biology. However, the author acknowledged that this work was not a research study. Finally, Tessier (2006) showed that students enrolled in a nonmajors ecology course significantly improved writing technical skills and committed fewer errors of fact regarding environmental issues in response to a writing treatment. Attitudes toward environmental issues also improved (Tessier, 2006). Although this study surveyed students at the beginning and the end of the academic term and also tracked student progress during the quarter, instrument validity and reliability were not provided. The generalizability of results was further limited because of an overreliance on student self-reports and small sample size.

Each of the studies described above peripherally supports a relationship between writing and critical thinking. Although not explicitly an investigation of critical thinking, results from a relatively recent study support a stronger connection between writing and reasoning ability (Daempfle, 2002). Ebert-May *et al.* (1997) used a modified learning cycle instructional method and small group collaboration to increase reasoning ability in general education biology students. A quasi-experimental pretest/posttest control group design was used on a comparatively large sample of students, and considerable thought was given to controlling extraneous variables across the treatment and comparison groups. A multifaceted assessment strategy based on writing, standardized tests, and student interviews was used to quantitatively and qualitatively evaluate student content knowledge and thinking skill. Results indicated that students in the treatment group significantly outperformed control group students on reasoning and process skills as indicated by the National Association of Biology Teachers (NABT) content exam. Coincidentally, student content knowledge did not differ significantly between the treatment and control sections, indicating that development of thinking skill did not occur at the expense of content knowledge (Ebert-May *et al.*, 1997). Interview data indicated that students experiencing the writing and collaboration-based instruction changed how they perceived the construction of biological knowledge and how they applied their reasoning skills. Although the Ebert-May study is one of the more complete investigations of writing and critical thinking to date, several questions remain. Supporting validity and reliability data for the NABT test was not included in the study, making interpretation of results somewhat less certain. In addition, the NABT exam is designed to assess high school biology performance, not college performance (Daempfle, 2002). Perhaps more importantly, the NABT exam does not explicitly measure critical thinking skills.

Collectively, it appears that additional research is necessary to establish a more defined relationship between writing and critical thinking in science (Rivard, 1994; Tsui, 1998, 2002; Daempfle, 2002). The current study addresses some of the gaps in previous work by evaluating the effects of writing on critical thinking performance using relatively large numbers of students, suitable comparison groups, valid and

reliable instruments, a sizable cadre of covariables, and statistical analyses of covariance. This study uses an experimental design similar to that of the Ebert-May *et al.* (1997) study but incorporates valid and reliable test measures of critical thinking that can be used both within and across different science disciplines.

Purpose of the Study

Currently there is much national discussion about increasing the numbers of students majoring in various science fields (National Research Council, 2003; National Academy of Sciences, 2005). Although this is a necessary and worthwhile goal, attention should also be focused on improving student performance in general education science because these students will far outnumber science majors for the foreseeable future. If college instructors want general education students to think critically about science, they will need to use teaching methods that improve student critical thinking performance. In many traditional general education biology courses, students are not expected to work collaboratively, to think about concepts as much as memorize facts, or to develop and support a written thesis or argument. This presents a large problem when one considers the societal role that general education students will play as voters, community members, and global citizens. By improving their critical thinking skills in science, general education students will be better able to deal with the broad scientific, economic, social, and political issues they will face in the future.

The problem addressed by this study was to discover whether writing could improve student critical thinking performance in general education biology courses. How might writing in general education biology affect the analysis, inference, and evaluation skills that are inherent to critical thinking? What level of critical thinking skill do students bring to nonmajors biology courses? Can their critical thinking skills be measurably improved using writing? What other factors affect development of critical thinking skills? When do student critical thinking skills begin to change, and how much? In this study, the effect of writing on critical thinking performance was investigated using the California Critical Thinking Skills Test (CCTST) at the beginning (pretest) and end (posttest) of 10 sections of general education biology at a regional comprehensive university in the Pacific Northwest. Several research questions framed this investigation:

Does writing in laboratory affect critical thinking performance in general education biology?

Does the development of analysis, inference, and evaluation skills differ between students who experience writing versus those who experience traditional laboratory instruction?

What measurable effect do factors like gender, ethnicity, and prior thinking skill have on changes in critical thinking in general education biology?

If critical thinking skills change during an academic quarter, when does that take place?

MATERIALS AND METHODS

Study Context

The study took place at a state-funded regional comprehensive university in the Pacific Northwest. All participants were nonmajor undergraduates who were taking biology to satisfy their general education science requirement. Ten total sections of general education biology offered over three academic quarters (one academic year) were included in the study. Four of the 10 sections implemented a writing component during weekly laboratory meetings ($N = 158$); six traditional quiz-based laboratory sections served as a nonwriting control group ($N = 152$). Only scores from students who had completed both the initial (pretest) and end-of-quarter (posttest) critical thinking assessments were included in the data analysis. A breakdown of participant demographics for the writing and nonwriting groups is provided in Table 1.

Each course section included a lecture component offered four times per week for 50 min and a laboratory component that met once a week for 2 h. Course lecture sections were limited to a maximum enrollment of 48 students, with two concurrent lab sections of 24 students. Two different instructors taught five writing sections and five other instructors taught 11 traditional sections over three consecutive quarters. Each course instructor materially participated in teaching laboratory with the help of one graduate assistant per lab section (two graduate students per course section). None of the instructors from treatment sections had implemented writing in the laboratory before the start of this study. Writing instructors were chosen on the basis of personal dissatisfaction with traditional laboratory teaching methods and willingness to try something new.

Strong efforts were made to establish equivalency between writing and nonwriting course sections a priori. Course elements that were highly similar included common lecture rooms, the use of similar (in most cases identical) textbooks, and a lab facility coordinated by a single faculty member. More specifically, three similarly appointed lecture rooms outfitted with contemporary instructional technology including dry erase boards, media cabinets, a networked computer, and digital projection were used to teach the nonmajors biology courses. The same nonmajors biology textbook was used across the writing and most of the nonwriting sections. All laboratory sections used a common lab facility and were taught on the same day of the week. Although the order in which specific labs were taught differed among sections, a common laboratory manual containing prescriptive exercises covering the main themes of biology (scientific method, cellular biology and genetics, natural selection and evolution, kingdoms of life, and a mammalian dissection) was used across all writing and nonwriting lab sections.

Primary course differences included a writing component in the laboratory, and how much time was devoted to laboratory activities. Those sections that experienced the writing treatment completed the prescriptive lab exercises in the first hour and engaged in writing during the second hour of the lab. Nonwriting sections allocated 2 h for the prescriptive lab exercises and included a traditional laboratory quiz rather than a writing assignment. The degree to which the writing and nonwriting sections included small group collaboration in laboratory varied and all course sections differed with regards to individual instructor teaching style. Although all course sections used traditional lecture exams during the quarter to assess content knowledge, the degree to which rote memorization-based exam questions were used to evaluate student learning varied.

Description of the Writing Treatment

On the first day of lecture, students in the writing treatment group were told that their laboratory performance would be evaluated using collaborative essays instead of traditional quizzes. A brief overview of the writing assignments was included in associated course syllabi. During the first laboratory session of the quarter, students were grouped into teams of three or four individuals, and the criteria for completing weekly writing assignments were further explained.

The decision to use collaborative groups to support writing in the laboratory was partly based on existing literature (Collier, 1980; Bruffee, 1984; Tobin *et al.*, 1994; Jones and Carter, 1998; Springer *et al.*, 1999) and prior research by Quitadamo, Brahler, and Crouch (unpublished results), who showed that Peer Led Team Learning (one form of collaborative learning) helped to measurably improve undergraduate critical thinking skills. Small group learning was also used in the nonwriting treatment groups to a greater or lesser extent depending on individual instructor preference.

