TOWARD UNDERSTANDING WRITING TO LEARN IN PHYSICS: INVESTIGATING STUDENT WRITING

DISSERTATION

Presented in Partial Fulfillment of the Requirements for
the Degree Doctor of Philosophy in the
Graduate School of The Ohio State University

By

Dedra Demaree, B.S., M.S.

* * * * *

The Ohio State University

2006

Dissertation Committee:
Gordon Aubrecht, Co-Adviser
Lei Bao, Co-Adviser
Evan Sugarbaker
Dongping Zhong

Approved by

Co-Adviser

Co-Adviser

Graduate Program in Physics
ABSTRACT

It is received wisdom that writing in a discipline helps students learn the discipline, and millions of dollars have been committed at many universities to supporting such writing. We show that evidence for effectiveness is anecdotal, and that little data-based material informs these prejudices. This thesis begins the process of scientific study of writing in the discipline, in specific, in physics, and creates means to judge whether such writing is effective. The studies culminating in this thesis are an aggressive start to addressing these complex questions.

Writing is often promoted as an activity that, when put into classrooms in specific disciplines, not only helps students learn to write in the methods of that discipline but also helps students learn content knowledge. Students at the Ohio State University are being asked to write more in introductory courses, and the Engineering schools want their students to have more writing skills for the job market. Combined with the desire of many educators to have students be able to explain the course content knowledge clearly, it would seem that writing activities would be important and useful in physics courses. However, the question of whether writing helps learning or whether students learn writing within a non-English classroom helps learning in the discipline are open to debate, and data are needed before such claims can be made.

This thesis presents several studies aimed at understanding the correlation of writing and content, and tracking and characterizing student writing behaviors to see
how they are impacted by writing in physics courses. It consists of four parts: summer and autumn 2005 focus on writing in introductory physics labs with and without explicit instruction, while winter and spring 2006 focus on tracking and analyzing student writing and revising behavior in Physics by Inquiry (PbI).

With these related projects, we establish three main results. First, there is a need for quantitative studies of Writing to Learn, and in specific of Writing to Learn within physics. Second, we have also made progress in characterizing student behaviors in an effort to quantify the study of writing: the link between writing and learning content is not obvious, and we have shown that students may not even be learning to write through practice in the context of physics. Third, we have developed a valuable new tool, a novel program to track and analyze student writing, that supplies quantitative information about student writing.

In Summer quarter, 2005, introductory calculus-based physics lab students wrote essays, some sections with and some without explicit writing instruction. When analyzing student writing, we found that an essay’s grade based solely on English correlated strongly with its grade based on physics, and that explicit writing instruction improved the physics content more than writing practice alone. In addition, we have studied the location and type of comments made by both physics and English instructors on individual student essays. We find a strong correlation between the location and type of comments made by both instructors, and find that when students struggle overall with the content of an essay they make more mistakes with the writing. A similar study in Autumn quarter, 2005, reduced the writing to one specific physics content area, and increased the content knowledge testing to look for changes in content knowledge with the writing activities, and differences in content knowledge
based on the inclusion of explicit writing instruction. Limited impact of the writing or instruction on content knowledge was observed; students who completed the writing activities in place of traditional lab activities did better on lab quizzes immediately following the activities, but not on quizzes and tests taken later. However, this null result indicates that time can be devoted to writing within an introductory physics class without reducing the focus on physics content. Differences between those who wrote with and without explicit writing instruction were minimal, which may be due to the shorter time spent on writing during this project.

In Winter quarter, 2006, our focus shifted toward tracking writing behavior, characterizing student writing, and looking for any impact of the writing on these behaviors. Students in Physics by Inquiry wrote essays requiring two drafts; the second being half the length, which required the students to make significant revisions. Although we learned a lot about how to do such studies, we did not observe many consistent significant correlations. It did not appear there were strong correlations between writing behaviors and essay grades, nor were there changes in student behavior over time.

We developed a novel tracking program that allows us to obtain much more data in a more ideal fashion than existing programs, and also automates part of the analysis process, allowing larger studies needed for generalizing behaviors because they are no longer be time prohibitive. We used this tracking program for the same student population and assignment type in Spring quarter, 2006. We were able to better characterize student writing and observe behavior consistent with the literature: students who had better grades exhibited more revisions, and more non-trivial revisions. However, these results were not consistent week to week, and we saw no change in writing
behavior through the quarter. Case studies presented help us validate the information we obtain from the tracking, and suggest changes we can make for improved further studies.
ACKNOWLEDGMENTS

The author would like to thank her co-advisors: Dr. Aubrecht and Dr. Bao. Having two advisers has allowed this work to have more breadth and depth with complementary suggestions and contributions from each. Their support through the last few years has also made the many months of working way too much worthwhile and bearable.

The author also could not have completed this project without the guidance of Scott Franklin and Lisa Hermsen at the Rochester Institute of Technology. They originated the idea of providing writing instruction and looking at the correlations between the writing and physics comments, as well as the idea of the shortening an essay in the second draft that we used in Physics by Inquiry. Their initial work inspired this thesis, and they acted as consultants throughout each project.

The author would also like to thank Cat Gubernatis for the large amount of time and effort she gave for developing writing lesson plans, giving instruction within the physics classroom, and grading essays. The summer and autumn quarter projects would not have been possible without her help. In addition the author thanks Jessica Hanzlik who coded the essay comments from summer quarter, which was tedious but important for this project. Also the author thanks Christopher Manion, who consulted on the project from the perspective of a writing across the curriculum
specialist. The author also thanks Wenhui Zhao for writing the tracking program we used in the spring quarter data.

From a research perspective the author would like to thank the Ohio State University Physics Department for cooperation with implementing changes to the courses for these studies. Specifically, Kathy Harper for allowing us to do this project in her summer quarter class, and putting a research question on the final exam. The author also thanks Dr. Sugarbaker and Dr. Adelson for approving and being supportive of me testing writing in the introductory physics labs. In addition, the Physics by Inquiry instructors provided access to their students and spent time helping develop appropriate essay topics.

The author would also like to thank her friends and family for general support and assistance with aspects of life that would have otherwise been neglected while writing this thesis. I may have been impossible for me to finish this thesis, move across the country, and start my new job in such a short time span without the support of all four parents, big sister, and friends.

Most importantly, the author would like to thank the students for their participation and cooperation with this project.
VITA

September 1, 1974.........................Born - Washington, D.C., USA

1999.................................B.S. University of Maryland Baltimore County

2002.................................M.S. The Ohio State University

1999-2003..........................Graduate Teaching Associate,
                               The Ohio State University.

2003-Present........................Lab Sergeant,
                               The Ohio State University.

PUBLICATIONS

D. Demaree, Y. Lin “Assessing ISLE labs as an enhancement to traditional large-lecture courses at the Ohio State University”. Physics Education Research Conference Proceedings, Salt Lake City, 2005.


FIELDS OF STUDY

Major Field: Physics
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The Importance of Writing to Learn</td>
<td>1</td>
</tr>
<tr>
<td>1.1</td>
<td>Is Writing to Learn Effective?</td>
<td>2</td>
</tr>
<tr>
<td>1.2</td>
<td>National History of WAC and WID</td>
<td>4</td>
</tr>
<tr>
<td>1.3</td>
<td>Writing at the Ohio State University</td>
<td>6</td>
</tr>
<tr>
<td>1.4</td>
<td>The Debate Surrounding Writing to Learn</td>
<td>8</td>
</tr>
<tr>
<td>2</td>
<td>Theories and Studies Regarding Writing to Learn</td>
<td>12</td>
</tr>
<tr>
<td>2.1</td>
<td>How Writing May Help Learning</td>
<td>13</td>
</tr>
<tr>
<td>2.2</td>
<td>Cognitive Theories of Writing Processes</td>
<td>19</td>
</tr>
<tr>
<td>3</td>
<td>Writing in Science Studies</td>
<td>26</td>
</tr>
<tr>
<td>3.1</td>
<td>Teresa Larkin-Hein [46, 55, 54]</td>
<td>28</td>
</tr>
<tr>
<td>3.2</td>
<td>John Jewett, Jr. [48]</td>
<td>30</td>
</tr>
<tr>
<td>3.3</td>
<td>Larry Kirkpatrick [50]</td>
<td>31</td>
</tr>
<tr>
<td>3.4</td>
<td>William Mullin [66]</td>
<td>33</td>
</tr>
<tr>
<td>3.5</td>
<td>Scott Franklin and Lisa Hermsen[38]</td>
<td>34</td>
</tr>
</tbody>
</table>
3.6 Related Physics Research ........................................ 35
3.7 Writing in Mathematics ........................................ 36
3.8 OSU’s Writing in Physics Studies .............................. 40

4. Summer 2005: An English Instructor in the Physics Classroom ...... 43

4.1 Implementation Details ........................................ 43
  4.1.1 Writing Activities ...................................... 45
  4.1.2 English Writing Instruction .............................. 46
  4.1.3 Grading ................................................... 47

4.2 Data Obtained ................................................ 47

4.3 Difficulties We Observed ..................................... 51

4.4 Results ...................................................... 52
  4.4.1 Observed Changes in Writing .............................. 52
  4.4.2 Quantitative Data ...................................... 53
  4.4.3 CSEM and EBAPS Results ............................... 53
  4.4.4 Data from Essay Content and Grading ................... 56
  4.4.5 Final Exam Data ........................................ 61
  4.4.6 Other Possible Analysis ................................ 65

4.5 Conclusions ................................................ 65

5. Autumn 2005: Can we Teach One Concept through Writing ......... 67

5.1 Introduction ................................................ 67

5.2 Implementation Details ..................................... 68
  5.2.1 Writing Activities ...................................... 69
  5.2.2 English Writing Instruction .............................. 70
  5.2.3 Grading ................................................... 71

5.3 Difficulties We Observed ................................... 72

5.4 Results ...................................................... 72
  5.4.1 Essay Grades ........................................... 73
  5.4.2 Lab quizzes .............................................. 76
  5.4.3 CSEM and DIRECT ....................................... 80
  5.4.4 Recitation Quizzes and Course Exams ................... 81

5.5 Discussion and Conclusions ................................ 83

6. The Study of Revisions ........................................ 85

6.1 Specific Studies on Revisions ................................ 89
  6.1.1 Huub Van Den Bergh and Gert Rijlaarsdam [21] .......... 90
  6.1.2 Lillian Bridwell [23] .................................... 90
  6.1.3 Eva Lindgren and Kirk P.H. Sullivan [58] ................ 92
11. Conclusions ......................................................... 230
   11.1 Results of Our Studies ...................................... 231
   11.2 Future Directions ........................................... 237

Appendices:

A. Essay Assignments and Lesson Plans for Summer Quarter .............. 240
   A.1 Essay and Prompt for the First Assignment ...................... 240
   A.2 Essay and Prompt for the Second Assignment .................... 244
   A.3 Essay and Prompt for the Third Assignment ...................... 247
   A.4 Essay and Prompt for the Fourth Assignment .................... 250
   A.5 Essay and Prompt for the Fifth Assignment ..................... 253
   A.6 First Explicit Writing Instruction ............................. 255
   A.7 Second Explicit Writing Instruction ............................ 257
   A.8 Third Explicit Writing Instruction ............................. 259
   A.9 Fourth Explicit Writing Instruction ............................. 262

B. Essay Assignments and Lesson Plans for Autumn Quarter .............. 265
   B.1 Essay and Prompt for the First Assignment ...................... 265
   B.2 Essay and Prompt for the Second Assignment .................... 266
   B.3 Essay and Prompt for the Third Assignment ...................... 268
   B.4 First Explicit Writing Instruction ............................. 268
   B.5 Second Explicit Writing Instruction ............................ 270
   B.6 Third Explicit Writing Instruction ............................. 271

C. Quiz and Exam Questions, Fall 2005 .................................. 276
   C.1 First Lab Quiz .................................................. 276
   C.2 Second Lab Quiz ................................................ 277
   C.3 Third Lab Quiz .................................................. 279
   C.4 Course Exam Questions ......................................... 280

