Article

The Utility of Writing Assignments in Undergraduate Bioscience

Julie Libarkin[*][†] **and Gabriel Ording**[†][‡]

*Geocognition Research Lab and Center for Integrative Studies in General Science and ‡Department of Entomology and Center for Integrative Studies in General Science, Michigan State University, East Lansing, MI 48824

We tested the hypothesis that engagement in a few, brief writing assignments in a nonmajors science course can improve student ability to convey critical thought about science. A sample of three papers written by students ($n = 30$) was coded for presence and accuracy of elements related to scientific writing. Scores for different aspects of scientific writing were significantly correlated, suggesting that students recognized relationships between components of scientific thought. We found that students' ability to write about science topics and state conclusions based on data improved over the course of three writing assignments, while the abilities to state a hypothesis and draw clear connections between human activities and environmental impacts did not improve. Three writing assignments generated significant change in student ability to write scientifically, although our results suggest that three is an insufficient number to generate complete development of scientific writing skills.

INTRODUCTION

Fundamental skills, including reading, writing, and mathematics, are generally targeted in both core and ancillary college-level courses. Many institutions offer "across-the-curriculum" approaches to instruction, with basic skills becoming integrated into courses that traditionally focus on specific content, such as science. Writing across the curriculum, for example, is a well-respected national movement to make writing an integral part of any course (Townsend, 2002). Many scientists recognize the value of written communication for science, and have embraced the infusion of writing about science in courses for majors and nonmajors alike. Certainly, the ability to read and understand scientific writing is an important skill for both scientists and the general public.

National organizations, such as the American Association for the Advancement of Science (AAAS), have long acknowledged the importance of writing in scientific literacy by incorporating explicit discussion of writing into their literacy frameworks and standards (AAAS, 1989; Aaron, 1996). In addition to targeting reading and writing skills within K–12 science classrooms, educators in colleges and universities also acknowledge the importance of these skills within the sciences. The ability to connect ideas and reason about the importance of scientific claims aligns with what Norris and Phillips (2003) would consider to be scientific literacy. For science majors, writing skills are often enhanced through term papers or lab reports requiring critical thinking, information processing, and writing skills (Allie *et al.*, 1997; Krest and Carle, 1999). Incorporation of significant writing assignments is less common in nonmajors courses, particularly for large sections in which grading of such assignments is prohibitive. Writing in nonmajors courses, if it exists, is generally limited in both frequency and magnitude of the assignment.

Discussion of the importance of writing in science courses abounds and suggests the effectiveness of writing for the building of writing or other general skills (Koprowski, 1997; Krest and Carle, 1999; Peat *et al.*, 2005), scientific literacy (Hand *et al.*, 1999; Quitadamo and Kurtz, 2007), reasoning (Edwards *et al.*, 2001; VanDerZanden, 2010), and conceptual understanding (Greenbowe and Schroeder, 2008). These studies indicate that, in general, writing can be an effective use of time in college science courses.

J. Libarkin and G. Ording

Despite the utility of writing for building a suite of valuable abilities, undergraduate students generally have few opportunities to engage in writing, whether within or outside of the science major (Motavalli *et al.*, 2007). This remains true despite national effort to incorporate writing into the broader undergraduate curriculum (Townsend, 2002), including science (Huerta and McMillan, 2000). Faculty are generally reluctant to incorporate writing assignments into undergraduate courses. This reluctance may stem from inherent limitations on faculty time for reviewing and commenting on writing, the ambiguity associated with grading written work, or a general lack of understanding about how to effectively modify instruction to target writing skills.

While significant literature stresses the importance of writing for building of general scientific literacy and suggests the importance of writing skills within the context of science courses, very little work has provided empirical evidence for the effectiveness of writing activities in undergraduate science classrooms (e.g., Quitadamo and Kurtz, 2007). In particular, the value to nonmajors of writing about science, especially when time permits only a few such exercises per course, is clear from studies of pre/post changes in critical thinking (VanDerZanden, 2010) and conceptual understanding (Greenbowe and Schroeder, 2008). The utility of science writing exercises for enhancing student ability to think scientifically, as reflected in their ability to write about science, is not as clearly understood.