Baseline critical thinking performance was established in the academic quarters preceding the writing experiment to more specifically attribute changes in critical thinking to the writing treatment. Concurrent nonwriting course sections were also used as comparison groups. The historical baseline provided a way to determine what student performance had been before experiencing the writing treatment, whereas the concurrent nonwriting groups allowed for a direct comparison of critical thinking performance during the writing treatment. Pretest scores indicating prior critical thinking skill were also used to further establish comparability between the writing and nonwriting groups.

Laboratory activities were coordinated for all sections by a single faculty member who taught in the nonwriting group. All faculty and graduate assistants met regularly to discuss course progress,

Table 1. Demographics for the writing and nonwriting groups

Sample	Class distribution (%)					Gender distribution (%)		
	Fr	So	Jr	Sr	2nd Sr	M	F	
Writing (158)	44.9	33.5	15.2	3.8	2.5	38.6	61.4	
No writing (152)	53.3	28.3	7.2	9.2	2.0	38.2	61.8	
Overall (310)	49.0	31.0	11.3	6.5	2.3	38.4	61.6	
Sample	Ethnic distribution (%)						Asian	Other ^a
	Caucasian	Hispanic	African American	Native American				
Writing (158)	84.8	1.9		2.5	0	4.4	6.3	
No writing (152)	81.6	4.6		1.3	1.3	5.9	5.3	
Overall (310)	83.2	3.2		1.9	0.6	5.2	5.8	

Demographics profile for the study sample. n values in parentheses.

^a Other includes the "choose not to answer" response.

laboratory procedure, and coordinate resources. Nonwriting faculty drafted quizzes that addressed laboratory content knowledge. Writing faculty collaboratively crafted a consensus essay, or thought question, designed to elicit student critical thinking and ability to apply content knowledge. Each thought question was designed so that students had to apply lecture concepts and build on their conceptual understanding by integrating actual laboratory experiences (see Supplemental Appendix 1, available online) for thought question examples). Weekly thought questions became progressively more difficult as the term progressed. Initial planning meetings took place just before the beginning of the academic quarter and included graduate assistant training to help them learn to consistently evaluate student writing using a modified thesis-based essay rubric (see Supplemental Appendix 2; Beers *et al.*, 1994). A range of sample essays from poor to high quality was used to calibrate graduate assistant scoring and ensure consistency between assistants from different laboratory sections within the writing group. All graduate assistants and course instructors applied the thesis-based rubric to sample essays and worked toward consensus. Initial training ended when all graduate assistants scored within 0.5 points of each other on at least two sample essays.

Students were given weekly thought questions before beginning laboratory to help them frame their efforts during laboratory exercises. Students completed the prescriptive lab activities during the first hour, and then each student group relocated to an assigned computer lab in the same building and worked around a common computer terminal to draft a collective response to the weekly thought question. Students were allowed to use any suitable information or materials (laboratory observations, laboratory manuals, lecture notes, textbooks, the Internet, etc.) to help them address their thought question. Internal group discussions allowed students to argue individual viewpoints as they worked toward group agreement on each thought question. Essay responses to thought questions were answered using a standard five-paragraph format. Each essay included an introduction with a group-generated thesis statement, two to three body paragraphs that provided sufficient detail to support the thesis statement, and a summary paragraph that concluded the essay. Students were not allowed to work on essays outside of the laboratory environment.

Initial essay drafts were composed in Microsoft Word and submitted to the graduate assistant by the end of the laboratory period using the campus e-mail system. Graduate assistants evaluated each group's essay (typically six per lab section) and assigned an initial grade based on the thesis-based essay rubric. Graduate assistants made comments and suggestions electronically using Microsoft Word revising and track changes tools. Evaluated essays were e-mailed back to each student group, which addressed comments and suggestions during the subsequent week's laboratory writing time. Each student group submitted a final draft that was re-evaluated and assigned a final grade. During the second week, students both revised their essay from the previous week and then generated an initial draft for the current week's thought question, all within the lab writing hour. This was done to help students become more proficient writers within a short period of time. Overall, students in the writing group completed eight essays that, along with lab book scores, constituted 25% of their overall course grade. An identical percentage was used to calculate traditional quiz and lab book scores in all nonwriting course sections.

At the end of the quarter, each writing group member completed a peer evaluation for all group members, including themselves (see Supplemental Appendix 3). This was done to help students reflect on and evaluate their own performance, maximize individual accountability within the group, and make sure students received credit proportional to their contributions. The average peer evaluation score for each student was included as 5% of the final course grade.

Collectively, this approach to writing and evaluation was used to 1) help students reflect on and discuss deficiencies in their collective and written work, 2) provide an opportunity for students to explicitly address deficiencies in thesis development and general writing

skill, 3) provide a suitable reward for student efforts to revise their work relative to established performance benchmarks, 4) improve individual accountability within each group, and 5) help students develop more efficient and effective writing skills that collectively might lead to improved critical thinking skill.

Assessment of Critical Thinking

Using critical thinking to indicate student learning performance is particularly useful because it can be measured within and across disciplines. Various instruments are available to assess critical thinking (Watson and Glaser, 1980; Ennis and Weir, 1985; Facione, 1990b; Center for Critical Thinking and Moral Critique, 1996); however, only the CCTST measures cognitive and meta-cognitive skills associated with critical thinking, is based on a consensus definition of critical thinking, and has been evaluated for validity and reliability for measuring critical thinking at the college level (Facione, 1990a; Facione *et al.*, 1992, 2004). The CCTST measures cognitive skills of analysis, inference, evaluation, induction, and deduction, with results expressed as raw scores or national percentile equivalents based on a national norming sample of students from 4-yr colleges and universities. Construct validity for the CCTST is high as indicated by greater than 95% consensus of the Delphi panel experts on the component skills of critical thinking. Test reliability (calculated using the KR-20 internal consistency method) is 0.78–0.84 for the form used in this study, a value considered to be within the recommended range for tests that measure a wide range of critical thinking skills (Facione, 1991). The CCTST norming sample for 4-yr colleges and universities is based on a stratified sample of 2000 students from various disciplines, with approximately 30% of the norming sample comprised of science and math students. Approximately 20,000 college students complete the CCTST each year (Insight Assessment and Blohm, 2005).

The CCTST contains 34 questions and is a 45-min timed assessment of critical thinking. An online version of the CCTST was administered in this study, which allowed the researchers to collect student demographics data including gender, ethnicity, age, and several others at the same time critical thinking skill was measured. Total critical thinking skill as well as analysis, inference, and evaluation component critical thinking skills (Facione, 1990c) were determined for each CCTST administration and compared across the writing and nonwriting groups.

Research Design

A quasi-experimental pretest/posttest control group design was used for this study to determine whether critical thinking performance in the writing group differed significantly from the nonwriting group. This design was chosen in order to compare critical thinking performance between intact groups, and because it was not feasible to randomly assign students from one course section to another within the sample. Frequency distributions of pretest/posttest changes in total critical thinking skill and analysis, inference, and evaluation component critical thinking skills were constructed to provide some indication of sample randomness and to inform assumptions for subsequent statistical analyses of covariance (see Figure 1, A–D).