D. Essay Assignments for Winter and Spring Quarters ...................... 282
## LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1 Division of time on lab activities; Summer 2005</td>
<td>44</td>
</tr>
<tr>
<td>5.1 Division of time on lab activities; Autumn 2005</td>
<td>69</td>
</tr>
<tr>
<td>5.2 Correlations between English and physics grades for selected essays</td>
<td>76</td>
</tr>
<tr>
<td>7.1 Correlation coefficients for Winter 2006 essay grades. Significance levels are in parenthesis, and those marked with an asterisk are $\rho &lt; 0.05$.</td>
<td>110</td>
</tr>
<tr>
<td>7.2 Results from Revision Analysis on First Drafts for Spring quarter, 2006</td>
<td>119</td>
</tr>
<tr>
<td>7.3 Results from Revision Analysis on First Drafts for Spring quarter, 2006</td>
<td>120</td>
</tr>
<tr>
<td>7.4 Results from Revision Analysis on First Drafts for Spring quarter, 2006</td>
<td>121</td>
</tr>
<tr>
<td>9.1 Correlation coefficients for Spring 2006 essay grades. Significance levels are in parenthesis, and those marked with an asterisk are $\rho &lt; 0.05$. Note that no significant differences are given for correlations with the change in grade from assignment two because only two data points were available for that analysis.</td>
<td>149</td>
</tr>
<tr>
<td>9.2 Results from Mann-Whitney U and Kurskal-Wallis tests on Essay Grades for Spring quarter, 2006</td>
<td>151</td>
</tr>
<tr>
<td>9.3 Results from Revision Analysis on First Drafts for Spring quarter, 2006</td>
<td>153</td>
</tr>
<tr>
<td>9.4 Results from Revision Analysis on Second Drafts for Spring quarter, 2006</td>
<td>158</td>
</tr>
<tr>
<td>9.5 Results from Revision Analysis on First Drafts for Spring quarter, 2006</td>
<td>162</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Bereiter and Scardamalia model of “knowledge telling” strategy [20]</td>
<td>21</td>
</tr>
<tr>
<td>2.2</td>
<td>Bereiter and Scardamalia model of “knowledge transforming” strategy [20]</td>
<td>22</td>
</tr>
<tr>
<td>2.3</td>
<td>Flower and Hayes model of composition [44]</td>
<td>24</td>
</tr>
<tr>
<td>4.1</td>
<td>A sample of student writing with grading comments based on physics.</td>
<td>49</td>
</tr>
<tr>
<td>4.2</td>
<td>A sample of student writing with grading comments based on English.</td>
<td>50</td>
</tr>
<tr>
<td>4.3</td>
<td>Weekly essay grades based on physics.</td>
<td>54</td>
</tr>
<tr>
<td>4.4</td>
<td>Percentage of times physics and English comments at same location were of the same type.</td>
<td>58</td>
</tr>
<tr>
<td>4.5</td>
<td>Percentage of English positive-based comments for both groups.</td>
<td>59</td>
</tr>
<tr>
<td>4.6</td>
<td>Percentage of sentences of each type for the WI group.</td>
<td>62</td>
</tr>
<tr>
<td>4.7</td>
<td>Percentage of speculative sentences for both groups.</td>
<td>63</td>
</tr>
<tr>
<td>5.1</td>
<td>English grades for writing activities.</td>
<td>74</td>
</tr>
<tr>
<td>5.2</td>
<td>Physics grades for writing activities.</td>
<td>75</td>
</tr>
<tr>
<td>5.3</td>
<td>Weekly lab quizzes for each circuits lab and the previous week.</td>
<td>78</td>
</tr>
<tr>
<td>7.1</td>
<td>This shows the number of events versus time for students who scored high, medium, and low on their final exams.</td>
<td>119</td>
</tr>
</tbody>
</table>
7.2 This shows the number of events over each second draft for students who had high, medium, and low scores on their final exam.

7.3 This shows the time on task over each second draft for students who had high, medium, and low scores on their final exam.

7.4 This shows the number of events versus time for students who scored high, medium, and low on their essay.

7.5 This shows the number of events versus time for students who scored high, medium, and low on their fourth essay assignment.

7.6 This shows the number of events versus time for students who had high, medium, and low changes in their grade from the first to the second draft.

7.7 This shows the number of events versus time for students who had high, medium, and low changes in their grade from the first to the second draft.

7.8 This shows the number of events of each size for students who had high, medium, and low scores on their final exam.

7.9 This shows the number of events of each size for students who had high, medium, and low scores on their detailed essay grades.

7.10 This shows the number of events of each size for each of the four second draft essay assignments.

8.1 Example of S-notation.

8.2 Snapshot of the online version of EditNav.

9.1 This shows the number of events versus the time on task for all first drafts. The data are grouped based on the students’ final exam scores.

9.2 This shows the number of revision events for each of week of the first draft essay assignments. The data are grouped based on the students final exam scores.
9.3 This shows how much time on task was spent for each of week of the first draft essay assignments. The data are grouped based on the students’ final exam scores.

9.4 This shows the number of events versus the time on task for all second drafts. The data are grouped based on the students’ final exam scores.

9.5 This shows the number of revision events for each of week of the second draft essay assignments. The data are grouped based on the students’ final exam scores.

9.6 This shows how much time on task was spent for each of week of the second draft essay assignments. The data are grouped based on the students’ final exam scores.

9.7 This shows the number of events versus the time on task for all first drafts. The data are grouped based on the students’ essay grades.

9.8 This shows the number of events versus the time on task for all second drafts. The data are grouped based on the students’ essay grades.

9.9 This shows the number of events versus the time on task for all second drafts. The data are grouped based on the change in the essay grades from the first to the second draft.

9.10 This shows the average size of each revision event for all first drafts. A 1-word revision includes when students fix typos while writing. The data are grouped based on the students’ final exam scores.

9.11 This shows the average size of each revision event for all first drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ total essay grades.

9.12 This shows the average size of each revision event for each first draft. A 1-word revision includes students fixing typos while writing. The data are grouped based on the assignment number, which also represents progression of time through the quarter.

9.13 This shows the average location of each revision event for all first drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ final exam scores.
9.14 This shows the average location of each revision event for all first drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ total essay grades. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 180

9.15 This shows the average location of each revision event for each first draft. A 1-word revision includes students fixing typos while writing. The data are grouped based on the assignment number, which also represents progression of time through the quarter. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 181

9.16 This shows the average size of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ final exam scores. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 183

9.17 This shows the average size of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ total essay grades. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 184

9.18 This shows the average size of each revision event for each second draft. A 1-word revision includes students fixing typos while writing. The data are grouped based on the assignment number, which also represents progression of time through the quarter. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 185

9.19 This shows the average size of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the change in grade from the first to the second draft. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 186

9.20 This shows the average location of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ final exam scores. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 188

9.21 This shows the average location of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ total essay grades. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 189

9.22 This shows the average location of each revision event for each second draft. A 1-word revision includes students fixing typos while writing. The data are grouped based on the assignment number, which also represents progression of time through the quarter. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 190

xviii
9.23 This shows the average location of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the change in grade from the first to the second draft. ............................... 191

10.1 Sample of approaches to first draft writing. ............................... 199

10.2 This graph shows the location of the revisions (as the percent of the way through the text written up to that point) in order of the sequence of revision events for C1. The size of the dots represents the number of words that were added or deleted during that revision event. ............................... 200

10.3 Sample of approaches to first draft writing. ............................... 202

10.4 This graph shows the location of the revisions (as the percent of the way through the text written up to that point) in order of the sequence of revision events for C1. The size of the dots represents the number of words that were added or deleted during that revision event. ............................... 203

10.5 This graph shows how C1 revised her essays when writing the second drafts. ............................... 205

10.6 This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C2. The size of the dots represents the number of words that were added or deleted during that revision event. ............................... 209

10.7 This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C2. The size of the dots represents the number of words that were added or deleted during that revision event. ............................... 210

10.8 This graph shows how C2 revised his essays when writing the second drafts. ............................... 212

10.9 This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C3. The size of the dots represents the number of words that were added or deleted during that revision event. ............................... 216
This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C3. The size of the dots represents the number of words that were added or deleted during that revision event.

This graph shows how C3 revised her essays when writing the second drafts.

This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C4. The size of the dots represents the number of words that were added or deleted during that revision event.

This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C4. The size of the dots represents the number of words that were added or deleted during that revision event.

This graph shows how C4 revised her essays when writing the second drafts.

A.1 Electrostatic Precipitator

A.2 Mini-light Shunt

B.1 Two circuits for first writing prompt, Autumn quarter

B.2 Two circuits for second writing prompt, Autumn quarter

C.1 Circuit for quiz I question 1

C.2 Circuit for quiz II question 1

C.3 Circuit for quiz II question 3

C.4 Circuit for quiz II question 4

C.5 Circuit for quiz III question 3
CHAPTER 1

THE IMPORTANCE OF WRITING TO LEARN

From elementary school through college, writing is strongly emphasized as a skill students need to learn. The Association of American Universities has committed millions of dollars toward improving student writing skills. Despite this emphasis, many students do not have the writing skills needed for entering the professional work force. The United States Department of Education states that 73% of employers rate the writing skills of recent high school graduates as fair or poor [1]. Since around 1970 universities across the country have promoted Writing in the Disciplines (WID), then Writing Across the Curriculum (WAC) (around 1980). Current estimates state that between 20 and 50 percent of higher education institutions have WAC programs [72]. These programs emphasize the use of writing outside the English Department, often promoting discipline-specific writing-intensive courses. However, the educational effectiveness of these programs typically has not been rigorously tested, and most claims in the literature about the benefits of writing to learn are accepted as received wisdom rather than supported scientific fact.

The purpose of introducing writing in the classroom is two-fold: for students to improve their writing abilities, and to gain discipline-specific knowledge through writing activities. “Advocates of Writing Across the Curriculum have long contended

1
that incorporating writing into classes in any discipline can enrich content instruction rather than supplant or burden it [2].” WID is learning to write in the ways specific to disciplines (“learning to write”), while WAC is writing about the subjects a discipline studies (“writing to learn”). In the 1970s and 1980s, WID was seen as integral to WAC, but starting in the 1990s, WID started to become seen as something separate from WAC. However, the split between WAC and WID can not be complete, because what counts as good writing in a field is largely determined by the questions, goals, methods, and epistemologies of that field. This keeps writing to learn tied deeply to the understanding of the field itself.

1.1 Is Writing to Learn Effective?

Emig claims that “by its nature writing forces integration of ideas, requires the establishment of relationships, provides immediate and tangible feedback, and forces personal involvement with the material [29].” Other authors suggest that writing involves conscious exploration of information, and/or the joining of information into relationships, creating new knowledge. Ref. [76] gives a brief review of evidence for writing to learn and quotes Langer “few studies have been undertaken to learn what people learn from writing, what different kinds of learning result from different kinds of writing experiences, or how writing can be used to help students understand and remember the material they have read [52].”

As for specific writing activities, there is little research-based evidence to show whether they aid learning within the disciplines. The majority of publications discussing innovative uses of writing in the classroom consist of discussing the activities, stating why the instructor feels they are beneficial, and giving anecdotal evidence
for their benefit such as student statements they were a good addition to the course. There are two major problems with this type of publication. First, we do not actually know if the activities were successful; there is usually no support for the concluding statements made. Second, if they actually were successful, we do not know why or what specific aspect of the activity was the key, which makes it impossible to know if this activity could also be successful in a different environment.

It is easy to fall into this trap of stating how helpful writing is without providing evidence. I’m sure each of us has experienced a moment of clarity while trying to explain something, whether orally, say to a friend, or in writing while working on our own publications. As students I have interviewed have stated: writing would help learning, it’s like peer tutoring, when you have to explain something to someone else it helps you understand it yourself. However, there is good reason to question this. First, we are educated people; we are used to reflecting on what we are saying. Do our students reflect when we ask them to write something? Second, we usually already have the basic facts before we have that “ah hah” moment. After all, we don’t try tutoring a peer unless we already have some idea of how to answer the question. Our students may need a grasp on the subject before they can use writing to solidify that grasp. But how much of a grasp do they need before writing can be effective? Third, students exhibit a range of learning style preferences. Some learn well reading and writing while others do not. Further, many students do not learn the way the average faculty member prefers to learn, so what is to dictate that they will have this same experience that we have?