We consider two research questions to add to the discourse about writing in nonmajors science instruction:

1. Can students improve their science writing skills through integration of a limited number (three) of short writing exercises into a nonmajors classroom?
2. Which aspects of science writing can be improved through repeated, yet limited, opportunities?

We hypothesized that general science writing skills, including all aspects of science writing, would improve over the course of a semester. To address this hypothesis, we report on the results of an investigation of the impact of writing on a random sample of students enrolled in a large, general education science course at a midwestern university in North America.

METHODS

This study is observational in nature and took advantage of normal classroom practice for data collection; this was not an experimental design. We applied mixed-methods approaches in our work and in establishing study validity and reliability. In particular, we analyzed three short papers written over a semester by a subsample of 30 students enrolled in a general education biology course. This analysis considered the extent to which student ability to write effectively changed over the course of these writing assignments. We applied a Likert-scale scoring rubric to analyze five key characteristics of scientific writing exhibited within student papers, and utilized standard classical statistics to document interrater reliability, statistical correlations between scores for individual writing characteristics, and change in individual writing characteristics and total writing scores over time. Appropri-

ate institutional review board approval was received for this study.

Participants

The study population ($n = 30$) consisted of a random sample within a stratified population of students enrolled in an entry-level nonmajors biology course taught by the second author during Fall 2006 at a midwestern North American university. Ten students each from high-, medium-, and low-performing groups, differentiated by final course GPA, were randomly selected to minimize student ability effects. Random sampling occurred through generation of a random set of numbers and extraction of papers corresponding to those numbers from each of the three groups. Papers from these 30 students were de-identified, and each of three raters was given a randomly ordered set of papers to score. In general, students enrolled in this course are almost evenly divided between males and females, often with between 43% and 57% gender distribution. Students tend to have mixed levels of college experience, with the majority in their second year of college ($\sim$50%). Students in the first, third, and fourth (or later) years generally make up about 15%, 25%, and 10% of the enrollment, respectively. At the start of the term, less than 10% of the students are undecided in their major, with most students enrolled in diverse nonscience majors, including marketing, criminal justice, anthropology, psychology, and education. Very few students are declared science, math, technology, or engineering majors, although a small number (<10%) of such students do occasionally enroll in the course.

Course-Embedded Data Collection

Students completed three papers over the course of the semester; each paper was similar in design and requirements. At the same time, students were learning how to write scientifically as they engaged in these assignments. Although all three papers required students to recognize and to illustrate ecological and social connections, only the second and third papers explicitly required students to make decisions and recommendations. Each paper assignment was preceded by a brief in-class discussion that modeled the assignment.

Paper 1 was assigned in week 2 and collected in week 4 of a 15-wk semester. Assignment of this paper coincided with a unit on insect ecology. On the day of the assignment of paper 1, a presentation consisting of 10 slides was shown to the students. This presentation modeled identifying ecological relationships in nature and unintended human impacts on those connections and the ecological system in general. Specifically, the instructor modeled two different ecosystem relationships as these ecosystems are impacted by human activity. For example, the relationship between gypsy moth outbreaks, caused by human introduction of gypsy moths into the eastern United States, and human Lyme disease was conveyed to students (e.g., Jones *et al.*, 1998). The instructor led students through this series of relationships: 1) Increased gypsy moth population causes mast failure (i.e., decreased acorn production by oak trees). 2) A reduction in acorns results in decreased mouse and deer population density, since mice and deer feed on acorns. 3) A decrease in mouse and deer populations causes a decreased infection rate of black-legged ticks on deer and mice. 4) A decreased infection rate causes

a decrease in the transmission of Lyme disease to the ticks. 5) A decrease in Lyme disease in ticks results in a decrease in Lyme disease in humans. This is an engaging example for students, since the invasive gypsy moth, although harmful to trees, has a beneficial impact on humans.

Paper 2 was assigned in week 5 and completed in week 7 and followed instruction related to population dynamics and ecology. Building from the earlier experience with paper 1, students were introduced to the paper's topic during class. Students were reminded of the structure and content of the first paper, and then presented with a new paper topic. In this case, no additional modeling was provided. Students were required both to make ecosystems connections and to develop a position for human engagement in ecosystem-impactful activities.