The pretest/posttest control group design was also used in order to minimize internal validity threats that could potentially compete with the effects of the writing treatment on student critical thinking performance. This design is widely used in educational research, and generally controls for most threats to internal validity (Campbell and Stanley, 1963). Internal threats that remain a concern include history, maturation, pretest sensitization, selection, and statistical regression toward the mean. In the current study, history and maturation threats were minimized to the extent that the CCTST pretest and posttest were administered only 9 wk apart, and class standing and age covariables that indicate maturation were included in the statistical analysis. Pretest sensitization and selection are larger concerns for this design. Pretest sensitization was

minimized in several ways: 1) prior critical thinking skill indicated by the CCTST pretest was used as a covariable in statistical analyses, 2) pretest/posttest to posttest only comparison studies conducted by Insight Assessment indicate CCTST pretest sensitization is minimized (Facione, 1990a), and 3) neither the students, instructors, nor the test administrators have access to the correct answers on the CCTST, so repeat performance on the posttest is less likely. Selection threats were also reduced by using CCTST pretest scores in the statistical analyses, thereby making it more difficult to detect statistically significant differences in critical thinking performance between the writing and nonwriting groups. Statistical regression toward the mean, which was observed to some extent in this study, was minimized because this study used a valid and reliable instrument to assess critical thinking (Facione, 1990a). Regression threats were also minimized to the extent that students with higher initial scores regressed much less than students with lower initial scores.

The generalizability of study results is limited because all data were collected at a single university. Specific threats to external validity include selection-treatment interaction and treatment diffusion. These threats were minimized because writing was mandatory for all treatment group participants, thereby minimizing volunteer effects. Because the writing also took considerable student effort, it is less likely that treatment diffusion occurred. In summary, the pretest/posttest control group design was used to minimize internal and external validity threats and maximize the ability to determine the effects of writing on student critical thinking performance.

Study Variables and Data Analysis

Effect of Writing on Critical Thinking Performance. General education biology students were divided into writing and nonwriting groups (independent variable). Changes in CCTST pretest/posttest scores (dependent variable) were determined to discover whether writing influenced student critical thinking performance. Two CCTST outcome measures were used to statistically test for writing effect: 1) raw scores for total critical thinking skill, and 2) raw scores for analysis, inference, and evaluation component skills. Results were reported using raw scores and corresponding national percentile rank so that critical thinking performance outcomes would be more meaningful and intuitive. Conversion of CCTST raw scores to national percentile ranking was done using SPSS (SPSS, Inc., Chicago, IL) statistical software and a linear estimation conversion script based on an equivalency scale from Insight Assessment (Millbrae, CA).

Several covariables were included in the analysis to increase statistical accuracy and precision, and to more specifically isolate the effects of writing on critical thinking performance. CCTST pretest scores were used to indicate initial critical thinking skill. Gender and ethnicity helped to account for male/female or race-specific changes in critical thinking performance and were also used to identify potential sources of performance bias. Academic term and time of day were used to account for critical thinking differences due to the time of year each course was offered and the time of day each student took the course, respectively. Class standing and age were used to indicate maturation related to time in college and chronological age, respectively. Finally, the instructor covariable was used to account for performance differences due to individual teaching styles.

Statistical Analysis of Effect of Writing. Several statistical analyses were conducted to determine the effects of writing on critical thinking performance in general education biology. An analysis of covariance (ANCOVA) test provided insight regarding differences in overall critical thinking performance between the writing and nonwriting groups. Change in CCTST total raw scores and national percentile ranking was used as composite measures of critical thinking (Facione, 1990c) in this initial analysis. Second, changes in particular component critical thinking skills (analysis, inference, and evaluation) were evaluated using a multivariate analysis of covariance (MANCOVA) test because of the three dependent

variables. The ANCOVA and MANCOVA tests also provided some insight into the effect the covariables had on critical thinking performance in general education biology. Collectively, these statistical tests allowed for a more accurate and precise analysis because variance associated with the covariables could be more specifically isolated from the writing treatment. Mean, SE, and effect size were also compared between the writing and nonwriting groups. Effect size, represented in standard units, was used to compare the magnitude of writing effect in the study.

Analysis of Thought Question Performance. Performance on weekly thought questions was analyzed to discover specifically when and how much student critical thinking skills changed during the academic term. This analysis also provided context for CCTST critical thinking performance measures. Specifically, average scores from a representative sample of writing course sections (approximately 100 students) were used to compare initial essay drafts across the weeks of the term to discover when students began to show changes in their first attempt at each essay. Weekly performance on final revised essays was also compared to determine how student final submissions changed over time. Finally, the weekly difference between each initial essay and each final essay was compared to determine how much the revision process changed during the term. These calculations collectively helped to provide a profile of critical thinking performance over time.

RESULTS

Participant Demographics

Student demographics provided in Table 1 indicated an overall distribution of approximately 49% freshmen, 31% sophomores, 11% juniors, and 9% seniors. Approximately 74% of the writing group students were freshmen and sophomores, whereas 82% of the nonwriting group was underclassmen. Overall, 61% of the sample was female and 39% male, with near identical gender distribution across the writing and nonwriting groups. The predominant ethnicity in the sample was Caucasian (>83%), with Asian American (5%), Latino/Hispanic (3%), African American (2%), and Native American (1%) students comprising the remainder of the sample. About 6% of the sample classified themselves as having some other ethnicity or chose not to identify their ethnic heritage.

Statistical Assumptions

Analysis of covariance and multivariate analysis of covariance tests were used to compare critical thinking performance between the writing and nonwriting groups. The evaluated assumptions for the ANCOVA and MANCOVA tests were homogeneity of slopes, homogeneity of covariances, and normality. An analysis evaluating the homogeneity of slopes assumption indicated that the relationship between the covariables and the critical thinking performance dependent variable did not differ significantly by the writing/nonwriting independent variable for the ANCOVA test, $F(2, 307) = 1.642, p = 0.195$, power = 0.346, partial $\eta^2 = 0.011$, or the MANCOVA test, $F(6, 610) = 1.685, p = 0.122$, power = 0.645, partial $\eta^2 = 0.016$. These results confirmed that both analyses of covariance met the homogeneity of slopes assumption. The homogeneity of covariance assumption was tested using Levene's and Box's tests. Levene's test results for the ANCOVA indicated that error variances were not equal across writing and nonwriting groups,

$F(1,308) = 7.139, p = 0.008$. Similarly, Box's test results indicated that covariance was not equal for the writing and nonwriting groups, $F(6, 684,530) = 4.628, p = 0.000$. These results indicated that the ANCOVA/MANCOVA tests did not meet the homogeneity of covariance assumption. To more fully evaluate this assumption, distributions of total and component critical thinking skill were constructed (see Figure 1, A–D). Furthermore, the writing and nonwriting groups were highly similar in size and no post hoc tests were conducted. On the basis of these data, it was determined that the ANCOVA and MANCOVA tests were the best statistical measures to answer the research questions. Finally, the normality assumption was evalu-

ated using the previously constructed frequency distributions for total change in critical thinking (Figure 1A) as well as change in analysis (Figure 1B), inference (Figure 1C), and evaluation (Figure 1D) critical thinking skills. Frequency distributions of total and component critical thinking dependent variables indicated that each approximated a standard normal curve.