There is reason from other areas research to believe that writing may help learning. Research shows that it is helpful for students to have a positive epistemology [28,
Writing may help students understand that physics is about drawing logical connections. This could help students get out of the attitude that to succeed in physics one just needs to memorize equations. Also, research supports the need to teach scientific abilities and thought processes explicitly [31]. Making logical written arguments may help develop these abilities. There is evidence that it is helpful for students to build a hierarchy of ideas [53, 62], and the logical structure of writing may help with that. Research supports the need for active engagement to improve learning [41]. Writing could be an active engagement activity, and some people who use writing in their classes claim that it is.

Based on the literature and knowledge about active learning, my speculation on these possible benefits is that writing won’t be as helpful unless students are reflective while they write. Getting students to be reflective, however, is not a trivial matter! Hillocks [39] devotes an entire text to how to teach writing as a reflective practice. He discusses theories and cites examples showing writing that lacks reflection and writing that contains reflection. It is his premise not only that writing must be taught in order to be a method for creating meaning, but that it must be taught explicitly as a reflective process.

1.2 National History of WAC and WID

In early American education settings, explicit writing instruction was not emphasized. Instead, the focus was on oratory skills. In fact, classroom education consisted primarily of students orally responding to a teacher’s questions. In addition, the curricula were not specialized as they are today; there were not departments for different subjects. All students took the same curriculum, and a professor might teach each
of the offered courses: Latin, Greek, mathematics, rhetoric, and moral philosophy. 

After the Civil War, enrollment in colleges drastically increased, and new colleges and land-grant universities were founded. These schools moved toward a system in which students could choose which courses to take, and since students were being trained for a variety of professions, education became more departmentalized. With each specialization there developed a unique style of writing specific to the terminology, publications, and communications of that field. With the enlargement of the universities also came the advent of written exams instead of oral exams, since there were too many students to devote time to individual examinations [72].

With these changes, it became evident that students needed explicit instruction in writing, and the first composition courses started appearing in the early 1870s. By the late 1800s, the education system had shifted to one we are more familiar with today: including writing courses for students over the entire span of college years. However, while explicit instruction in writing became more prevalent, the number of courses with a strong component based on communication skills decreased drastically by the start of the 20th century (with the exception of liberal arts schools, which tended to keep an emphasis on these). “As writing became one more subject among many, it ceased to be a central part of all of them. Despite pious pronouncements about every teacher being an English teacher, responsibility gradually shifted from the whole faculty to the English department [72]” By around 1910, “the academic community had ceased to be a community in the sense that those raised in the oral, face-to-face culture of rural and small-town America understood the term. The college had become divided, rationalized, efficient, with knowledge committed to specialized
writing, though students had no clear means of mastering those written conventions [72]."

The academic view toward writing remained fairly steady until the 1970s, when people started to emphasize the need to teach discipline-specific writing, as well as espouse the use of writing as an educational tool within disciplines. WAC initiatives received many grants in the 1980s and programs became prevalent throughout colleges and universities in the United States. However, by the ‘90s education priorities shifted toward other issues such as assessment, and there was far less funding available for WAC programs.

Since the 1990s, the most popular WAC courses have essentially been regular courses within a discipline with the addition of writing requirements (similar to Physics 367 at OSU). It is common for a university to have 2 or more such courses as graduation requirements. Also in this period, writing centers started to collaborate or manage the WAC programs. Traditionally, WAC programs seldom involved full departments; rather they focused on involving individual faculty members and developing specific courses. However, more programs now focus on working with the whole department to set goals.

1.3 Writing at the Ohio State University

At the Ohio State University (OSU), we mirror the larger picture painted above. In addition to a strong writing center, we have had an active WAC program at OSU since the 1980s, with a full time coordinator, Dr. Christopher Manion. OSU undergraduates are required to take 3 writing courses, with only the first of those given in the English Department. This structure has been the standard at OSU since
around 1988. The physics department at OSU offers 2 writing courses: Physics 367 and Physics 595. Physics 367 is primarily a course for non-physics majors, in which students learn presentation skills and write one long paper while studying the use of energy resources in the US. Physics 595 is for physics majors, and aims to give them communication skills needed specifically for the physics job market, including scientific paper writing, conference-style talks, and interview skills. Despite these writing-intensive courses offered by the physics department, writing is not emphasized in any other lower-level physics courses, and is virtually absent from the introductory courses.

From my observations at OSU I can speculate on why writing is not emphasized in our introductory courses despite the general attitude that it would be helpful. First, traditionally our introductory courses put focus on problem solving; not only is the bulk of assignments based in this, but so is the bulk of the testing. If we were to give writing assignments it would take away from the time students spend doing problems, and consideration would have to be given on how to modify the course exams. Written work also takes longer for people to grade, and in my experience the grading can be very subjective, so grading by teaching assistants would have to be monitored more carefully than for standard problems. In addition, although instructors will often state that writing is helpful, it is probably not one of the things they would list under most important goals for their course. This may be because of a lack of awareness that content can often be learned in the context of skill-building exercises. Another difficulty in implementing writing is that students often lack the prerequisite skills needed to write effectively. Because writing activities are already time consuming, instructing students in the necessary writing skills is certainly not easy to implement.
Despite OSU’s overall commitment to a strong writing program, there is evidence that students are not graduating with strong enough writing skills. In the words of Dr. Manion: “I have been approached by professors, instructors, and administrators in the sciences and engineering who are concerned that their students have difficulty communicating knowledge in their discipline. Furthermore, they are frustrated that students cannot critically apply the scientific methodologies they are teaching; they can “plug in” numbers to equations and memorize textbook facts, but they have a hard time coming to a critical understanding of scientific concepts.” In addition, OSU’s College of Engineering, recently polled its alumni and their employers about the quality of education graduates received in preparing them for the work force. The college overwhelmingly found that both groups wished graduates were better prepared to communicate the knowledge they had learned and to critically use that knowledge to solve new problems. The Engineering College wants to see more writing in the general education requirements to help make up for this deficiency. In addition, as part of a total revamping of the core curriculum required of OSU undergraduates, the McHale Committee has recommended an increase in the university’s emphasis on writing.

1.4 The Debate Surrounding Writing to Learn

There is a larger debate about writing to learn within the writing community. WAC and WID have been the focus of serious debate, review, and many publications. Some of these papers espouse the benefits of writing to learn, and show examples of innovative uses of writing in the classrooms. Others criticize the lack of evidence to show if writing helps learning. General reviews of these programs as a whole are open
to the following criticisms. Because these programs emphasize the use of writing outside the English Department, they tend to not have a strong management or backing, since it can be unclear who is responsible for them. Also, these courses tend to be developed by individual faculty members, and once they move on (or teach other courses) the courses tend to be dropped or their focus altered. Perhaps the biggest problem with these courses is they tend to require a very large time commitment, both from the instructors and the students. Writing activities tend to take longer for students, require more feedback to be successful, and take more time to grade than standard exercises. For reviews of writing to learn at multiple institutions, see references [12, 71]. We are fortunate at OSU to have a coordinator for our WAC program, but even for him, it is difficult monitoring and supporting courses across the campus.

In a paper titled “The Promise of Writing to Learn” by John Ackerman [12], he reviews publications that study the use of writing outside English. He concludes that “writing as a mode of learning is at best an argument yet to be made” and “the mixed message in the empirical research could be a clear one if we reject the inevitability of writing and learning.” In other studies, it is found that writing to learn efforts in science, math, and economics have not been analyzed in terms of either the collective importance of their findings or in terms of dissemination. “Whether or not the professions are moving, or will move, in the direction of writing in the discipline is an important but unanswered question [25].” There has also been more push lately from administrators and funding agencies to assess the effectiveness of WAC programs. “This posed thorny problems as it is devilishly difficult to measure either
writing or learning and even more difficult to measure whether the changes are due
to the WAC efforts [72].”

An interesting thing to note about research on writing and learning in the 1990s
is the shift toward “naturalistic” qualitative studies. Russell ([72] states that this is
because of the confusing results that came from quantitative studies. These studies
found that writing does not automatically improve learning or thinking. Qualitative
research on writing and learning to write is influenced by cultural-historical ethnographic,
and discourse-analysis studies of literacy. Studies, for example, have shown
how what students expect often does not agree with what teachers expect; studies
also help to understand how students gain authority through writing, and other stud-
ies show how students transition to becoming professional writers. The consequence
of this is a dearth of quantitative studies reported in the literature.

“The consequences of not exploring a broad range of methodologies as McCarthy
could be dire for the study of writing pedagogy. In a recent article provocatively ent-
titled ‘The War on Scholarship,’ Richard Haswell complains that Composition’s most
prominent professional organizations, the National Council for Teachers of English
and the Conference on College Composition and Communication, have systematically
ignored and de-valued empirical research, which he defines as replicable, agreeable,
and data-supported research. Without such research, he argues, the field of composi-
tion can not build ‘a coherent body of testable knowledge’ to convince scholars from
other disciplines that the claims we make about writing are more than just ‘private
epideictic,’ a pie-eyed fad in higher education that will go away when people decide
it does not work. Indeed, his diagnosis for the field is grim: ‘As when a body un-
dermines its own immune system, when college composition as a whole treats the
data-gathering, data-validating, and data-aggregating part of itself as alien, then the whole may be doomed[43]' [2]."
CHAPTER 2

THEORIES AND STUDIES REGARDING WRITING TO LEARN

Despite the reluctance to approach writing to learn in a quantitative manner, there are models in the literature for ways to analyze student writing quantitatively. One often-cited researcher, Marlene Scardamalia [75], has done extensive work on teaching writing to children in sixth grade. The instructions she gives are based on having them generate ideas, improve them, elaborate on them, identify their goals, and form a cohesive whole from their ideas. She uses the cognitive apprenticeship model, and works from first demonstrating the writing techniques to the students, to having them use the techniques with the help of cue cards. For example, one planning cue is “a good point on the other side of the argument is...,” and another is “I could describe this in more detail by adding...” With her students, she has measured an increased frequency of idea-level revisions, increased time spent planning, and increased reflectivity about the writing (as determined by interviews with the students). In addition, some authors have attempted to determine how writing may help learning, both from a basis English studies and Cognitive Science studies.
2.1 How Writing May Help Learning

In [76], Gary Schumacher and Jane Nash emphasize a need to formulate theories for ways writing may help learning, and to use better methods for measuring the types of learning that may be taking place. They rely heavily on research from cognitive science to make these lists. They note that the subject is not easy to study as different kinds of writing may result in different kinds of learning. They suggest a question that I have also pondered with respect to this topic “what do we mean by learning or knowledge change?” To address this, they suggest different types of learning: incorporating new information into existing knowledge structures, modifying or refining existing knowledge structures, and restructuring knowledge entirely. Writing probably does not help with the first task, since writing itself does not help with fact gathering.

One way of changing knowledge structure is through cognitive conflict. New structures must be created to incorporate new knowledge without conflict with existing structures. Whether this occurs depends on whether students realize why there is a conflict, their interest in resolving that conflict, and their realization that it can not be resolved within their current structure [76]. The authors suggest writing may help with “mentally wrestling with alternative interpretations of a set of findings” in certain situations. They suggest giving writing tasks in which the writers are presented with conflicting information or require them to synthesize conflicting positions.

Another method to support knowledge structure change is through analogies, metaphors, and models. Metaphors make information available by implicit comparison between domains of knowledge. The authors in [76] note that analogies have been
powerful for theory construction in science. They are also helpful for transfer of knowledge. Writing may force a comparative process, especially with comparison/contrast essays. Writing that involves dealing directly with analogies or metaphors may help with conceptual change, for example, having students explain or criticize an analogy.

Another way to cause change in knowledge structure is through construction of multiple representations. Ref. [76] claims that writing involves this on four levels: nonverbal and procedural representations (illustrations), metaphors, the writing plan, and the final text itself. They say that it is also important to create a flexible knowledge base to aid with transfer of knowledge. If knowledge representations involve a high degree of interconnectedness then it is easier for it to be transferred to new contexts. Ref. [76] suggests that having the new perspective of the content in the context of writing aids in this.