Finally, paper 3 was assigned in week 10, with a due date in week 12; during this time, students were learning about water quality and disease transmission. As with the second paper, students were reminded of the content and structure of the first two papers and were presented with a final unique scenario related to human–ecosystem interactions. Students were not provided with any additional modeling, although they were explicitly instructed to consider comments on the first two papers in developing their argument for paper 3.

Students were not limited to a particular paper length, although four pages, double-spaced, was a recommended minimum. In this course, paper 1 tends to be four to five pages, with an increase to an average length of six to seven pages by paper 3. The three papers build from an initial introduction to the relationship between humans and pests, through consideration of how humans can reduce impacts of pests, to culminate in a debate as to whether we should utilize certain techniques to combat pests in the face of potential environmental impacts.

Paper 1: Deforestation and Malaria. Students were asked to investigate potential implications and connections between human activities associated with globalization and disease. Specifically, students considered the ramifications of widespread deforestation on the global incidence of malaria. This provided opportunities for students to clearly recognize and articulate their understanding of the relationships between human activities and insect-vectored diseases.

Paper 2: Gypsy Moth and Bacillus. Students were asked to determine whether the use of a natural pesticide was an appropriate mechanism for combating an invasive species. In particular, the use of *Bacillus thuringiensis* to reduce ecosystem impacts associated with invasive gypsy moths was considered. This provided opportunities for students to consider the benefits and drawbacks to specific pest-management techniques.

Paper 3: Malaria and DDT. Students were asked to investigate the history of the use of DDT as an insect control mechanism. In this paper, students were again asked to weigh the pros and cons of the use of a specific pesticide to combat malaria on a global scale. This provided students with a capstone writing experience in which to draw together the lessons learned from paper 1 and paper 2 in the context of a novel scenario.

Students were provided with the scoring rubric (Figure 1) for each paper upon assignment. Each of the elements of the global rubric (Figure 1) was embedded in the rubrics targeted to each individual paper assignment. Students were therefore aware that papers would be assessed based on five key characteristics that correlate to a model of scientific thinking. In particular, students were required to 1) provide a scientific

---

PAPER RUBRIC
Each of the following categories has been designed so as to provide clear mechanisms to objectively assign a numeric rank of 0 – 3 points. 3 = Excellent; 2 = Average; 1 = Poor; 0 = Absent.

Scientific Background (SCI)
The student has provided an organized summary of information, from multiple scientifically reputable sources, describing the assigned human activities and relevant biological topic(s).

Hypothesis / Position (HYP)
A concise and testable statement indicating whether the described / proposed human activity can be linked with ecological / environmental impacts (positive or negative).

Connections (CONN)
Clear links have been identified between the assigned human activities and any resulting ecological / environmental impacts associated with the assigned topic. A clear cause and effect relationship has been demonstrated.

Investigations and Data (DATA)
Multiple scientifically reputable sources have been utilized to gather specific examples of investigations and data that directly support or refute the hypothesis.

Conclusion (CONC)
A clear summary of the collected information and data that supports and / or refutes the stated hypothesis has been provided (cost / benefit analysis). This summary leads to a clear conclusion (decision / recommendation associated with future activities).

Figure 1. Coding rubric used to evaluate student papers. Students were provided with paper-specific versions of this rubric as part of each assignment.

background with reference to reputable sources of information (SCI), 2) state a testable hypothesis or position statement (HYP), 3) identify clear connections between human activities and ecological or environmental impacts (CONN), 4) gather data that would support or refute the hypothesis (DATA), and 5) draw a conclusion about the original hypothesis based on collected data (CONC). Each characteristic was rated on a Likert scale of 0–3, where 0 represented a missing component and 3 represented excellence. An overall paper score was calculated as the sum of the characteristic scores, for a maximum paper score of 15.

Data Analysis

All 90 papers, three per each of 30 students, were scored based on each of the five characteristics by three independent raters hired to provide unbiased analysis. Raters were three graduate teaching assistants with the necessary exposure to entomology and ecology to approach the papers as experts. The raters and the second author discussed the global rubric and its interpretation prior to concurrent scoring of six papers. Throughout the rating process, none of three raters was told the order in which students had been assigned papers. This preserved the ability to compare changes in student writing over time without introducing unintentional bias to the study. Initial disagreement among the raters after pilot coding of three papers was addressed through modification of the global rubric. Analysis of ratings for three additional papers brought agreement to at least 80% across the three pilot raters.