Effect of Writing on Total Critical Thinking Performance

The ANCOVA test of total critical thinking performance showed that writing and nonwriting groups differed significantly,

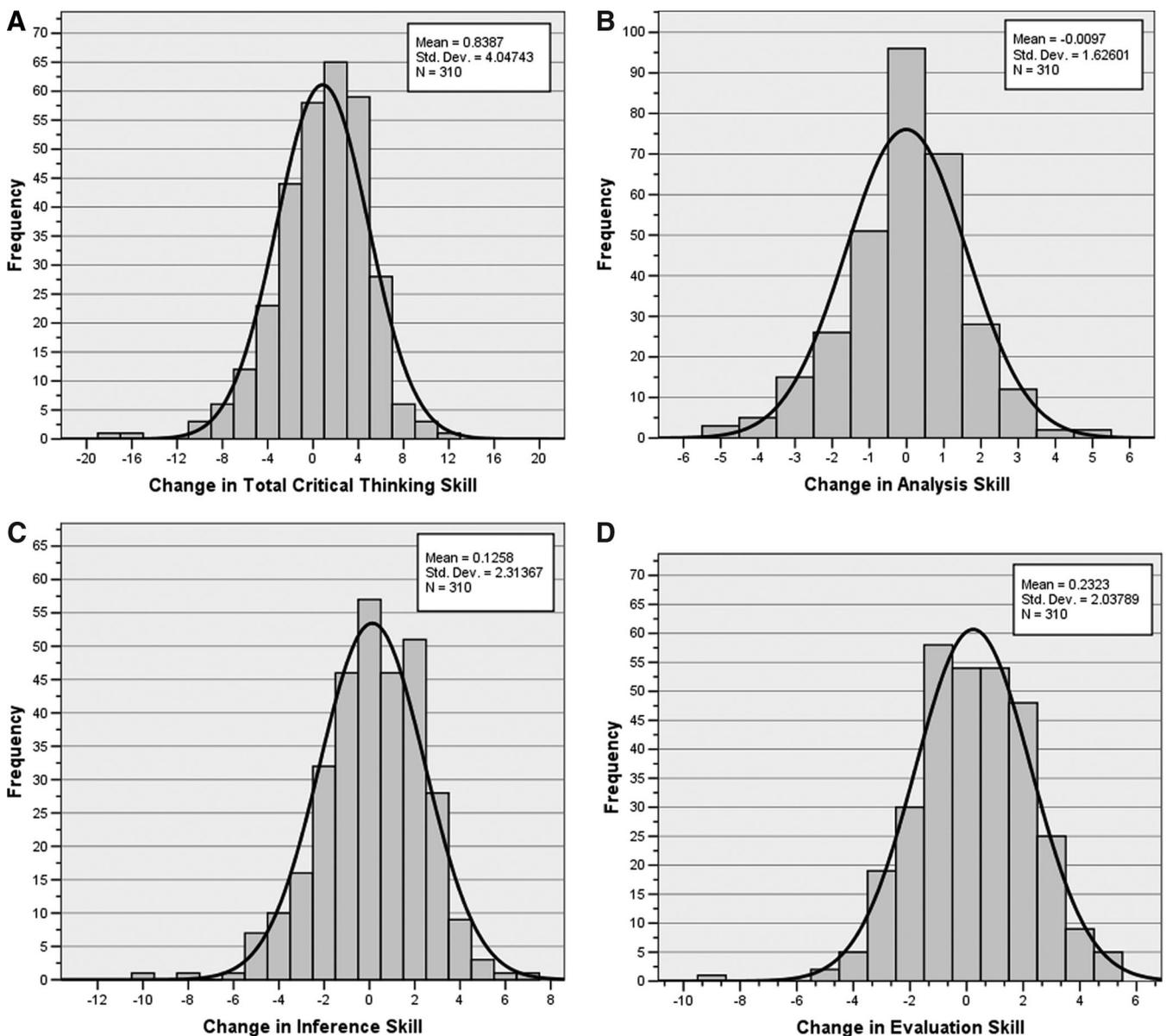


Figure 1. (A–D) Frequency distribution of change in critical thinking skills. Distribution of change in critical thinking skill for the experimental sample. Changes are indicated using raw scores from CCTST pre- and posttests for total critical thinking skill (A) as well as analysis (B), inference (C), and evaluation (D) component critical thinking skills.

Table 2. ANCOVA results for total critical thinking performance

Treatment	F	df	P	Power	Effect size
Writing	19.357	300	0.000 ^a	0.992	0.061
CCTST pretest	19.713	300	0.000 ^a	0.993	0.062
Instructor	7.745	300	0.006 ^a	0.792	0.025
Time of day	6.291	300	0.013 ^a	0.705	0.021
Gender	0.226	300	0.635	0.076	0.001
Ethnicity	2.326	300	0.128	0.330	0.008
Age	0.453	300	0.502	0.103	0.002
Class standing	0.002	300	0.962	0.050	0.000
Academic term	2.387	300	0.123	0.338	0.008

Analysis of covariance for the writing and nonwriting groups. Tested covariables included gender, ethnicity, class standing, age, prior critical thinking skill (CCTST pre-test), academic term, time of day, and instructor.

^a Significance tested at 0.05 level.

$F(1, 300) = 19.357, p < 0.0001, \text{power} = 0.992, \text{partial } \eta^2 = 0.061$ (see Table 2). The strength of the relationship between the writing/nonwriting groups and critical thinking performance was modest but significant, accounting for more than 6% of the variance in critical thinking performance.

Descriptive statistics of total critical thinking performance in the writing and nonwriting groups were also calculated (see Table 3). The writing group showed an average CCTST raw score change of 1.18 compared with the nonwriting group, which showed an average raw score change of -0.51 . These critical thinking raw scores equated to gains in national percentile rank of 7.47 (45th to 53rd percentile) for the writing group and -2.09 (42nd to 40th percentile) for the nonwriting group. Critical thinking improvement in the writing group was approximately nine times greater than the nonwriting group (see Figure 2).

The ANCOVA test of total critical thinking skill indicated that gender, ethnicity, age, class standing, and academic term did not significantly affect critical thinking performance (see Table 2). Covariables that significantly affected total critical thinking performance included 1) CCTST pretest score, $F(1, 300) = 19.713, p < 0.0001, \text{power} = 0.993, \text{partial } \eta^2 = 0.062$, 2) instructor, $F(1, 300) = 7.745, p < 0.006, \text{power} = 0.792, \text{partial } \eta^2 = 0.025$, and 3) time of day, $F(1300) = 6.291, p < 0.013, \text{power} = 0.705, \text{partial } \eta^2 = 0.021$.

The effect of prior critical thinking skill (CCTST pretest) was moderately strong, accounting for more than 6% of the variance in total critical thinking performance. The effect of instructor and time of day were smaller, accounting for 2.5 and 2%, respectively, of total critical thinking performance variance. Critical thinking improvement associated with CCTST pretest score was approximately 2.5 times greater than for instructor and nearly three times greater than for time of day.

Effect of Writing on Component Critical Thinking Performance

The MANCOVA test indicated that analysis, inference, and evaluation critical thinking skills differed significantly between the writing and nonwriting groups, Wilks $\lambda = 0.919, F(3, 296) = 8.746, p < 0.0001, \text{power} = 0.995, \text{partial } \eta^2 = 0.081$ (see Table 4). The strength of the relationship between writing and component critical thinking performance was modest but significant, accounting for more than 8% of the variance in critical thinking performance.

Specifically, significant gains in analysis and inference skills were observed in the writing group but not the nonwriting group. No statistically significant gains in evaluation skill were observed in either group (see Table 5). National percentile rank equivalents for CCTST component raw scores indicated the writing group gained 10.51 percentile in analysis skill (42nd to 52nd percentile), 6.05 percentile in inference skill (45th to 52nd percentile), and 5.16 percentile in evaluation skill (46th to 52nd percentile). The nonwriting group showed a national percentile rank change of -4.43 percentile in analysis skill (47th to 42nd percentile), -2.23 percentile in inference skill (42nd to 40th percentile), and 1.37 percentile in evaluation (44th to 45th percentile; see Figure 3). Critical thinking performance for the writing group was 15 times greater for analysis and 8 times greater for inference skills than for the nonwriting group. Although neither the writing nor the nonwriting group showed significant gains in evaluation skill, the writing group showed more than 3 times greater improvement than did the nonwriting group.

The MANCOVA test of analysis, inference, and evaluation skills indicated that gender, ethnicity, age, class standing, academic term, and time of day did not significantly affect critical thinking performance. Critical thinking performance was affected by prior analysis, inference, and evaluation skill (CCTST component pretest scores)

Table 3. Writing effect on total critical thinking performance: CCTST raw scores

Treatment	Mean raw score		SEM		Raw CT change
	Pre	Post	Pre	Post	
Writing (158)	15.84	17.02	0.32	0.36	1.18 ^a
Nonwriting (152)	15.46	14.95	0.34	0.43	-0.51
Overall (310)	15.65	16.00	0.33	0.40	0.34

Comparison of writing and nonwriting group performance based on CCTST raw scores. CCTST raw score range was 0–34; n values in parentheses.