The authors in [76] also discuss some methods to test knowledge gain via testing knowledge structure. One method is to measure reaction time: to see how fast a writer responds to stimulus in pre- and post-writing. This helps to see whether and how quickly new knowledge structures form through writing, and how stable they are. Other methods suggested attempt to probe the knowledge structures themselves. One of these is ordered trees, which reflect the organization of existing knowledge structures. Another is multidimensional scaling, in which the dimensions that describe the content show the depth of the knowledge structure. Another is hierarchical clustering, which is similar but assumes knowledge is clustered, but not lying on separate dimensions. Last, the authors discuss weighted networks, which are similar but do not rely on a hierarchical structure. They note that structures of experts are more simple than those of novices, and the links are highly integrated.
“Writing is one of the most appealing of instructional activities. At the very least, frequent composing should help students to become better communicators... advocates of writing-to-learn share the conception that ‘writing’ means creating texts that explore the relationships among ideas. They often contrast composing with restricted writing, in which students write brief texts within narrowly prescribed formats... [they] also share overlapping conceptions of ‘learning.’ Most refer to students coming to understand relationships among ideas... ‘knowledge transformation’, ‘conceptual change’, ‘discovery’, or the ‘construction of meanings’ [51].” This makes use of the idea of constructivism; for students to understand important curriculum material, they need to reason actively about it. Arthur Applebee [14] in a review article concludes “the question of whether writing contributes to learning had seldom been investigated, and that many published research studies fail to have control groups, pre- and post-measures, or other techniques that would allow them to rigorously demonstrate the effects of writing.”

Articles since Applebee’s have continued to show mixed and unclear results, but the instructional value of writing has been discussed more in the following 15 years [51]. Perry Klein [51] quotes an article by Nancy Fellows that shows writing activities improved conceptual understanding more than other activities for sixth grade science students [33]. Klein also discusses an article [47] that describes how first year university biology students were assigned to write summaries of lectures, and those students showed greater comprehension and problem solving on post-tests. Another article [61] describes how instructors had some students write journals in addition to conventional lab reports. Students described the content, reflected on how they learned it, and evaluated the effectiveness of this learning. The journal students did
score higher on a multiple-choice content-based final exam. These are the studies cited that show writing helps learning. Ref. [51] states that most studies have null results on at least some measures. Klein discusses a study [15] in a second-year physics class that claims students became progressively invested in scientific sense making, yet in the results, the students have more sense-making articulations in the middle of the semester, but revert back to “story-telling” by the end of the semester.

Instead of focusing on whether writing helps learning, Ref. [51] focuses on “when writing contributes to learning, how does it do so?” Klein discusses the main cognitive processes through which writing-to-learn may operate, and the evidence to support them in the research. He organizes the processes into a continuum from spontaneous to deliberative, from requiring little compositional expertise to requiring more, from involving little self-monitoring to extensive monitoring, and from a range of formalities such as personal journals and formal essays. He names four main types of hypothesis. The first is “at the point of utterance”: this states that knowledge is generated as it is communicated. The second is called “forward search”: writers externalize their ideas in text, then reread this text and make new inferences. The third involves using genre structures to organize relationships among elements of text, and linking elements of knowledge. The fourth is called “backward search”: writers set rhetorical goals, then derive content subgoals, and transform their knowledge to accomplish these.

The “shaping at the point of utterance” hypothesis comes out of literature from the early 1980s, and has influenced the WAC movement. Klein’s literature review shows that evidence to support this hypothesis is sparse and largely indirect. People who support this claim that expressive writing is better than drafting and revising, because the spontaneity is necessary for “language to shape meaning.” Bereiter
and Scardamalia challenged this hypothesis with their idea that novice students use “knowledge telling” as a writing technique, which will not itself generate knowledge. The content students retrieve “is already associated linked in long-term memory.” There is also the critique that writing spontaneously would do little if anything to help students confront misconceptions, because it is unreflective and uncritical. In 1970, it was suggested that that speech plays a major role in learning because children organize, explore, and investigate through speech. Since speech is a starting place for writing, this carries over to writing-to-learn. This is in support of the idea that writing spontaneously generates knowledge.

The “forward search” hypothesis also has little empirical support reported in the literature. However, there is indirect support from other research. This hypothesis is that when writers review while writing they “transform their ideas iteratively through operations such as deriving inferences and detecting contradictions.” An idea supporting “forward search” is that there is only so much working memory for processing writing and ideas. Once some ideas are down on paper, that frees up some memory for reviewing and rethinking about ideas. The idea is that this process can actually generate new knowledge. This hypothesis is called “forward” because the writers are constantly transforming their work forward toward a goal. This is in contrast to the “backward search” hypothesis in which the writer sets rhetorical goals, but then revises them to fit the content. This is the only major hypothesis that relies on writing, as opposed to other forms of communication such as speech. Critiques of this hypothesis include a study that showed college students could dictate letters as coherent as those they wrote, and a unique study by Bereiter and Scardamalia in which they had a range of age groups write with inkless pens so they could not
review their work, and found that writing was not hindered. Both of these studies used fairly short writing tasks that may not require as much revision, so they do not disprove the hypothesis.

The “backward search” hypothesis describes a situation in which the writer constructs knowledge by refining rhetorical goals based on the generated content. This is very similar to “forward search” except that the goals are driven by the content, rather than the content being revised by the goals. This is a method used by expert writers, and has been shown to help novice writers when they are taught to apply it; however, the idea that this helps transform knowledge has not been tested empirically. In fact, some research indicates that reflective operations such as goal-setting do not necessarily occur during the writing process [37]. In addition, genres such as argumentation require consistent supportive evidence, but if writers are fitting the goals to the content, they may not build a sound argument.

The “genre” hypothesis states that learning is genre-dependent; for example argumentation requires students to process information deeply and construct relationships. This has been well researched, and is supported by data showing that students must navigate “critical junctures” when they write in specific genres. If this hypothesis is true, it implies that having students work with the same content in multiple genres will encourage them to construct new relationships among ideas. A serious difficulty with applying this hypothesis is that students may not be familiar enough with the genre assigned to make full use of the assigned writing activities. This hypothesis does not stand alone; for example, it does not take into account how prior knowledge is used, and what logical operations occur during the writing process. Literature
supports the idea that different types of writing have different effects on student learning.

There are direct implications of these hypothesis toward using writing in the classrooms. All but the “shaping at the point of utterance” require some knowledge of writing, and therefore writing instruction. Instruction used include prompts such as guiding questions used during the writing process, concept maps or other representations composed before the formal writing, providing sample text or some sentences to build around, or providing explicit instruction about writing and genres.

2.2 Cognitive Theories of Writing Processes

Bereiter and Scardamalia [20] developed and tested two cognitive theories of writing that are cited frequently in the literature. The first of these is the “knowledge telling” strategy. This is a model of how discourse production can occur with the only cues for text production coming from the topic of the writing, the discourse schema, and the text they have already written. Fig. 2.1 shows the structure of this model. It should be noted that knowledge telling is not simply dumping information from memory onto paper: this is not possible since working memory holds much less text than a typical writing assignment. However, this writing does not rely on reflective processes, and is unlikely to generate new knowledge. The second model is the “knowledge transforming” model, with problem-solving spaces in which both belief and knowledge can be worked out through a content problem space, and issues of achieving a goal can be worked out through rhetorical problem space. This model is diagrammed in Fig. 2.2. The output of either of these spaces can be the input to the other, for example a realization of a rhetorical need to clarify can lead to defining a
It should be noted that good writing can come from the “knowledge telling” process, especially if the writer already has the content in mind and organized before writing. It is also believed that both models are used by people, and, in fact, other models may also be used, and different models can even be used by the same person at different times, and in different contexts. The purpose for defining the models is to analyze different tracks that cognitive development can take. Another important point to be made about “knowledge telling” is that this structure is supported in traditional education. Consider the common practices of testing only the content recently taught, presenting reviews and test items in chronological order, assigning papers requiring knowledge assembly (such as summarizing women’s rights), and rewarding students for “showing they learned something” rather than the ability to address new problems. It seems that the “knowledge telling” strategy is reinforced since it is helpful in standard educational settings, and one would have little reason to expand beyond it unless explicitly taught to do so. Success has been achieved both through the use of procedural facilitation (for example, prompts to cue parts of a more advanced writing model), and through goal concretization (prompting the writer to evaluate the text in terms of specific goals). Most importantly, the full extent of the composing process should be modeled by the instructor, with the clear
Figure 2.1: Bereiter and Scardamalia model of “knowledge telling” strategy [20]
Figure 2.2: Bereiter and Scardamalia model of "knowledge transforming" strategy [20]
goal of obtaining higher levels of competence. It is suggested in Ref. [20] that “if students can develop powerful knowledge-transforming skills in writing, this should help them to become more active builders of their own knowledge in all domains.”

It has been noted that there are significant difficulties involved in any writing process. One of these problems is being able to evaluate one’s own text from the reader’s point of view. An example of this is the failure of writers to see what is unclear or vague in their text, since they are already aware of the point they are trying to make (but the reader may not be). In order to write, the writers must be able to generate text without external signals, and to activate the relevant portions of their memory. It is even more difficult to trace down branches of memory until all needed content for conveying a message is retrieved. For people to write verbatim from a memory construct, writing also requires planning at a larger level. In addition to these, are the mechanics of writing: handwriting, spelling, and grammar. Being able to coordinate all the tasks and content involved in writing is difficult, and studies have shown that few people can coordinate more than four ideas at a time. This begs the question “do younger writers follow a knowledge-telling model of composing because they can only coordinate two ideas at a time in writing, or do they coordinate only two ideas at a time because they are following a knowledge-telling model that does not require any higher level of coordination [20]?” Another concern stemming from this is that foreign-language students can coordinate less while writing, because they must also deal with more language issues.

Another popular model comes from Hayes and Flower [44], and is shown in Fig. 2.3. This model shows a framework for what goes on during the composition process. Ref. [44] includes additional models for the structure within the writing sub-processes.
such as generating and organizing, and the area related to generating text is quite similar to the “knowledge telling” process. They have used “think-aloud” protocols to test their model and found their model consistent with real observations. They caution that the models are not “stages”: it is a recursive model in that any part of the writing processes can happen at any time. They also caution that their model is for competent writers, and other writers may fail to use sub-processes within the model such as organizing. According to Flower and Hayes [37], the planning stage is where a person should be able to tackle the cognitive strain of the constraints of writing, so fewer aspects can be addressed during the actual writing.
Epistemic Writing is considered by Bereiter [19] as the most advanced stage of writing development: this is writing in which the knowledge gets modified in the process of being written down. He suggests that writing can be epistemic because of the ability to store, review, and revise information. He believes this occurs when people use their skills of reflective thought and integrate them with writing. This suggests that if a writer is not skilled or instructed in reflective thought, the writing may not help the writer to develop new knowledge.
CHAPTER 3

WRITING IN SCIENCE STUDIES

There are very few studies on writing in physics classes; all research studies that could be found from physics are included here. Select studies from related fields are also included. A common link among these studies is that the authors want to include writing in their classes because they believe it will be beneficial to their students. Another common link is that almost none of the studies cited here provide any data to support their conclusions about the effects of writing to learn. Those that do include data either have not done controlled studies, or have not provided statistical analysis of their data to show whether the results are significant.

“Writing to learn is less about formal uses of writing to display memory and test mastery than it is about informal writing; about language that is forming meaning; about writing that is done regularly in and out of class to help students acquire a personal ownership of ideas conveyed in lectures and textbooks... The writing-to-learn movement is fundamentally about using words to acquire concepts [25].” This goes beyond simply learning to write in different disciplinary modes, but actually using writing to aid learning. The relevance to writing in mathematics and science is that, like writing, these topics are about having a successful way of addressing questions
and working on problems. Science and writing are both more about the process more than the answer.