Scores were on a 4-point scale (0–3), where 3 represents excellent, 2 represents adequate, 1 represents poor, and 0 represents an absent characteristic (Figure 1). A total score for each paper was determined by summing the five characteristic scores, giving a maximum possible score of 15. A suite of statistical tests was run in PASW Statistics 18.0; these tests are discussed in detail in the *Results* section. A reliability analysis was run to determine the interrater reliability of the three raters who scored the papers on five measures; these five variables were summed to give a total score (a sixth measure). We calculated the single-measures, intraclass correlation across six measures on 30 papers coded by each of three raters; the single-measures, intraclass correlation was 0.89 (minimum = 0.88; maximum = 0.91). Since a perfect positive correlation is reflected by an intraclass correlation of 1.0, a correlation of 0.89 indicates that all three raters were in good agreement across codes, allowing us to use average ratings for additional analysis.

Validity and Reliability

In this work, we have carefully considered aspects of rigor related to study validity, the extent to which we are measuring the true value of outcomes, and reliability, the reproducibility or repeatability of our measures and study (e.g., Trochim and Donnelly, 2007). Coupled with consideration of trustworthiness for qualitative aspects of our study was our concern to address essential aspects of overarching validity and reliability as appropriate for mixed-methods research (Table 1; Lincoln and Guba, 1985; Litwin, 1995; Trochim and Donnelly, 2007). Table 1 describes those approaches to validity and reliability explicitly considered in this study.

Limitations

As with any study, this work has limitations inherent in the study design. This study was observational and did not utilize a control group of students. Students were not randomly assigned to this course, although the subsample of students included in the study was chosen at random. The use of a subsample arose from the limitations of the amount of data three different raters could reasonably evaluate; investigation of a different subsample has a small probability of resulting in different findings. The subsample was too small to determine the effect of demographic variables, such as gender or ethnicity, on study outcomes. Finally, the applicability of this research to other instructional settings is limited to general findings, rather than specific relationships between variables.

The papers analyzed in this work were part of the formative assessment practice of the course. As such, they were not perfectly designed to illuminate changes in the variables of interest. In particular, paper 1 did not require students to engage in the same depth of decision making as did papers 2 and 3.

RESULTS

A simple review of the three sets of papers illustrates that students' ability to discuss scientific topics clearly improved over the course of the three writing assignments. For example, the sophistication with which students discussed the course's scientific material (SCI) increased from paper 1 to paper 2 to paper 3; this is illustrated in Table 2 through discussion of the characteristics of three papers written by a single student. This student is representative of the larger subsample. For example, we observed improvement in the presentation and use of data (DATA) by students over the three assignments. Other paper characteristics, such as statements of purpose or hypothesis (HYP), did not change as clearly, although papers appeared to improve overall. These raw data, while suggestive, do not provide concrete evidence for actual change in student ability. We utilized a series of quantitative statistical measures to determine 1) relationships between paper characteristics, 2) which characteristics saw improvement over the three assignments, and 3) the magnitude of this improvement.

Consideration of both subsample performance (Figure 2) and average scores (Table 3) over each of the five paper characteristics illustrates changes in student performance over time. These data suggest that large, positive changes occurred in student performance for some variables, while minor or negative changes occurred for other variables. Both DATA and CONC, as well as the overall total SUM score, showed systematic increases in student performance over the three papers, with both higher averages (Table 3) and more students performing at adequate or excellent levels (Figure 2). Student performance in the SCI category increased from paper 1 to paper 2, and remained steady for paper 3, while CONN performance only showed an increase from paper 2 to paper 3. The HYP category produced anomalous effects, with an initial decrease in performance from paper 1 to paper 2, and then an increase back to paper 1 performance levels from paper 2 to paper 3. This qualitative look at score distribution and averages is suggestive, although only statistical

Table 1. Validity and reliability considerations applied to this study[a]