^a Significance tested at 0.05 level.

and instructor (see Table 4). Specifically, component pre-test scores had a large effect on critical thinking, accounting for 38% (analysis), 32% (inference), and 39% (evaluation) of critical thinking performance variance. The effect of instructor was smaller, accounting for 4.4% of variation in critical thinking skill. The effect of prior component critical thinking skill was approximately 4.5 times greater than the effect of writing, and nearly 9 times greater than the effect of instructor.

Student Thought Question Performance

Critical thinking performance on student essays was evaluated by applying a thesis-based essay rubric (see Supplemental Appendix 2) on initial submissions and final revised essays. Average weekly performance during the academic term is shown in Figure 4. A comparison of initial essays indicated that students improved 53.3% from week 1 (average score of 27.9%) to week 7 (average score of 81.2%). A similar comparison of final essays showed that students improved 32.5% from week 1 (average score of 54.1%) to week 7 (average score of 86.6%). The largest changes between initial and final essays occurred in week 1 (change of 26.2%), and decreased each week thereafter (24.8, 23.9, 18.8, 8, 7.8, and 5.4% for weeks 2 through 7, respectively). These results showed that students produced little evidence of critical thinking skill in their writing early in the term, but improved dramatically on both initial and revised essay submissions by the end of the term.

DISCUSSION

The purpose of this study was to discover whether writing could measurably influence critical thinking performance in general education biology. Results indicated that students from the writing group significantly outperformed their nonwriting peers in both total critical thinking skill and the

component critical thinking skills of analysis and inference. The writing and nonwriting groups were highly similar initially and began the academic term with comparable critical thinking ability (45th and 42nd national percentile for writing and nonwriting, respectively). By the end of the term, writing students had improved their critical thinking skill to above the 52nd percentile whereas nonwriting students decreased to below the 40th percentile. In addition to writing, prior critical thinking skill and course instructor significantly affected critical thinking performance, with prior critical thinking skill having the largest effect on critical thinking gains of any variable tested. Further analysis of the writing group showed that the largest gains in critical thinking occurred during the first few weeks of the term, with graduated improvement during the remainder of the term. A comparison of average critical thinking performance on initial essays and revised essays showed that thinking skills improvement was greater on initial essays (53%) than on final essays (33%). Collectively, the results of this study indicated that students who experienced writing in general education biology significantly improved their critical thinking skills.

The covariance analysis that was conducted provided a partial means to separate out the effects of writing, prior critical thinking skill, instructor, and multiple covariables from total and component critical thinking gains. The analysis of total critical thinking skill indicated that writing students changed their critical thinking skill from below the national average to above the national average within an academic quarter, whereas nonwriting students remained below the national average. This observation is important because it shows that students can develop critical thinking skills within a fairly short 9-wk period of time, and that writing can play a role in that process. A similar study showed critical thinking skills improve over 15 wk (Quitadamo, Brahler, and Crouch, unpublished results); however, this study provided no insight into whether critical thinking skills could be changed over a shorter period of time, in a

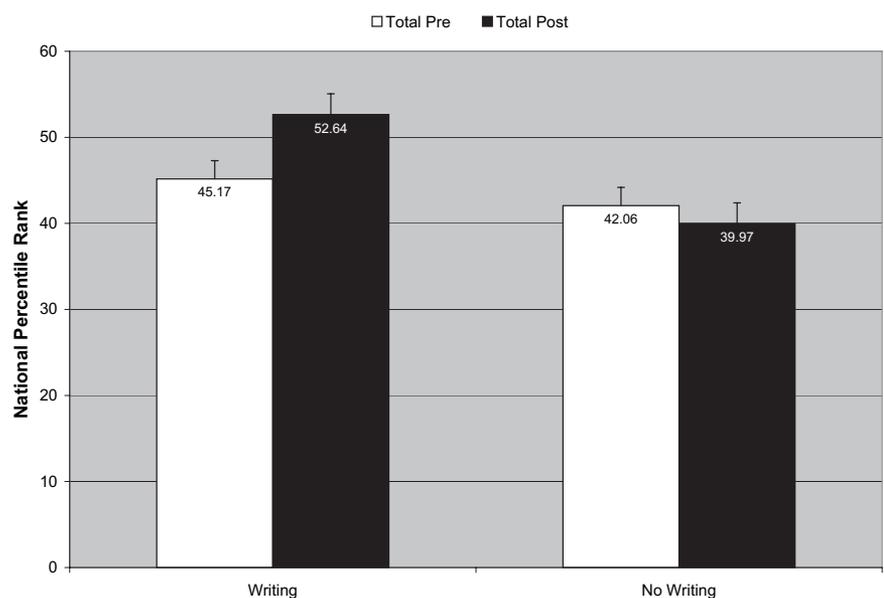


Figure 2. Effect of writing on total critical thinking national percentile rank. Comparison of total critical thinking national percentile gains between writing and nonwriting groups. Percentile ranking was computed using CCTST raw scores, an equivalency scale from Insight Assessment, and a linear conversion script in SPSS.

Table 4. MANCOVA results for component critical thinking performance

Treatment	Wilks λ	F	df	P	Power	Effect size
Writing	0.919	9.746	296	0.000	0.995	0.081
Analysis pretest	0.623	59.737	296	0.000	1.000	0.377
Inference pretest	0.681	46.222	296	0.000	1.000	0.319
Evaluation pretest	0.613	62.398	296	0.000	1.000	0.387
Gender	0.984	1.602	296	0.189	0.420	0.016
Ethnicity	0.983	1.756	296	0.156	0.456	0.017
Age	0.988	1.153	296	0.328	0.456	0.012
Class standing	0.978	2.186	296	0.090	0.553	0.022
Instructor	0.956	4.508	296	0.004	0.880	0.044
Quarter	0.991	0.899	296	0.442	0.246	0.009
Time of day	0.980	2.022	296	0.111	0.517	0.020

Multivariate analysis of covariance for the writing and nonwriting groups. Tested covariables included gender, ethnicity, class standing, age, prior critical thinking skill (CCTST pretest), academic term, time of day, and instructor.

^a Significance tested at 0.05 level.

different academic setting, or in response to instructional variables such as writing.

Although critical thinking gains were influenced by writing, they did not appear to be affected by gender, ethnicity, class standing, or age. In fact, statistical results indicated that these variables collectively had a very small effect on critical thinking performance. Gender distribution was nearly identical across the writing and nonwriting groups, and was predominantly female (nearly 62%). Ethnic distribution was also highly similar across the writing and nonwriting groups, but the sampling was largely Caucasian (>84%). Class standing varied a little more across the writing and nonwriting groups, with the sample largely comprised of underclassmen (70%). Although nearly three-quarters of the sample was between 18 and 21 years of age, nearly 10% was

Table 5. Effect of writing on component critical thinking performance

Component skill	Mean raw score and change			
	Writing (n = 158)		Nonwriting (n = 152)	
	Raw score	SEM	Raw score	SEM
Analysis (pre)	4.22	0.11	4.35	0.11
Analysis (post)	4.54	0.10	3.99	0.13
Analysis (change)	0.33 ^a	0.11	-0.36	0.14
Inference (pre)	7.42	0.18	7.07	0.20
Inference (post)	7.91	0.20	6.83	0.23
Inference (change)	0.48 ^a	0.16	-0.24	0.21
Evaluation (pre)	4.20	0.14	4.04	0.15
Evaluation (post)	4.57	0.15	4.13	0.17
Evaluation (change)	0.37	0.15	0.09	0.17

Comparison of writing and nonwriting group performance based on critical thinking component skill raw scores (CCTST subscales). Score range was 0–7 (analysis), 0–16 (inference), and 0–11 (evaluation).