Often, science course writing such as in lab reports becomes about reporting the correct results, and not about the observations and wondering. “When all the students in the class obtain the same results to an activity, and there is only one scientifically acceptable outcome, the learners quickly realize that they must somehow generate, copy, or paraphrase the knowledge claim that is desired by the teacher.” “The important feature of education becomes saying the right words, not learning how to use one’s own words [25].” There is also a strong tendency to memorize and mimic language, because students believe the language itself holds the knowledge, and they need to participate in using it in order to show their knowledge. A report discussed in Ref. [70] involved having students write 15 explanations in different contexts and found that after doing this students used more scientifically accurate language, and were able to express abstract ideas better. In science, the conclusions should be the same for everyone, therefore the focus of writing should be to understand scientific processes rather than just reporting results. Allowing more non-traditional writing may help students think for themselves, even if they are imprecise at first. Writing should promote the construction of explanations of how and why the world is a certain way, as opposed to language that repeats dogma.

There have been some innovative uses of writing in physics courses. Writing has been used in science courses for education majors [16], and is similarly a part of OSU’s Physics by Inquiry course. Journals are used for reflection on effective teaching strategies, enforcing activity-based, inquiry-mode methodology. The idea is to help future elementary science teachers to understand that “questions are the very
‘stuff’ of science, and that investigating questions is what scientists do ... and even established scientific facts are subject to an ongoing process of questioning, testing, and refinement [16].” In addition to journals, students are asked to write about their understanding of the concepts studied. Asking questions such as “what did you observe, and how do you explain your observations” and encouraging students to state what is unclear allows students to have more awareness of their knowledge. In addition, this can help both the students and instructors recognize and confront misconceptions.

Details of studies on the use of writing in physics classes are presented here. They are organized based on the individual authors and each given a detailed description. This is possible due to the small number of published studies on writing in physics. It is also important to establish the minimal information actually known about the effectiveness of writing in physics classes.

3.1 Teresa Larkin-Hein [46, 55, 54]

Dr. Larkin-Hein is a faculty member in the physics department at American University. Among physicists, she has the most publications on the use of writing in physics. Her publications describe courses in which she has very strongly emphasized the use of writing. Her students keep portfolios of written work that significantly adds to her weekly grading load. One of her more extensive projects involved organizing a mock conference at which students gave formal presentations, which replicated most features of a real conference. Several of her papers start with strong and unsupported statements. One of these is “The educational benefits of adapting a writing approach in the classrooms have been widely documented.” The publications she cites with this
statement (some of which are her own) do not offer any research or data to support this. She claims that writing helps her students learn by making use of the fact that students have different learning styles, but she offers no support for that claim. She often collaborates with people at the University of Pittsburgh, which has a strong commitment to WAC. These projects are supported by the engineering school, the engineering libraries, the writing center, and the academic advising center.

One statement Dr. Larkin-Hein makes is “the use of writing in introductory classes may be an effective vehicle for allowing students to enhance their critical thinking [and] assist students with the identification and confrontation of personal misconceptions.” This statement is one that many of us would want to believe, but it is weak in the absence of any evidence. To support the critical thinking claim, she states “my experience with using writing in introductory physics classes for non-science majors suggests that it can be an effective vehicle for allowing students to develop their critical thinking skills.” She does not offer any explanation of how she came to this decision, and certainly does not provide any data.

To back up the claim about confronting misconceptions, she references her article on “folder activities. [46]” These activities are similar to portfolios, which are a type of formative assessment generally respected as a way to enrich the learning experience. Her activities seem like a good idea. The students write 1-2 pages per week explaining a problem or concept from class, or they can design an exam problem, solve it, and explain the solution. She reads them and gives thorough feedback; it takes her roughly 6-8 hours per week for 60 students. She claims they are a non-intimidating way of learning, because they’re not marked for correctness of physics, helping students to not be afraid to write what they are thinking. However, writing
itself can be intimidating, so it is quite possible the intimidation explanation is correct instead. She claims that students are able to play the role of “learner as teacher” when explaining themselves. This an idea instructors and students have reported to me during interviews and discussions, but once again, the article offers no attempt to study this claim.

Larkin-Hein is able to adjust her lectures based on problems she observes in the writing, which is one of the benefits of formative assessment. She also says that providing feedback to students on these activities allows them to see their mistakes before the exams and adjust their own studying. With her feedback, she “intends for students to think very deeply about [her] comments and then do whatever they need to do to correct the flaws in their thinking.” However, there is no grade given for the folder activities, so it is unclear why students would put much emphasis on the feedback she gives. The only evidence given in the paper for the effectiveness of these activities is her statement that she thinks their work is good quality, and that students have commented that the feedback is helpful. Despite this, at the end of the paper she states “folder activities have proved to be an effective way of helping students make deeper and more personal connections to the physics content under study.”

3.2 John Jewett, Jr. [48]

Dr. Jewett gave writing assignments to his introductory physics students “with the goal of improvement of student writing skills with the learning of physics content.” The writing assignments varied in the audience and type of writing, each encouraging a different skill. One was to write a memo, another was a report, and a third was
a newspaper article. He did do a controlled study. One of his lecture sections did standard problem-based homework and no writing, while the other lecture section did the writing with reduced homework for equity. He found that the writing students did slightly better on writing questions given on the final exam (about 3%), and they did slightly worse (again about 3%) on standard exam problems. He also noted that the two groups have slightly different GPAs, but he does not present any statistics as to whether the two treatment groups are significantly different or not.

He gave a feedback survey to the students, and they had negative attitudes about writing, and especially writing in a physics class. However, students did think that writing is important for doing science. Even more telling, the drop-out rate was two times higher for the lecture section that did the writing assignments, and many students skipped doing the writing assignments as the semester progressed. Dr. Jewett’s overall conclusion was that the writing did not hurt the physics, but probably also did not help their writing, either. These conclusions are dubious in light of the drop-out rate; it seems there was an observed negative impact on some students and thus the writing may have hurt the course.

3.3 Larry Kirkpatrick [50]

Dr. Kirkpatrick’s use of writing in the physics classroom has the most parallels to some of our own projects. He brought a lecturer from the English department into his classroom to help implement the writing assignments and instruct the students in science writing. They gave short essay questions that asked students to analyze a problem or explain a physical phenomenon. They gave a template for writing they called “RAFT,” which defined the Role, Audience, Format, and Task. They
also broke down the writing into key ideas, general ideas, specific cases, and frosting (optional additional insights.) The belief of the English lecturer was that the students had the basic writing skills needed for the tasks, but did not know how to apply them in the subject matter. His philosophy was “if teachers believe a skill is important, it must be explicitly included in the course. That is, not only must the skill be taught, but it must also be tested and the results included in the tabulation of the grade.” He graded the essays both on clarity and style of writing, in addition to the physics content, in order to encourage students to work on the clarity of meaning, logical constructions, and readability of their work.

His reason for implementing writing in the classrooms echoes my own reasons for my choice of thesis topic. He has observed that even ‘A’ students often show a serious lack of understanding when they discuss relationships between concepts. Dr. Kirkpatrick never did any study of the effectiveness of the writing or the impact of having an English expert in the classroom. He states in his publication that “no conclusive study has yet been done to measure the effects of writing on cognitive development or on mastery of course material.” He continued to use writing assignments in his classes up until his retirement (though not with the English lecturer), and feels they are beneficial to students, but he never collected any data from his classes.

He provides several remarks based on his classroom observations. His students were initially resistant to writing in their physics class. He felt the quality of their writing improved significantly with the inclusion of the English lecturer. The students answers were more extensive, there were fewer disjoined answers and more logical connections between ideas, and it was easier to see whether the physics was correct or not. An interesting observation he made was that the best essays were about the
same quality as the ones he saw without the presence of the English lecturer, but there were a lot fewer essays of really poor quality. He never attempted any quantitative measurement of the writing quality, however. The students said the hardest part of writing the essays was getting the physics right; that it was not possible to “bluff” the answer. They also said that once they understood the physics they found it easier to get their ideas on paper. Student attitudes toward writing improved by the end of the course, with 91% of students stating the writing helped them think about and understand the physics content.

### 3.4 William Mullin [66]

Dr. Mullin’s publication describes a physics course developed at the University of Massachusetts as part of a WAC mandate. The course was a junior-level writing course within the physics department. Over the course of the semester, students wrote six papers ranging from three to four pages in length. The papers required library research and multiple drafts. Each had a different focus, for example, explaining to a freshman how an airplane flies, or writing a newspaper article on a science talk. He feels that writing encourages heuristic thinking and learning, and can generate interest in new topics. An interesting note is that in his paper he explicitly mentions that he has a set of beliefs about the benefits of writing that he holds without any firm proof. These include that writing helps students learn the content and style of technical writing, they gain confidence in their writing, they learn to organize and revise their writing, that they gain awareness that physics is highly dependent on intuitive arguments, that they learn about different styles of writing, and that they learn about some aspects of contemporary physics research.
Dr. Mullin argues strongly, without providing evidence, for the use of intuitive (verbal) arguments along with mathematical formulations and proofs for teaching both old and new physics [65]. He notes that Richard Feynman was known for the use of intuitive arguments in his legendary instruction. He also reports frequent complaints by his colleagues that students know the formulas but don’t understand the physics behind them. He stresses the fact that written and oral formulations are part of both physics tradition and modern research, but that this is often hidden from the view of students. For example, notes and discussions lead to publications, but the articles themselves are formal and rigorous.

3.5 Scott Franklin and Lisa Hermsen[38]

Our research collaborators at the Rochester Institute of Technology (RIT) have completed pilot studies in Writing to Learn that parallel some that are presented in this thesis. They have used essay writing in introductory physics classes at RIT with English instruction. They develop two methods for analyzing student writing. The first is called “PrimaryTrait Analysis,” which is a procedure to categorize the assignment-specific observable traits in written documents. They graph occurrences of sentences coded as motivation, procedure, observation, inference, or fact. They find more sophisticated writers have more inference and speculative sentences, and less sophisticated writers rely on “procedure-observation” cycles, where students write “we did this and saw this.” Our use of similar categories in the Summer quarter 2005 study does not define speculative sentences the same way; here they are usually considered good components of the essays. The second method they develop is called “Revision-based Assessment,” where they use Microsoft Word’s “Track Changes” to
assess the revisions students make in writing activities qualitatively. Their writing activities consist of first drafts with a 250 word limit, and second drafts that must be no more than 125 words. This forces students to do major revisions. This activity inspired those used in the Winter and Spring quarters, 2006, presented in this thesis.

3.6 Related Physics Research

Kathy Harper and Sandra Doty at the Ohio State University looked for correlations in how students write their exam “equation sheets” and how well they did on exams. They found no correlation between people who organized and connected their ideas on the sheet and students who did well on the exams [42]. Another study focused on how physicists read journal articles [18]. Beriter and Scardamalia [20] explain that novice writers use a “knowledge telling strategy” in which they produce text by writing their first idea, then another one, then so on until they run out of ideas. The analogy between reading and writing is that both expert readers and writers first plan ahead, then read, then revise. In the case of reading journal articles, the planning is reading the title and abstract and thinking about what may be relevant in the article, and revising is going back and looking at more details including some references. Novice readers and writers tend not to realize that reading and writing are multi-step processes.

An interesting and related study reported by Rymer [73] reporting on how scientists write journal articles. She studied nine expert biochemists by using interviews and composing-aloud protocol. She concludes they use a range of writing approaches including very recursive models focusing on revisions. Revisions and polishing of drafts is common to all of the experts. Most of the scientists claimed to discover
new ideas about their results and what science means while composing their jour-
nal papers. Some of the scientists reported that having a highly planned research
and composing method was inhibiting, and in response the author suggests students
should be taught a variety of composing models.

3.7 Writing in Mathematics

More studies have been done on the use of writing in mathematics. These studies
typically involve more collaboration with experts from English and WAC programs,
but again are lacking in quantitative data to support their claims. Writing in physics
is related to writing in mathematics because problem solving is key to both fields, and
writing is uncommon in both at the introductory level. Implementations of writing
may be similar, and have many of the same difficulties and benefits. One important
similarity between writing in mathematics and physics is the needed precision with
word use and definitions. In mathematics, it is important that students be adept with
language, and that they are able to express things in multiple ways. For example,
knowing that “twelve divided by three is four” is the same as “three goes into twelve
four times [56].” Subtleties exist in mathematics language use; for example, if a
problem asks to show something is true for “an odd integer” it means students need
to show it for all odd integers, not some specific single odd integer.