Criteria	Description	This study
Content validity	Measure of whether or not items actually measure the trait they are intended to measure.	Components of a scientific paper evaluated here mirror those aspects of scientific writing considered necessary for scientific thinking by faculty with 10+ yr of science teaching experience.
Conclusion validity	Measure of ability to determine whether or not a relationship exists between the variables being studied.	We utilized only appropriate statistical measures in our correlation and difference analysis and chose an appropriate p value ($p \leq 0.05$) for statistical significance.
Construct validity	Measure of whether or not strong support for the content of items exists.	The raters hired to score student papers agreed that the paper characteristics were appropriate.
Communication validity	While researchers often assume that participants will interpret questions as intended, explicitly considering participant interpretation can generate important insights (e.g., Lopez, 1996).	The global scoring rubric was generated by one of the authors based on 10 yr of experience teaching nonmajor biology courses. This rubric is the culmination of years of revisions based on student feedback.
Transferability	A measure of the extent to which results can be generalized to populations outside the study.	Research findings are most directly applied to students enrolled in large, general education biology courses at the same university as the subsample. Findings can only be broadly applied to nonscience majors enrolled in biology courses nationwide.
Internal consistency reliability	The stability of test results across samples of similar populations, consistency in test results over time, and generation of similar results using slightly different forms all provide evidence that a survey is generating reproducible findings.	We considered each of the coded characteristics to represent an item on a scientific writing scale. For these five items, we calculated a Cronbach's alpha score of 0.78 using the average score for all three raters. This indicates that as a group, these characteristics are generally measuring a common variable.
Interrater reliability	Measure of the agreement between different raters analyzing the same data set.	Average intraclass correlation was calculated at 0.89, suggesting good agreement across codes.

[a]Except where noted, concepts of validity, reliability, and trustworthiness are adapted from Lincoln and Guba (1985), Litwin (1995), and Trochim and Donnelly (2007). Table is modeled after Clark and Libarkin (2011).

comparisons can provide evidence of significant correlations between categories or change over time.

We chose Spearman's rho (two-tailed) as an appropriate metric for statistical evaluation of the relationship between paper measures. The Spearman's metric was chosen over Pearson's or other approaches, because scores were based on a semiquantitative metric (ranking from 0 to 3), rather than quantitative scalar values. Statistically relevant correlations were observed between most measures, indicating that performance on one measure is a good predictor of performance on another. In general, SCI, CONN, DATA, and CONC measures were correlated with one another. Correlation coefficients between SCI, CONN, DATA, and CONC were all high, ranging from 0.61 to 0.83 (all at $p \leq 0.001$). Only the HYP measure did not correlate with all other measures, showing statistically significant correlations with only SCI, CONC, and SUM. In cases where HYP did correlate, the correlation coefficient was low (0.22–0.57) relative to the high correlations observed elsewhere (Table 4).

Finally, a Kruskal–Wallis test, a nonparametric equivalent for analysis of variance, was performed to compare the average characteristic and sum scores achieved by the subsample of students between the first, second, and third papers. Both HYP and CONN characteristics showed statistically insignificant change at $p \leq 0.10$. SCI, DATA, CONC and SUM all showed statistically significant differences between the three papers at $p \leq 0.05$; SCI was statistically significant at $p \leq 0.001$.

DISCUSSION

The importance of writing in college courses has been made evident by a wide array of research studies performed across disciplines, including the sciences. For science majors, the value of writing is clear; scientists need to be able to communicate effectively about science to one another and the general public. Courses for science and engineering majors, particularly at higher levels, often incorporate writing as an important measure of performance (e.g., Yalvac *et al.*, 2007). While writing in science courses for nonscience majors is considered important (e.g., Firooznia and Andreadis, 2006), the ability to incorporate effective writing exercises into nonmajors courses is essentially limited by the size of such courses and the overwhelming grading load associated with evaluating writing. Despite these limitations, writing about science provides students with essential skills. To fully comprehend the role that science plays in their lives, students need to be able to clearly articulate relationships between observations and the real world, as well as connections between science and the real world.