^a Significance tested at 0.05 level.

over 21, with a fair number of older nontraditional students represented. It is possible that a more diverse sample would have produced different results, or it may be that the individuals participating in this study responded particularly well to writing. Although further investigation of these variables is necessary and important, it was beyond the scope of the current study.

The analysis of component skills provided greater insight into the particular critical thinking skills that students changed in response to writing. Specifically, writing students significantly improved their analysis and inference skills whereas nonwriting students did not. Writing students also improved their evaluation skills much more than nonwriting students, although not significantly. These results indicate that the process of writing helps students develop improved analytical and inference skills. Prior research indicates that the writing to learn strategy is effective because students must conceptually organize and structure their thoughts as well as their awareness of thinking processes (Langer and Applebee, 1987; Ackerman, 1993; Holliday, 1994; Rivard, 1994). More specifically, as students begin to shape their thoughts at the point of construction and continually analyze, review, and clarify meaning through the processes of drafting and revision, they necessarily engage and apply analysis and inference skills (Klein, 1999; Hand and Prain, 2002). In this study, the process of writing appears to have influenced critical thinking gains. It also seems likely that writing students experienced a greater cognitive demand than nonwriting students simply because the writing act required them to hypothesize, debate, and persuade (Rivard, 1994; Hand and Prain, 2002) rather than memorize as was the case in nonwriting control courses.

Conversely, the lack of any significant change in analysis, inference, or evaluation skills in the nonwriting group indicated that the traditional lab instruction used in the general education biology control courses did not help students develop critical thinking skills. Based on the results of this study, it could be argued that traditional lab instruction actually prevents the development of critical thinking skills, which presents a rather large problem when one considers how frequently these traditional methods are used in general education biology courses. One also has to consider that the critical thinking gains seen in the writing group might also have resulted from the relative absence of traditional lab instruction rather than writing alone. Additional research will be necessary to gain further insight into this question. Either way, changes to the traditional model of lab instruction will be necessary if the goal is to enhance the critical thinking abilities of general education biology students.

The variable that had the largest impact on critical thinking performance gains was prior critical thinking skill. This phenomenon was previously observed by Quitadamo, Brahler, and Crouch (unpublished results) in a related study that investigated the effect of Peer Led Team Learning on critical thinking performance. That study focused on science and math major undergraduate critical thinking performance at a major research university, and found that, in addition to Peer Led Team Learning, prior critical thinking skill significantly influenced critical thinking performance (Quitadamo, Brahler, and Crouch, unpublished results). Specifically, students with the highest prior crit-

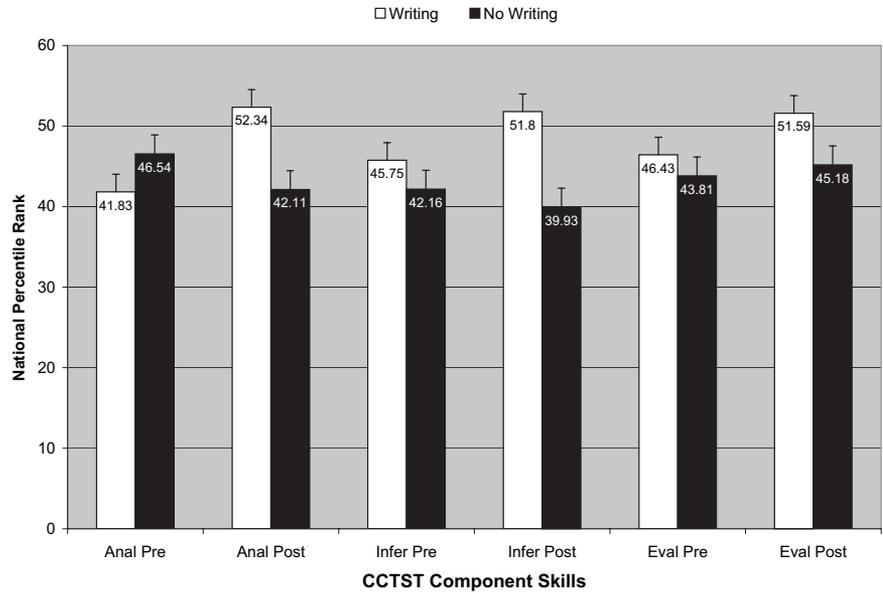


Figure 3. Effect of writing on component critical thinking national percentile rank. Comparison of component critical thinking national percentile gains between writing and nonwriting groups. Percentile ranking was computed using CCTST raw scores, an equivalency scale from Insight Assessment, and a linear conversion script in SPSS.

ical thinking skill showed the largest performance gains, whereas students with low initial skill were at a comparative disadvantage. The fact that prior critical thinking skill also had a large effect on critical thinking performance in this study increases the generalizability of the observation and underscores its importance. Simply put, students who have not been explicitly taught how to think critically may not reach the same potential as peers who have been taught these skills, not because they lack the cognitive hard-wiring to perform but because they lack the tools to build their knowledge. Is it reasonable or just to expect otherwise comparable students to perform at similar levels when only some of them have the keys for

success? If we hope to improve the perception of science in this country, we need to educate people on *how* to think about important scientific issues, and not simply argue a position based on one school of thought. By helping general education students to develop critical thinking skills, it is hoped that they will be better able to think rationally about science.

The observation that students who come to general education biology with greater critical thinking skills leave with the largest skill gains has important implications for the K–12 school system as well. If a high proportion of students are coming to institutions of higher education lacking critical thinking skills, why are these skills not being explicitly

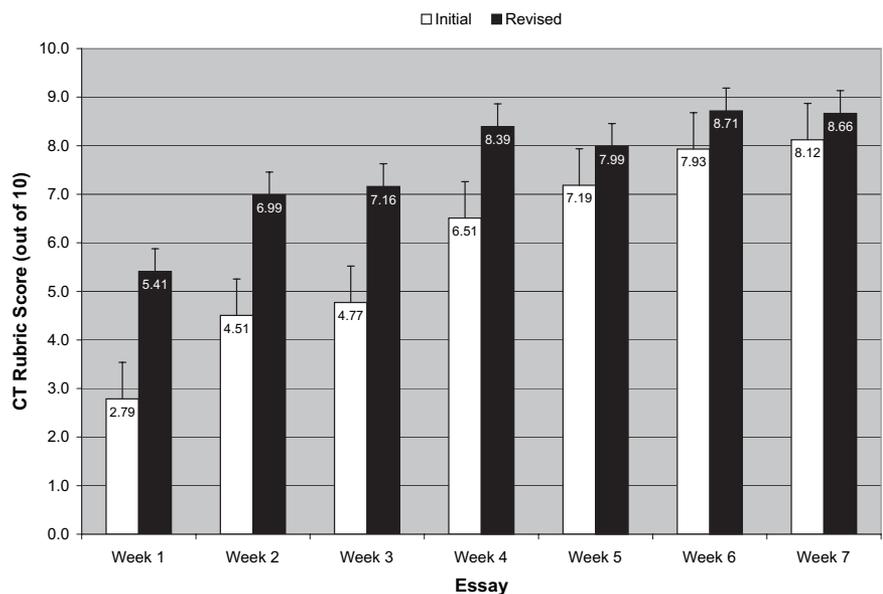


Figure 4. Profile of change in critical thinking performance in writing group. Comparison of student writing performance on weekly initial and revised essays. Essay scores were derived using a thesis-based critical thinking rubric (see Supplemental Appendix 2). Average essay scores were computed across writing sections.

taught in the K–12 system? Ideally, students would learn the foundational tenets of critical thinking at an earlier age, and be able to refine and hone these skills as they progress through the K–20 education system. The results of this study reinforce the idea that students should be explicitly taught critical thinking skills and be expected to practice them as early and often as possible.