According to Ref. [70], transactional writing (writing meant to be read by an
audience, for example, a teacher) is commonly used in mathematics classrooms. This
includes summaries, reports, projects, essays, and notes aimed at helping students ex-
press their mathematical understanding. Sometimes expressive writing is also used,
for example, freewriting, to generate thoughts about concepts and processes, and
journal writing, to record problems, questions, and feelings. “Journals are deliberate exercises in expanding the awareness of what is happening, personally and academically.” One activity used in mathematics classes is having students construct their own word problems, or rewrite problems into word problems. When students do this, they often apply the mathematics to situations they are familiar with. The claim is this helps them see the relevance of mathematics in their daily lives. It is also claimed this helps students practice being clear, specific, and detailed, in order to write a well-defined problem.

“Some professors resist writing assignments [in mathematics] both because of the anticipated time needed to read and grade the work and because they feel untrained as a composition teacher. Students are also reticent [reluctant] about writing, since most were never required to write in previous mathematics courses [70].” It is proposed that in order for writing in mathematics to be effective, professors should participate in the writing assignments, students should be given help and encouragement and regular feedback, and the teacher needs to establish an atmosphere of trust. We are warned in [22] that “writing by itself is not a cure-all. Writing exercises are no more beneficial than tests unless they allow the student to explore, think, test, take risks, and learn through the process.”

One specific writing activity from a mathematics course as cited in Ref. [22] required students to write about the purpose of proofs, what is is to do a formal proof, and how they think mathematicians use them. They also had to outline, in prose, an induction proof for a problem and then write the proof in two ways: using just mathematics symbols, and just English phrases. It was felt, without evidence for support, that this assignment helped students reflect on what they were doing
and why they were doing it. Students reported having a deeper understanding and better retention of the concepts after writing. In addition, such assignments helped the instructor see where student thinking went wrong, and help redirect it. “Writing allows students to engage deeply in the content of a course in a way that tests and quizzes simply cannot, while providing a new forum in which to engage in discussion about their mathematical thinking, perception of the course, and interest in the subject.”

The work of Tobias is familiar to the Physics Education Research community. She showed in Ref. [79] that students who do not traditionally succeed well in mathematics and science are held back for reasons other than intellectual capability, for example, learning styles, attitudes, and confidence levels. She argues that problem solving is very similar to writing; it causes students to “come to grips with their incomplete understanding in an active and self-stimulating fashion.” However, she argues for the benefits of using writing especially in terms of reaching students who would otherwise be turned off to mathematics and science. Students feel spending several hours to produce a draft of writing is productive, while spending several hours trying to solve a problem is not a successful use of time. They also do not see that both tasks require reviewing and revising, so they are not necessarily expected to solve a problem in the first step. The task in the “noncomfort zone” is given impossibly high expectations by the students; they feel they are not successful if they do not achieve the task immediately. Students who are allowed to do informal mathematics writing, paying attention to their past experiences, thoughts and errors, become less afraid of mathematics. She points out that textbooks in these subjects use a very terse and precise language, which can be difficult for students. She cautions that writing not
be viewed as a substitute for traditional work in these fields, but rather as a scaffold between language the students are comfortable with and the useful symbolic code they will need to eventually understand.

Kenyon [49] notes that the mathematics community has been reluctant to participate in WAC programs. He finds this ironic, since he feels strongly that the stages of writing are very similar to the stages of problem solving; essentially he states writing is problem solving. The process of problem solving can be summarized by four steps: being aware of the situation, recognizing it requires action, wanting or needing to act upon it, and coming to a resolution that might not be immediately obvious to those acting on it. “Verbalization is a technique that requires the learner to report on the thinking process during problem solving. It is based on the assumption that the problem-solving process will be understood by the learner only when he or she can explain it to others [49].” In group problem solving, verbalization is used to help the problem solver and others observe and evaluate the work. Ref. [49] claims it is also helpful to write down the thoughts and procedures involved in problem solving when working alone, as the steps are visible for evaluation, review, and self-monitoring. This also helps students clarify their thoughts and identify the concepts more clearly; the writing process becomes an integral part of the thought process.

Ref. [49] states that in writing, the first stage is prewriting, or planning. This is a place where the student attempts to understand the conditions of the problem and how to achieve the goal. In this state, the memory is searched for data, strategies, and techniques, and ideas for possible solutions are generated and organized. In the composing stage, ideas are translated into text. In the revising and reviewing stage, where ideas can be moved around and more can be added if needed for clarity.
Kenyon has tried mathematics writing where students have to write out solutions with no references to the specific numbers or variables. He has noted observing that students are generally able to write shorter, more precise responses as they practice such writing activities. Once students become comfortable with writing, they are given more advanced assignments where they must explain why things work, or why a certain solution method is chosen.

### 3.8 OSU’s Writing in Physics Studies

The bottom line from the literature review is that there are no data to show whether or how writing helps learning. It is necessary to apply a scientific study to determine whether there is merit to the common beliefs and claims about writing. In the words of Carl Wieman, “science teachers tend to be very unscientific about their teaching. Teachers should not decide what’s right and wrong by tradition, or superstition, or anecdotes... that 2 out of 100 students told you they liked it. We know how to evaluate these things better [3].”

An ongoing collaboration between members of both the Physics and English departments at the Ohio State University and Rochester Institute of Technology has been working toward understanding writing to learn in the context of physics. Our main long-term goal is to do controlled studies to quantify whether writing helps learning. However, there are multiple questions within this topic, all of which need to be understood before larger issues can be tackled.

The broader question “Can/do students gain conceptual understanding in their discipline (in specific, physics) by writing?” can be broken up into the following questions:
1. Does student writing in a physics course improve with practice?
   - Does explicit writing instruction improve students’ writing?
   - Does the displayed physics content improve with practice?

2. Are students reflective when they write?
   - What can we understand about the nature of student writing by looking at samples of their writing?
   - What can we learn about student processes of making revisions?
   - Does forcing students to make revision force reflection?

3. Can we correlate the quality of the English usage with the physics content?

   In order to start addressing these questions, it is also necessary to develop a scientific methodology that can be used to study these issues. In addition, the available tools for investigating writing are limited. Our aim is also to develop a tool for tracking student writing, and use it to characterize student writing behaviors. These aspects are addressed in more detail in Chapters 6 and 7. Briefly, we investigate student writing and revision behaviors such as:

   1. The percentage and number of revisions, types, and sizes of revisions.

   2. What content students cut or rephrase from their first drafts, and how they approach the word limit.

   3. The quality of the text on both drafts and how this correlations with revision behavior.
4. Where they pause, and how often they pause.

5. How long it takes the students to write the entire text.

The thesis is comprised of four major studies. The first two use traditional methods to try to address questions from the first list above on whether writing impacts learning directly. Although interesting results are obtained from these studies, it was found the research methods were insufficient for determining whether there were actually any effects of the writing assignments on student writing or on their learning. The latter two studies use an existing and a new tracking program to refine the research. This allows us to continue addressing questions from the first list above, while developing new research methods for the general field of writing to learn using the second list of questions.
CHAPTER 4

SUMMER 2005: AN ENGLISH INSTRUCTOR IN THE PHYSICS CLASSROOM

The focus of this chapter is on a project that took place during Summer quarter, 2005, at the Ohio State University (OSU). The main issues addressed involve the use of these data to address the question: can we correlate quality of students writing English with quality of their physics explanations (and types of mistakes)? In specific we consider the following research questions:

1. Is there a correlation between instructors’ comments on English usage and instructors’ comments on physics made in grading?
2. What types of comment occur more frequently?
3. Do explanations of more difficult content engender more comments?
4. Do any of these factors change with instruction and/or practice?

4.1 Implementation Details

This project was implemented in the electricity and magnetism segment of OSU’s introductory calculus-based physics sequence. There were three laboratory (henceforth, lab) sections of this class, one with six students, one with eleven students,
and one with seventeen students. The design of our experiment included having two
treatment groups: one with explicit writing instruction (WI) and the other with no
writing instruction (NI). These two treatment groups did a shortened version of the
standard lab plus writing activities. The largest section was the WI group, the mid-
sized section was the NI group, and the smallest section was the control: writing no in
class assignments, but instead doing the entire standard lab and standard end-of-lab
quiz. These size groupings were not intentional; it was a matter of scheduling, and
unfortunately the control was not reliable because after a few students dropped the
population was too small for proper statistical analysis. These details are summarized
in Table 4.1.

Although there were fewer students in the NI group, and therefore more time
could be spent helping each student, that section was in the period prior to the WI
section. Therefore the instructors were more aware of the problems students had and
how to quickly address concerns in the WI section. Therefore the effect of having a
larger class to address, yet being more prepared probably cancel each other out.

<table>
<thead>
<tr>
<th>Group</th>
<th>Lab</th>
<th>Writing</th>
<th>Extra Instruction</th>
</tr>
</thead>
<tbody>
<tr>
<td>NI: No Instruction</td>
<td>Regular, 1 h</td>
<td>1/2 h in 2\textsuperscript{nd} part of lab</td>
<td>5 minutes general instruction plus physics help during writing</td>
</tr>
<tr>
<td>WI: With Instruction</td>
<td>Regular, 1 h</td>
<td>1/2 h in 2\textsuperscript{nd} part of lab</td>
<td>As above, plus 15 minutes English instruction</td>
</tr>
<tr>
<td>Control</td>
<td>Regular, 1</td>
<td>none</td>
<td>15 minute end-of-lab quiz</td>
</tr>
<tr>
<td></td>
<td>1/2 h</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Division of time on lab activities; Summer 2005
4.1.1 Writing Activities

The writing activity consisted of filling in a missing paragraph of a pre-written essay, which was four to five paragraphs in length. The missing paragraph was replaced by an explicit prompt that included cues indicating what information was missing. The topics included why a car is safe during lightning, how electrostatic precipitators work, how holiday lights are wired, and how solar particles are trapped in Earth’s magnetic field. These essays are in Appendix A. The missing paragraphs required explanations of some aspect of the phenomenon based on content from that day’s lab. The type of writing required of the students was argumentation: constructing an argument about the way certain physical processes operate in the world. The argument should show, by careful reasoning, that there is a valid solution for the proposed problem, not just stating facts.

The students had six labs during the quarter. In the first lab, we had three different paragraphs that could be filled in by students. Because of the difficulty we experienced in assisting students with three simultaneous paragraphs, we chose to give only one missing paragraph to all students in every following lab. Because of this difference, the first essay was omitted from this study.

The essay topic for the fifth lab was particularly difficult for the students. I gave physics help during the writing activities, but I was not able to attend this lab, which may have contributed to some of the difficulty. Therefore instead of introducing a new topic for the sixth lab, the students were asked to make a second attempt at writing the topic given in the previous week (week five). This was the only time students had two chances to try an essay; consequently, the sixth lab essay rewrite was omitted. Therefore, only four weeks of the six are discussed in this study.
4.1.2 English Writing Instruction

The writing instruction consisted of lesson plans addressing both higher- and lower-order concerns, beginning with higher-order concerns and moving to progressively lower-order throughout the quarter. Higher-order concerns are universal issues in the paper such as topic, organization, or logical flow, while lower-order concerns comprise sentence or word level issues. In the first lab, the students took the Conceptual Survey of Electricity and Magnetism (CSEM) [59], so did not have time for a writing lesson.

In the second lab, we gave general tips for completing the assignment: (1) make sure the paragraph has all of the information the prompt asks for, (2) fit the paragraph so it flows with the rest of the essay, and (3) back up any generalization with specific examples. In the third lab, the English instruction focused on the organization needed to make an argument. The students were explicitly told that the purpose of the paragraph is to construct a physics argument that requires factual support. The students completed an activity in which they outlined the information needed to respond to the prompt from the previous essay, then compared their outlines to the essays they wrote. They then discussed what information was missing, and how the order of information could have been improved to strengthen their argument.