Table 2. Example responses from a single student over the three writing assignments

	SCI	DATA	HYP
Paper 1	One paragraph introduction	Four references cited, all online sources	"Deforestation causes large breeding grounds for mosquitoes as well as provides cleared areas where the spread of malaria is more common."
	Two paragraphs of malaria background Three paragraphs about deforestation and malaria One paragraph conclusions		
Paper 2	One paragraph introduction	Four references cited, all online sources	"The negative impacts caused by the gypsy moth are too great to discontinue use of Bacillus thuringiensis."
	Three paragraphs of gypsy moth background. Three paragraphs about defoliation and impact Three paragraphs about *Bacillus thuringiensis* One paragraph conclusion		
Paper 3	One paragraph introduction	Seven references cited, including one newspaper article and one journal article	"DDT is a particularly controversial pesticide because it is questioned for its safety by many, while at the same time advocates of the pesticide feel if it is used properly it can be extremely helpful to control malaria."
	Two paragraphs of malaria background One paragraph about DDT and malaria Three paragraphs pros/cons of DDT use One paragraph DDT background Three paragraphs argument with support One paragraph conclusion		
Comments	Use of sources to describe the biological topic and human activities *increased* over the three papers, along with better argumentation.	Use of data from appropriate sources to support position, rather than as background information only, *increased* over the three papers.	Explanation of the purpose or hypothesis under discussion *did not change* over the three papers.

In general, we found that student ability to write about scientific topics increased over the course of a limited number of assignments ($n = 3$). In particular, student ability to discuss scientific concepts (SCI), use data to support a position (DATA), and draw reasonable conclusions about future activity (CONC) all improved. Conversely, ability to state a purpose or a hypothesis (HYP) or draw clear connections between human activities and environmental impacts (CONN) did not improve over time. Hypothesis stating also did not correlate strongly with other variables at any time. We find these results encouraging, especially given both the brief time

spent on writing in this specific course and the significant value placed on writing across the undergraduate curriculum. At the same time, we propose underlying reasons for lack of improvement in or correlation between some paper elements.

The five characteristics of scientific writing chosen for analysis in this study reflect important elements of scientific discourse. Four of these five characteristics were highly correlated, suggesting that students recognized most connections between elements of scientific writing. Students' ability to state a hypothesis or purpose (HYP) correlated with

Table 3. Average scores ($n = 30$) across coders ($n = 3$) for five paper characteristics and total score[a]

	SCI	HYP	CONN	DATA	CONC	SUM
Paper 1	1.73 ± 0.39	1.99 ± 0.74	2.06 ± 0.53	1.73 ± 0.68	1.83 ± 0.55	9.34 ± 2.28
Paper 2	2.02 ± 0.28	1.68 ± 0.73	2.02 ± 0.44	2.09 ± 0.48	2.08 ± 0.42	9.89 ± 1.67
Paper 3	2.04 ± 0.36	1.97 ± 0.52	2.22 ± 0.48	2.22 ± 0.60	2.19 ± 0.46	10.65 ± 1.81

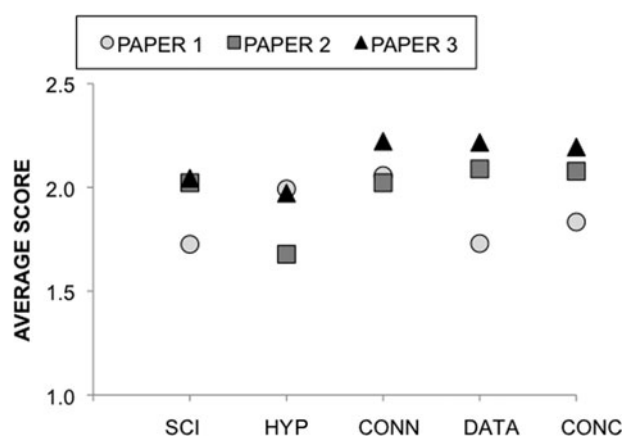[a]Standard deviation is 1 sigma.