Although its effect was smaller than writing or prior critical thinking skill, the instructor variable also played a significant role in student critical thinking performance, accounting for 2.5% of the total variance in critical thinking gains. Determining the particular qualities of each instructor that contributed to student critical thinking success and further separating instructor and writing effects will require additional research. Previous research indicates that teaching style positively influences certain aspects of student learning (Grasha, 1994; Hativa *et al.*, 2001; Bain, 2004), but the qualities that specifically influence student critical thinking gains have not been sufficiently investigated. Additional research in this area is necessary.

Faculty considering whether to use writing in the laboratory may wonder about how much time and energy it takes to implement, if efforts to change will translate into improved student learning, and how these changes affect disciplinary content. From a practical perspective, implementing writing did not take more time and effort per se; rather, it required faculty to reconceptualize *how* they spent their instructional time. Instead of individually developing course materials, writing faculty collaborated to a greater extent than nonwriting faculty on course design and assessments that required students to demonstrate their critical thinking skill. Interviews of faculty from the writing and nonwriting groups indicated that writing faculty felt the course was less work because they collaborated with colleagues and because students demonstrated improved thinking skill. Writing faculty generally became more comfortable with the new model after ~2–3 wk when students began to show observable changes in writing proficiency and critical thinking. Together, collaboration with colleagues and observed gains in critical thinking tended to create a positive feedback loop that helped to sustain writing faculty efforts. In contrast, nonwriting faculty similarly wanted their students to think better but were convinced that traditional methods would be more effective, and so remained closed to change. There were some logistical challenges with writing, like scheduling computer labs where students could draft and revise their weekly essay responses under instructor and teaching assistant supervision. Teaching assistants (and faculty) also needed to be trained on how to evaluate writing using a rubric. Finally, with regards to content coverage, no lecture or laboratory content was killed in order to implement writing because writing and nonwriting students both performed the same lab activities. Collectively, the benefits of using writing in laboratory should encourage faculty who want their students to learn to think critically to give it a try.

Future Directions

This study showed that writing affects student critical thinking skill in a nonmajors biology course, but the results have generated more questions than have been answered. How

does writing specifically produce gains in critical thinking performance? What factors influence student prior critical thinking skill? How do instructors specifically influence student gains in critical thinking? Future studies that analyze student essays in more detail would provide greater insight into how writing influences critical thinking skill. Using writing in other nonmajor science courses such as chemistry, geology, or physics could also be done to determine the transferability of this method. Additional studies that investigate student prior critical thinking skill and instructor variables are also necessary. These future studies would further contribute to the knowledge base in this area, and also address some of its identified limitations (Ebert-May *et al.*, 1997; Daempfle, 2002). Results from these studies would also increase the generalizability of the results from this study.

CONCLUSIONS

Building on existing research and on the basis of several lines of evidence presented in this study, we conclude that writing positively influences critical thinking performance for general education biology students. Those students with prior critical thinking skill may have a comparative advantage over other general education biology students who have not developed these same skills. To rectify that inequity critical thinking skills should be explicitly taught early and used often during the K–20 academic process. As it appears that particular instructors improve student critical thinking skills more than others, students should be discerning in their choice of instructors if they want to improve their critical thinking skills. Whether writing as a method to improve critical thinking skills will prove useful in other general education science courses will likely depend on a host of factors, but it has potential. Further study of writing in general education science will be necessary to verify these results and discover the breadth and depth of how writing affects critical thinking skill.

ACKNOWLEDGMENTS

We thank Drs. Holly Pinkart, Roberta Soltz, Phil Mattocks, and James Johnson and undergraduate researchers Matthew Brewer, Dayrk Flaugh, Adam Wallace, Colette Watson, Kelly Vincent, and Christine Weller for their valuable contributions to this study. The authors also acknowledge the generous financial support provided by the Central Washington University Office of the Provost and the Office of the Associate Vice President for Undergraduate Studies.

REFERENCES

- Ackerman, J. M. (1993). The promise of writing to learn. *Writ. Commun.* 10(3), 334–370.
- American Association for the Advancement of Science (1989). *Science for All Americans. A Project 2061 Report on Literacy Goals in Science, Mathematics, and Technology*, Washington, DC.
- Applebee, A. N. (1984). Writing and reasoning. *Rev. Educ. Res.* 54(4), 577–596.
- Association of American Colleges and Universities (2005). *Liberal Education Outcomes: A Preliminary Report on Student Achievement in College*, Washington, DC.