In the fourth lab, the English instruction focused on sentence-level problems. The students were given a handout with sample sentences from the preceding week’s essays. These samples illustrated problems with transitions, using equations, using concrete language, and mixing observations and explanations. Each of these problem types was addressed with a discussion, then the students rewrote the sentences to
correct them. In the fifth lab, students were given sample sentences from the preceding week to illustrate problems with specific word choices. Proper use of verbs, adverbs, and prepositions common in physics terminology were discussed, as well as the importance of using neutral words, and knowing when to use definite or indefinite language. The students were then given time to correct the sample sentences. Except for the case in which we were giving good writing examples, we pulled all sample sentences from the control group so that no students were uncomfortable seeing their work used as an example of bad writing. In the sixth lab, students again took the CSEM and had no explicit English instruction. The writing activities and lesson plans are given in full in Appendix A.

4.1.3 Grading

Each week, students’ essays were copied: one copy was given to Catherine Gubernatis and graded for the quality of the writing; students were not marked off for grammar or spelling mistakes. It is important to note that this grade was not biased by the physics content, since Gubernatis does not have any background knowledge of physics. The other copy was given to Dedra Demaree and graded for physics content. Each grader made comments on the papers as they were graded. The graded and commented papers were then copied, recombined, and returned to the student. The final grade students received for each essay was the average of the English and physics grades.

4.2 Data Obtained

In order to analyze the essays, the English and physics comments were coded by Jessica Hanzlik. Although the graders discussed the importance of students producing
a strong argument in their writing, each grader came up with her own grading rubrics separately. Jessica found that the English and physics comments were often very similar. Most statements could be grouped into five or six categories, which were then later re-grouped into three main English comment types and three main physics comment types. The English comments included external and internal language issues with comments such as “need transition”, “ideas seem disconnected”, and “awkward wording/confusing/unclear”. Other comments focused on content issues, such as “need more details/explanation/be more specific”, and “put this in context”. Both the English and physics essays included positive comments, which mainly consisted of “good”. Physics comments centered around clarity issues, such as “physics not clear”, and “no flow to explanation”. Other comments focused on the correctness of the physics, with “physics not completely explained/wrong logic to explanation”, and “physics is wrong”. A sample of both graded copies of one student’s essay are shown in Fig. 4.1 and Fig. 4.2.

The location of each physics and English comment was coded for each essay from lab weeks two through five. The basic content of each sentence was also coded, as well as the sentence type: motivation, observation, speculation, inference or fact. In addition to the coding and the essay grades, the students also took the CSEM diagnostic test, and the Epistemological Beliefs Assessment for Physical Science (EBAPS) [4]. The lecture instructor also put a question on the final exam requiring a written explanation of the motion of a charged particle in a magnetic field. This question was graded by Dedra for physics content for the sake of their final exam grade, and later also graded for comparison by Catherine for English.
Figure 4.1: A sample of student writing with grading comments based on physics.
Figure 4.2: A sample of student writing with grading comments based on English.
4.3 Difficulties We Observed

The main problems observed in the student writing included clarity, organization, and language. As regards to clarity, Catherine noted that she could not always gain a understanding of the ideas from reading their work. She also noted that students had problems organizing their papers: they either left out information, put information in the wrong order, or had a lot of confusing margin notes. She also noted that the language use in the writing was not always very scientific. Students also had problems showing the relationships between ideas, using equations and diagrams, and using physics terminology. It seemed students were often not thinking about these assignments as a process of constructing arguments, but instead thinking of them as describing facts.

The main problem we encountered was resistance of the students to writing, and resistance to having a non-physicist aid in instructing their course. Students were told of the fact that the Ohio State University’s Engineering departments want more writing in the core courses due to alumni feedback that their writing skills were insufficient for future jobs. (Most students in this course are engineering majors.) Students repeatedly complained that since they were not planning on going on to graduate school they did not need to practice writing. We suggest that if writing were used more often in introductory science classes students would be less resistant to viewing it as part of the pedagogy.

We also found that, although students were struggling with how to express their ideas, they seldom took advantage of Catherine’s presence in the room for help while writing. In contrast, Dedra was often asked explicitly about the physics concepts while students were writing. The students seemed to not believe that Catherine could tell if
their writing was making sense, despite being told weekly that the English grades were always quite similar to the physics grades because she could analyze the strength of the argument. If collaborators from the English Department can be used for designing lesson plans, but the lessons can be taught directly by the physics instructor in the classroom, possibly the students would be more open to the English feedback.

4.4 Results

Our results consist of both qualitative and quantitative data, based on essay grades, exam grades, surveys, observations, and written comments made by the essay graders.

4.4.1 Observed Changes in Writing

Through the quarter, students writing became easier to read, as well as clearer. Students also showed improvement in integrating diagrams and equations into their writing. However, only a few students improved greatly, with most obtaining a grade on their final essay within two points of the grade they had gotten on their first essay. This is not a great measure of improvement, however, because we also observed that the quality of writing was heavily dependent on how well the students understood the content. When students were confused, their writing was not as clear, understandable, or well organized. An example of this comes from one student who obtained a 12/15 on the first essay, then obtained several grades of 15/15. However, in the fifth lab (in which all students had difficulty), his grade was again 12/15.
4.4.2 Quantitative Data

One question we consider is whether the explicit writing instruction had an impact on the physics quality in the student writing. The essay grades given solely on physics content are graphed in Fig. 4.3. This shows that although the NI group started with higher physics grades, by the end of the quarter the WI group had higher grades. A Mann-Whitney U test was run to compare the essay grades between the WI and NI groups. In weeks one through three, the 2-tailed significance is no lower than $\rho = 0.4$. However, in week four $\rho = 0.1$, and by week five significance is reached with $\rho = 0.05$.

Each week, the physics and English grades correlated well, with a correlation coefficient ranging between 0.51 and 0.79. In addition, 68% of all essays had physics and English grades within 1 point of each other. Due to this, similar results from those shown in Fig. 4.3 can be shown with the total essay grades (with $\rho = 0.4$ in week 5). Two students stand out because of inconsistent English and physics grades. One of these students was a graduate student, and was very experienced as a writer. Although she sometimes did well with the physics content, in the weeks she struggled with the physics, she managed to write well enough to convince the English grader that her response was sensible, thus obtaining a high English grade and a low physics grade. An opposite case was a foreign student who struggled with written and spoken English, but generally did very well in physics.

4.4.3 CSEM and EBAPS Results

During both the first and last lab period, five students from the control group, nine from the NI group and 16 from the WI group completed the CSEM. Three students in the control group were outliers; they obtained higher normalized gain than any
Figure 4.3: Weekly essay grades based on physics.
student from the NI or WI groups. On the basis of the foregoing considerations, it is my belief this is a case of chance, and not any effect of this project. It has been observed at OSU that sections meeting earlier in the day often have more honors students due to the fact that they are allowed to register earlier. This is supported by the fact that the control group had significantly higher pretest scores on the CSEM ($\rho = 0.029$ between NI and control, and $\rho = 0.015$ between WI and control). These same data support comparing the NI and WI groups directly to each other, since their CSEM pre-test scores were not significantly different ($\rho = 0.846$). A Mann-Whitney U test shows no significant difference between any two groups on the normalized CSEM gain, though the difference between the control and the NI and WI groups are approaching significance (with $\rho = 0.082$ and $\rho = 0.052$, respectively). The difference between the WI and NI group is not significant, with $\rho = 0.552$.

The EBAPS measures shifts in students’ epistemological beliefs through the quarter. Four students from the control group, six from the NI group, and nine from the WI group took this test during their recitation section in both the first and last week of class. The test measures shifts along five axis: structure of knowledge, nature of learning, real-life applicability, evolving knowledge, and source of ability to learn. Although EBAPS does not explicitly address writing, it does involve explanations in several questions. It is possible that students would have had a more broad perspective on the nature of learning science if they had been exposed to new classroom activities, and writing arguments would (we hope) help students evolve past the idea that science is solely about memorizing facts. The EBAPS has not yet been used widely, so there is little data in the literature to compare our results to.
For the WI and NI groups we observed positive shifts across the structure of knowledge axis, and negative shifts across the evolving knowledge axis. Negative shifts for real-life applicability were observed for our control and WI groups, and are commonly seen in other surveys about student beliefs [5, 69]. Mixed results were seen across the other axes. No shifts across any axis (nor sum of all shifts) were found to be significantly different for any of the three groups. The most significant shifts between the WI and NI groups are across the nature of learning axis, where the WI group had a greater positive shift than the NI group (with $\rho = 0.113$). The control group had a significantly more negative shift across the structure of knowledge axis as compared to the NI group ($\rho = 0.019$) and a significantly more positive shift compared to the NI group across the nature of learning axis ($\rho = 0.038$). These results seem inconsistent and are hard to interpret given the small numbers of students, or more EBAPS data for comparison.

4.4.4 Data from Essay Content and Grading

The first data from the essay coding addresses the question: is there a correlation between the types of English comments and the types of physics comments made in grading? For every sentence that received either an English or physics comment during grading, 37% of the time there was both a comment from the English grader and the physics grader. For each group, and each week, the percentages of papers on which both graders commented ranged from 29% to 47%; this agreement was fairly consistent across the two groups and through the quarter.

Both graders made comments based on language and clarity, content and correctness, or made positive comments. Grouping comments accordingly, we found that,
averaged over all weeks, 63% of the time when both graders commented at the same location, the comments made were of the same type. This percentage ranged from 21% to 89% for the two groups for any specific week. It is interesting that except for week 2, the WI group had a higher agreement between the English and physics comments than the NI group. The average comment agreement in the NI group was 55%, while the average in the WI group was 72%. This is graphed by week in Fig. 4.4. It is possible that the writing instruction impacted this agreement. Perhaps students in the WI group were more cognizant of the writing issues, causing this difference. It is also possible that having both graders present while the students wrote somehow affected the way we commented the papers.

The second analysis addresses the question: what types of comment are more frequent? We looked at the coded comments to see what types of comments occurred in each group through the quarter. The percentage of English comments that were positive in week two for both groups was 9.5%. This gradually increased through the weeks until week four with 32%. This is consistent both with the improvements in the student writing, and the fact that the graders wanted to make sure we gave suggestions to correct student work early on. This probably artificially inflated the increase in positive comments observed from lab to lab during the quarter.

An interesting observation can be made in week five, when the WI group has a percentage of positive comments almost three times higher than the NI group. These percentages for each week and for each group are shown in Fig. 4.5. Although this essay was difficult for both groups, as is seen by the lower grades in week five and the reduced number of positive comments, the difference between the two groups is striking. This is the week when the difference in essay grades between the two groups
Figure 4.4: Percentage of times physics and English comments at same location were of the same type.
reached statistically significance. The striking difference in the number of positive comments for each group supports that result.

The percentage of language-based English comments decreased steadily throughout the quarter, with a steeper decrease for the WI group, though the decrease is not as striking as the increase in positive comments. A similar trend is seen in the content-based English comments, with the exception of week five, for which students had a lot of difficulty with the content, and the percentage of content-based comments almost doubled.

Similar analysis can be done with the physics comments, though no clear trends appear from this data. In contrast to the case with the English comments, most weeks the WI group had a higher percentage of physics comments based on correctness than the NI group. Again, there were more positive comments as the quarter progressed,
but there were more positive physics comments in week four than in week five. This may simply be due to the fact the students struggled with the content of the week 5 essays. The percentage of positive comments for each group were very similar, as well. Another interesting note is there was a large spike in clarity-based physics comments during week three, with percentages almost three times higher than any other week. This may simply be because the writing instruction for that week was based on organizing the argument structure, so the grading may have been biased toward reinforcing those comments.

The third analysis addresses the question: are there more comments when the content is harder? In order to address this question, one of the authors looked at the sentence content categories that were coded, and rated them as easy, medium, or hard. An easy comment would be something like the statement of a physical law when the prompt told students to use that law to show something. One hard sentence content-type was chosen for each essay, and was based mainly on instructor memory of what students struggled with understanding when writing the essays. Considering only the non-positive comments, we calculated the percentage of each sentence difficulty type that was commented.