Figure 2. Average score for the five paper characteristics (SCI, HYP, CONN, DATA, CONC) for papers 1, 2, and 3. Students had the most difficulty stating the hypothesis or purpose of their paper (HYP); this difficulty remained despite practice writing three papers. Student ability to provide a scientific background (SCI) remained fairly stable across the semester, while student ability to report and discuss data (DATA) improved significantly.

discussion of scientific background (SCI) and drawing of a conclusion about the original hypothesis (CONC) only. The correlation between HYP and SCI seems reasonable, in that reasonable hypotheses follow from a grounded understanding of the phenomenon under study. The correlation of HYP with CONC is also expected; the ability to pose a hypothesis is a prerequisite skill for drawing conclusions that are reasonable in the face of that hypothesis. Interestingly, these correlations suggest that accurate understanding of a hypothesis is not required for the collection of data to evaluate that hypothesis.

The ability of students to write about science topics and state conclusions based on data (SCI, CONC, DATA) each improved over the course of three writing assignments. The increase in ability to write cogently about science concepts (SCI) is not surprising. These students were enrolled in a bioscience course for nonmajors designed specifically to increase fundamental understanding of bioscience concepts of importance to nonscientists. Both correlation analysis and analysis of changes in characteristics over time indicate that

these three writing assignments were limited in their ability to improve students' ability to state hypotheses (HYP) or understanding of connections between human activities and environmental impacts (CONN). These findings align with the lack of strong correlation found between any variables and HYP.

We find the differential increases in other paper variables of much more interest and believe these point to areas to which science faculty in lower-division courses should pay careful attention. Fallahi *et al.* (2006) found that the science writing skills of psychology students did not improve until the fourth writing assignment given in a course. Although we cannot extrapolate this finding directly to nonmajors biology students, future implementations of writing assignments may need to move beyond the three-paper model. This recommendation must be considered in the context of the limitations faculty teaching large general education courses face, particularly in terms of time available to effectively grade assignments. Increasing the number of writing assignments in a course can be achieved through implementation of activities like student peer review, with the added benefit that peer review itself can help improve scientific writing skills (Gunersel *et al.*, 2008). Future studies should further consider the importance of how well students build connections between components of the scientific writing assignment, as touched on with the CONN variable discussed here.

The improved performance over the course of the three writing assignments may have been the result of factors unrelated to improved ability to write scientifically. The possibility that students were more motivated to perform well on the paper at the end of the semester could explain our overall observation of better performance on paper 3. This motivation could have arisen from the normal pressures to achieve a higher course grade. In addition, paper 3, as a novel scenario, may also have simply engaged student interest more deeply than papers 1 or 2, resulting in more thoughtful writing and higher scores.

This study provided insight into the core question of how much writing is enough writing for nonmajors science courses. We interpret our findings to indicate that engagement in a minimal number of brief writing assignments can improve students' ability to identify relevant background material, incorporate data into building an argument, and draw conclusions about the meaning of that data. At the

Table 4. Correlations and statistical significance between paper variables[a]

		HYP	CONN	DATA	CONC	SUM
SCI	Correlation coefficient	0.22*	0.64**	0.64**	0.66**	0.78**
	Sigma (two-tailed)	0.035	0.001	0.001	0.001	0.001
HYP	Correlation coefficient	—	0.20*	0.18	0.34**	0.57**
	Sigma (two-tailed)	—	0.05	0.08	0.001	0.001
CONN	Correlation coefficient	—	—	0.61**	0.63**	0.79**
	Sigma (two-tailed)	—	—	0.001	0.001	0.001
DATA	Correlation coefficient	—	—	—	0.61**	0.79**
	Sigma (two-tailed)	—	—	—	0.001	0.001
CONC	Correlation coefficient	—	—	—	—	0.83**
	Sigma (two-tailed)	—	—	—	—	0.001

*Correlation is significant at $p \leq 0.05$.

**Correlation is significant at $p \leq 0.001$.

[a] $n = 90$.

same time, despite engaging in three separate writing assignments, students were generally unable to clearly state hypotheses or draw connections between human activities and environmental impacts. These findings, in combination with other studies (Fallahi *et al.,* 2006), suggest that ability to write cogently about science will improve with just a few opportunities to write, although gaining full ability to communicate about science likely arises after multiple and repeated writing. Ultimately, we conclude that incorporating short writing assignments into large nonmajors college courses is both feasible and effective for improving student ability to communicate about science. The specific number of assignments needed to generate effective and complete improvement, however, is unclear, and is likely greater than three.