- Bain, K. (2004). *What the Best College Teachers Do*, Cambridge, MA: Harvard University Press.
- Beers, T., McIssac, C., Henderson, B., and Gainen, J. (1994). *Writing: thesis and support scoring guide*. http://www.insightassessment.com/pdf_files/RUB_WTHS.PDF (accessed 25 August 2006).
- Bruffee, K. A. (1984). Collaborative learning and the "conversation of mankind." *Coll. Engl.* 46(7), 635–653.
- Business-Higher Education Forum, and American Council on Education (2003). *Building a Nation of Learners: The Need for Changes in Teaching and Learning To Meet Global Challenges*, Washington, DC.
- Bybee, R. W., and Fuchs, B. (2006). Preparing the 21st century workforce: a new reform in science and technology education. *J. Res. Sci. Teach.* 43(4), 349–352.
- Campbell, D. T., and Stanley, J. C. (1963). *Experimental and Quasi-experimental Designs for Research*, Boston, MA: Houghton Mifflin Company.
- Carnevale A. P., and American Society for Training and Development (1990). *Workplace Basics: The Essential Skills Employers Want*, San Francisco, CA: Jossey-Bass.
- Center for Critical Thinking and Moral Critique (1996). *ICAT Critical Thinking Essay Test*, Rohnert Park, CA: Sonoma State University.
- Champagne A., and Kouba, V. 1999. Written product as performance measures. In: *Assessing Science Understanding: A Human Constructivist View*, ed. J. Mintzes, J. Wandersee, and J. Novak, New York: Academic Press, 224–248.
- Collier, K. G. (1980). Peer-group learning in higher education: the development of higher order skills. *Stud. High. Educ.* 5(1), 55–61.
- Daempfle, P. A. (2002). *Instructional Approaches for the Improvement of Reasoning in Introductory College Biology Courses: A Review of the Research*, New York: U.S. Department of Education.
- Ebert-May, D., Brewer, C., and Allred, S. (1997). Innovation in large lectures—teaching for active learning. *Bioscience* 47(9), 601–607.
- Ennis, R. H. (1985). A logical basis for measuring critical thinking skills. *Educ. Leadership* 43(2), 44–48.
- Ennis, R. H., and Weir, E. (1985). *The Ennis-Weir Critical Thinking Essay Test*, Pacific Grove, CA: Midwest Publications.
- Facione, P. A. (1990a). *The California Critical Thinking Skills Test—College Level. Technical Report 1. Experimental Validation and Content Validity*, Millbrae, CA: Insight Assessment.
- Facione, P. A. (1990b). *The California Critical Thinking Skills Test—College Level. Technical Report 3. Gender, Ethnicity, Major, CT Self-Esteem, and the CCTST*, Millbrae, CA: Insight Assessment.
- Facione, P. A. (1990c). *The California Critical Thinking Skills Test—College Level. Technical Report 4. Interpreting the CCTST, Group Norms, and Sub-Scores*, Millbrae, CA: Insight Assessment.
- Facione, P. A. (1991). *Using the California Critical Thinking Skills Test in Research, Evaluation, and Assessment*, Millbrae, CA: Insight Assessment.
- Facione P. A., and American Philosophical Association (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Research Findings and Recommendations*, Millbrae, CA: Insight Assessment.
- Facione, P. A., Facione, N. C., and Giancarlo, C. A. (1992). *Test Manual: The California Critical Thinking Disposition Inventory*, Millbrae, CA: Insight Assessment.
- Facione P. A., Facione N. C., and Insight Assessment. (2004). *Test of everyday reasoning*. <http://www.insightassessment.com/tester.html> (accessed 30 June 2006).
- Giancarlo, C. A., and Facione, P. A. (2001). A look across four years at the disposition toward critical thinking among undergraduate students. *J. Gen. Educ.* 50(1), 29–55.
- Grasha, A. F. (1994). A matter of style: the teacher as expert, formal authority, personal model, facilitator, and delegator. *Coll. Teach.* 42(4), 142–149.
- Hand, B., and Prain, V. (2002). Teachers implementing writing-to-learn strategies in junior secondary science: a case study. *Sci. Educ.* 86(6), 737–755.
- Hativa, N., Barak, R., and Simhi, E. (2001). Exemplary university teachers: knowledge and beliefs regarding effective teaching dimensions and strategies. *J. High. Educ.* 72(6), 699–729.
- Holliday, W. G. (1994). The reading-science learning-writing connection: breakthroughs, barriers, and promises. *J. Res. Sci. Teach.* 31(9), 877–893.
- Holmes, J., and Clizbe, E. (1997). Facing the 21st century. *Bus. Educ. Forum* 52(1), 33–35.
- Holyoak, A. R. (1998). A plan for writing throughout (not just across) the biology curriculum. *Am. Biol. Teach.* 60(3), 186–190.
- Howard, V. A. (1990). Thinking on paper: a philosopher's look at writing. In: *Varieties of Thinking: Essays from Harvard's Philosophy of Education Research Center*, ed. V. A. Howard, New York: Routledge, 84–92.
- Insight Assessment, and Blohm, S. (2005). Annual number of users for the CCTST form 2000. (accessed 8 December 2006).
- Jones, E. A., Hoffman, S., Moore, L. M., Ratcliff, G., Tibbets, S., and Click, B., III. (1995). *National Assessment of College Student Learning: Identifying College Graduates' Essential Skills in Writing, Speech and Listening, and Critical Thinking*. Final project report. University Park, PA: U.S. Department of Education, Office of Educational Research and Improvement. Report no. NCE-95-001.
- Jones G. M., and Carter G. (1998). Small groups and shared constructions. In: *Teaching Science for Understanding: A Human Constructivist View*, ed. J. J. Mintzes, J. H. Wandersee, and J. D. Novak, San Diego, CA: Academic Press, 261–279.
- Kelly, G. J., and Chen, C. (1999). The sound of music: constructing science as sociocultural practices through oral and written discourse. *J. Res. Sci. Teach.* 36(8), 883–915.
- Keys, C. W. (1999). Revitalizing instruction in scientific genres: connecting knowledge production with writing to learn in science. *Sci. Educ.* 83(2), 115–130.
- Keys, C. W., Hand, B., Prain, V., and Collins, S. (1999). Using the science writing heuristic as a tool for learning from laboratory investigations in secondary science. *J. Res. Sci. Teach.* 36(10), 1065–1084.
- Klein, P. (1999). Reopening inquiry into cognitive processes in writing-to-learn. *Ed. Psychol. Rev.* 11(3), 203–270.
- Klein, P. D. (2004). Constructing scientific explanations through writing. *Instr. Sci.* 32(3), 191–231.
- Klymkowsky, M. W. (2006). Can nonmajors courses lead to biological literacy? Do majors courses do any better? *Cell. Biol. Educ.* 4, 42–44.
- Kokkala, I., and Gessell, D. A. (2003). Writing science effectively: biology and English students in an author-editor relationship. *J. Coll. Sci. Teach.* 32(4), 252–257.
- Kurfiss J. G., and Association for the Study of Higher Education (1988). *Critical Thinking: Theory, Research, Practice, and Possibilities*, Washington, DC: George Washington University.
- Langer, J. A., and Applebee, A. N. (1985). Learning to write: learning to think. *Educ. Horizons* 64(1), 36–38.

- Langer, J. A., and Applebee, A. N. (1987). *How Writing Shapes Thinking: A Study of Teaching and Learning*. NCTE research report no. 22. Urbana, IL: National Council of Teachers of English.
- Lawson, A. E. (2001). Using the learning cycle to teach biology concepts and reasoning patterns. *J. Biol. Educ.* 35(4), 165–169.
- Malcom, S. M., Abdallah, J., Chubin, D. E., and Grogan, K. (2005). *A System of Solutions: Every School, Every Student*, Washington, DC: American Association for the Advancement of Science.
- Marzano, R. J. (1991). Fostering thinking across the curriculum through knowledge restructuring. *J. Reading* 34(7), 518–525.
- National Academy of Sciences, National Academy of Engineering, Institute of Medicine (2005). *Rising Above the Gathering Storm: Energizing and Employing America for a Brighter Economic Future*, Washington, DC: Committee on Prospering in the Global Economy of the 21st Century.
- National Research Council (1995). *National Science Education Standards*, Washington, DC: National Academy Press.
- National Research Council (2003). *Bio 2010, Transforming Undergraduate Education for Future Research Biologists*, Washington, DC: Committee on Undergraduate Biology Education to Prepare Research Scientists for the 21st Century.
- National Science Foundation (1996). *Shaping the Future: New Expectations for Undergraduate Education in Science, Mathematics, Engineering, and Technology*, Washington, DC: Directorate for Education and Human Resources.
- Office of Educational Research and Improvement (1991). *Striving for excellence: The National Education Goals*, Washington, DC.
- Project Kaleidoscope. (2006). *Transforming America's Scientific and Technological Infrastructure: Recommendations for Urgent Action*, Washington, DC: National Science Foundation.
- Resnick, L. B. (1987). *Education and Learning To Think*, Washington DC: National Academy Press.
- Rivard, L. P. (1994). A review of writing to learn in science: implications for practice and research. *J. Res. Sci. Teach.* 31(9), 969–983.
- Springer, L., Donovan, S. S., and Stanne, M. E. (1999). Effects of small-group learning on undergraduates in science, mathematics, engineering, and technology: a meta-analysis. *Rev. Educ. Res.* 69(1), 21–51.
- Steglich, C. S. (2000). A writing assignment that changes attitudes in biology classes. *Am. Biol. Teach.* 62(2), 98–101.
- Taylor, K. L., and Sobota, S. J. (1998). Writing in biology: an integration of disciplines. *Am. Biol. Teach.* 60(5), 350–353.
- Tessier, J. (2006). Writing assignment in a nonmajor introductory ecology class. *J. Coll. Sci. Teach.* 35(4), 25–29.
- Tobin, K. G., Tippins, D. J., and Gallard, A. J. (1994). Research on instructional strategies for teaching science. In: *Handbook of Research on Science Teaching and Learning*, ed. D. L. Gabel, New York: Macmillan, 45–93.
- Tsui, L. (1998). A review of research on critical thinking. ASHE annual meeting paper; 1998 November 5–8; Miami, FL.
- Tsui, L. (2002). Fostering critical thinking through effective pedagogy: evidence from four institutional case studies. *J. High. Educ.* 73(6), 740–763.
- U.S. Department of Education (1990). *National Goals for Education*, Washington, DC.
- Watson, G., and Glaser, E. M. (1980). *Watson-Glaser Critical Thinking Appraisal*, Cleveland, OH: The Psychological Corporation (Harcourt Brace Jovanovich).