ACKNOWLEDGMENTS

REFERENCES

Aaron DK (1996). Writing across the curriculum: putting theory into practice in animal science courses. J Anim Sci *74*, 2810–2827.

Allie S, Buffler A, Kaunda L, Inglis M (1997). Writing-intensive physics laboratory reports: tasks and assessment. Phys Teach *35*, 399–405.

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: AAAS.

Clark SK, Libarkin JC (2011). Designing a mixed-format research instrument and scoring rubric: case study of a plate tectonics instrument. In: Qualitative Inquiry in Geoscience Education Research, Geological Society of America Special Paper 474, ed. A Feig and A Stokes, Boulder, CO: Geological Society of America, 81–96.

Edwards R, White M, Gray J, Fischbacher C (2001). Use of a journal club and letter-writing exercise to teach critical appraisal to medical undergraduates. Med Educ *35*, 691–694.

Fallahi CR, Wooda RM, Austada CS, Fallahi H (2006). A program for improving undergraduate psychology students' basic writing skills. Teach Psychol *33*, 171–175.

Firooznia F, Andreadis DK (2006). Information literacy in introductory biology. J Coll Sci Teach *35*, 23–27.

Greenbowe TJ, Schroeder JD (2008). Implementing POGIL and the science writing heuristic jointly in undergraduate organic chemistry—student perceptions and performance. Chem Ed Res Prac *9*, 149–156.

Gunersel NJ, Simpson KJ, Aufderheide L, Wang L (2008). Effectiveness of calibrated peer review™ for improving writing and critical thinking skills in biology undergraduate students. JoSoTL *8*, 25–37.

Hand BM, Prain V, Lawrence C, Yore LA (1999). Writing in science framework designed to enhance science literacy. Int J Sci Educ *10*, 1021–1036.

Huerta D, McMillan VE (2000). Instruction by writing and library faculty: a two-tiered approach to the teaching of scientific writing. Iss Sci Technol Librarian, http://webdoc.sub.gwdg.de/edoc/aw/ucsb/istl/00-fall/article1.html (accessed 10 June 2010).

Jones CG, Ostfeld RS, Richard MP, Schauber EM, Wolff JO (1998). Chain reactions linking acorns to gypsy moth outbreaks and Lyme disease risk. Science *13*, 1023–1026.

Koprowski JL (1997). Sharpening the craft of scientific writing: a peer-review strategy to improve student writing. J Coll Sci Teach *27*, 133–135.

Krest M, Carle DO (1999). Teaching scientific writing: a model for integrating research, writing and critical thinking. Am Biol Teach *61*, 223–227.

Lincoln YS, Guba EG (1985). Naturalistic Inquiry, Beverly Hills, CA: Sage.

Litwin MS (1995). How to Measure Survey Reliability and Validity, Thousand Oaks, CA: Sage.

Lopez W (1996). Communication validity and rating scales. Rasch Meas Trans *10*, 482–483.

Motavalli PP, Patton MD, Miles RJ (2007). Use of web-based student extension publications to improve undergraduate student writing skills. J Nat Res Life Sci Educ *36*, 95–102.

Norris SP, Phillips LM (2003). How literacy in its fundamental sense is central to scientific literacy. Sci Educ *87*, 224–240.

Peat M, Taylor CE, Franklin S (2005). Re-engineering of undergraduate science curricula to emphasise development of lifelong learning skills. Innov Educ Teach Int *42*, 135–146.

Quitadamo IJ, Kurtz MJ (2007). Learning to improve: using writing to increase critical thinking performance in general education biology. CBE Life Sci Educ *6*, 140–152.

Townsend MA (2002). Writing across the curriculum. In: The Allyn and Bacon Source Book for Writing Program Administrators, ed. I Ward and W Carpenter, New York: Longman, 264–274.

Trochim WMK, Donnelly JP (2007). The Research Methods Knowledge Base, 3rd ed., Cincinnati, OH: Atomic Dog Publishing.

VanDerZanden AM (2010). Reflective writing as an assessment for student analysis and synthesis ability in a landscape design course. HortTech *20*, 668–671.

Yalvac B, Smith HD, Tloy JB, Hirsch P (2007). Promoting advanced writing skills in an upper-level engineering class. J Eng Educ *96*, 117–128.