The prediction was that harder sentences would have a higher percentage of comments than the easier sentences. However, this prediction was not observed for either the NI or WI group. In the week five essay with which students struggled, it is interesting to note that none of the NI students included in their essays the content that was coded as hard. However for the NI group the percentage of sentences commented in week five was greater than the previous weeks. Since they struggled with this content that indicates overall more negative comments were made when they struggled
with an essay, but not necessarily within any given essay. For both the WI and NI data, fewer comments were made in week four than in other weeks; students did well on that particular topic. No clear trends were observed in the WI data, though a higher percentage of hard sentences were commented in week two, but not in any other week.

Another coding that was done was whether the sentences students wrote were of one of the following types: motivation, observation, speculation, inference, or fact. A fact statement would be something such as the statement of a physical law. The difference between an inference and a speculation was determined by whether there was any basis of support for the statement. Between the WI and NI group the results were fairly similar. A graph in Fig. 4.6 shows the percentage of sentences of each type for the WI group. It can be seen that observation and inference sentences were most common, and the percentage of each sentence type was fairly steady throughout the quarter.

The most notable difference between the NI and WI groups was the amount of speculation, shown in Fig. 4.7. The NI students had considerably more speculative sentences than the WI group each week. Speculative sentences are not supported by argument and therefore should not appear in good science writing. It is interesting that both groups had more speculative sentences in week five when they struggled with content.

4.4.5 Final Exam Data

The final piece of data concerns the written explanation question on the final exam. This question required the physics used in the essay from week five (this
Figure 4.6: Percentage of sentences of each type for the WI group.
Figure 4.7: Percentage of speculative sentences for both groups.
was not planned). The WI and NI students had seen this topic twice in lab, while the control group students did not see it in lab. It was also covered in detail in lecture. The WI students had much higher physics grades on this problem than the NI students (the average for WI = 7.1/15, while the average for NI was 4.9). However, the English grades were nearly identical (10.6 for WI and 10.2 for NI). It should be noted that the physics grade was based on points for the inclusion of various details of the problem, while the English grade (which did not count toward their final exam grade) was graded as we had graded the writing in lab where the percentage score roughly corresponded to the grade we felt the essay deserved (for example a 12/15 would be a B- paper).

During the final exam, it seemed that the students did not focus at all on English, with almost no responses to the final exam question appearing in full sentence form, let alone essay form. It is not surprising that students did not transfer the idea of writing from lab to the final exam. The lab and lecture are fairly separate, and students seemed reluctant to write even during lab. The control group, which had stronger students, had slightly worse writing, and physics quality in between the NI and WI groups (physics grade averaged 6.3, and English grade averaged 9.8). There was no significant difference between physics grades for the NI and WI students on this problem, but there was a significant difference when the English and physics grades were averaged (the Mann-Whitney test found \( \rho = 0.11 \) for the physics grades and \( \rho = 0.027 \) for the averaged grade). Therefore, overall the writing instruction had a measurable positive impact on the exam problem. There was no significant difference between the two groups overall on the final exam (\( \rho = 0.89 \)).
4.4.6 Other Possible Analysis

Due to the small sample sizes, finding strong trends is difficult from these data. If additional studies were completed with more data, we suggest the following additional types of analysis. Analysis could also be done to see if sentence types or sentence content correlates with the overall essay grade. For example, are students who score better more likely to include the hard content and fewer likely to use speculative sentences? Also, it is possible to graph the types of sentences found in the essays: for example, do good students start with motivations and observations, then use inferences? We could also see if students with better grades received more positive comments and less comments of other types.

4.5 Conclusions

There are few strong conclusions that can be drawn from this data. There is little evidence that writing helped the students learn, though there is some evidence that writing instruction had a positive impact despite the negative student attitudes. The strongest evidence is that the quality of physics by the end of the quarter was significantly better for the WI group than the NI group. In addition, the WI students did better than the NI students on the written final exam problem. This did not transfer over to the CSEM, as may have been expected, since transfer of knowledge gained from one type of exercise to a very different type of problem is not easy [13, 53]. It is difficult to conclude whether students actually gained more physics knowledge, or if the WI students gained a better ability to explain their knowledge; hence producing better quality physics. The latter is supported by our observations, and is worthwhile even without the former, as improved writing skills are needed for the workplace.
We also provide evidence for the correlation between English quality and the physics content. This helps establish a connection between the ability to express content knowledge and writing, which provides (modest) support for the idea that writing is pedagogically beneficial. The data support the idea that students need explicit instruction in order to take full advantage of writing activities. Since the impact of writing instruction was not measurable until students had completed several assignments, we suspect that given more practice writing the results from our project would be stronger.

The project shows that an intervention in a student’s usage of English does have some correlation with their ability to explain physics ideas. The data support the idea that students should have explicit instruction in order to make full advantage of writing activities. However, since the English instruction could be conducted by a physicist, and students were reluctant to take instruction from a non-physicist in the physics class, we suggest that the writing instruction be led by the physics instructor.
CHAPTER 5

AUTUMN 2005: CAN WE TEACH ONE CONCEPT THROUGH WRITING

5.1 Introduction

In Summer, 2005, we saw evidence that having writing instruction in the classroom had a positive impact on the quality of physics the students produced during the writing activities. There is scant evidence that they actually learned the physics better; though students with writing instruction did better on a final exam problem of similar content there was no other measurable difference in content knowledge. In an effort to see whether we could measure a difference in conceptual knowledge gained, we decided to conduct a similar study during Autumn, 2005, in which we narrowed our focus to one concept area. In this study, we addressed the following questions:

1. Can we teach physics conceptual knowledge through writing activities?

2. Will there be a difference between conceptual understanding for those who have writing instruction compared to those who do not?
5.2 Implementation Details

This project was implemented in the electricity and magnetism segment of OSU’s introductory calculus-based physics sequence during Autumn quarter, 2005. There were sixteen lab sections for this class, taught by four teaching assistants, including the principal investigator of this study. Demaree’s lab section was omitted from the study, to guard against the possibility that she might bias the discussions with the students based on her research goals. Of the remaining fifteen sections, five sections did the traditional lab activities (control group), five had additional writing activities with English-specific instruction by an English PhD student (WI group), and the remaining five did the writing with no additional instruction (NI group). End of lab quizzes were given in each lab to all students, somewhat reducing the amount of time writing students could spend on the traditional portion of the lab. There were between sixty-five and seventy-eight students in each group who consented to allowing their course-work to be used for research purposes, and a total of 220 students who participated in this study. These details are summarized in Table 5.1. Considerable effort was put into choosing which sections to put in which treatment group. The sections were divided by time of day, by the three teaching assistants, and what day in the week they met. Each treatment group comprises an even mix of these three parameters.

There were a few major differences between the focus of the autumn project and our previous summer project. Most notable is our effort to focus on a specific content area. There were two existing labs on circuits: one focused on bulbs and batteries, the other focused on the effects of capacitors. We added a third circuits lab on Kirchhoff’s laws, and implemented the writing activities only during these three weeks. Due to
the heavy grading load (over 150 students were writing essays each of those weeks), we could not grade all of the essays and return them the following week. Instead of providing individual feedback, we looked at a sample of essays to choose feedback to give to the entire class, and from which to pull examples for the writing instruction.

In addition to doing the writing activities, students again completed the CSEM [59] during the first and last labs. They also completed the Determining and Interpreting Resistive Electric Circuits Concepts Test (DIRECT) [30] in the lab following the final circuits lab. All lab quizzes, recitation quizzes, and exams were collected as well, in order to search for any differences in content knowledge between the groups.

### 5.2.1 Writing Activities

The writing activities differed from Autumn quarter. In order to focus on specific content knowledge, we had students write on topics that were more directly related to the lab activities. The battery and bulb lab requires students to make several predictions about the relative brightness of bulbs in the circuit. Instead of doing several different circuit puzzles of this type, the students wrote an explanation for how to determine the relative brightness of bulbs within a circuit. Similarly with the
addition of capacitors, students had to explain how to determine how the brightness of the bulbs change with time. Students work on the labs in groups of two to five students. For each of these labs the students wrote individual essays, though they were encouraged to discuss them with their group members. For the Kirchhoff’s laws lab, students had to write an explicit set of instructions for using the law, including what the laws mean (for example, the junction rule is current conservation). Students wrote up these instructions as a group, sometimes with everyone contributing and one person writing, but other times they split up different aspects for each person in the group to write about. Students then exchanged papers with other groups, and had to apply the rules to solve for the currents in a circuit we provided. Students were given bonus points based on how well their rules worked. Students were allowed to write lists instead of essays for these assignments, since some students found this less intimidating and each of the assignments were of the instruction rhetorical type.

5.2.2 English Writing Instruction

The writing instruction borrowed from ideas we thought were successful during the summer 2005 project. For each day of writing instruction, the English instructor prepared a discussion including tips for how to address specific writing issues. In addition, there was a short activity for the students to work on associated with the discussion. For the first circuits lab, the writing instruction focused on how to move between generalities and specifics, and vice versa. This is because in order to write a general explanation for how to determine something, it is important to back up the ideas and give specific examples, so we wanted students to reflect on this before their writing activity. As an exercise, we gave them both general and specific prompts, and
they had to write a supporting sentence that was of the opposite type. For example, one prompt was “5 m from a point charge the measured electric field has a strength of 8 N/C, but 10 m from the charge the electric field has a strength of 2 N/C.” A sentence such as the following would be considered a good response: “Electric field strength from a point charge is inversely proportional to the square of the distance.”

In the second lab, the focus was on organization and forming a logical structure. We had students do an activity similar to that done in summer. We had them outline the steps necessary for responding to the previous week’s prompt, then had them organize some concepts into a clear argument. However, instead of returning their own essays, we gave them an essay from the NI group and circled three sentences from it that they were to focus on restructuring and rewriting. In the third lab, the discussion revolved around sentence-level issues: specific language use issues, using transitions, and putting equations into the essays. Students were then given sentences from the previous week’s essays and asked to correct them. For example “At zero seconds, all of the voltage will flow through bulb A” can be fixed simply by replacing the word voltage with current. All sample sentences were taken from the control group so no students would feel uncomfortable seeing their work used as an example of bad writing. The writing activities and instruction are given in detail in Appendix B.

5.2.3 Grading

As was previously mentioned, there were too many students doing the writing for us to return in a prompt fashion, so the essays were not graded on a weekly basis in the autumn project. Our time was very limited those three weeks, because we were
in the labs between ten to twenty hours per week in addition to our own teaching and
workloads. At the end of the project, once we had the final grades of all students,
we graded the essays of two average students from each of the five WI and each of
the five NI groups. These were done as in the summer project, with copies made for
grading the English and physics separately.

5.3 Difficulties We Observed

We observed fairly similar problems implementing writing in the lab as we had in
the summer project. Students struggled while writing essays, even though sometimes
they articulated the ideas out loud while working on the activity. Although we did not
grade every essay so we did not see all of the essays, we did not have difficulty finding
problematic sentences when looking for examples for the writing activities. Although
some students were very welcoming and enjoyable to work with, other students were
very resistant to having us in the lab. This experience was perhaps worse than it had
been in the summer, because we first appeared in the fourth week of the quarter to
work with the students, rather than working with them each week.

5.4 Results

Our results consist of quantitative data, based on essay grades for a subset of
essays, exam grades, surveys, lab and recitation quizzes, and course exams. For each
of the data analyzed below, except for the selected graded essays, there were at least
thirty grades for any individual group, on any specific question. Therefore, a t-test
is a valid measure for whether there is a significant difference between any of the
treatment groups. Unless otherwise specified, each of the significance values listed
are from two-tailed independent sample t-tests. Though the CSEM does not cover circuits, the pretest is a good measure of whether the treatment groups are directly comparable. The pretest scores for each group were not significantly different, but the difference between the NI and C group were significantly different ($\rho = 0.050$), with the NI group having higher scores.

### 5.4.1 Essay Grades

Due to the small subset of essays that were chosen for grading, a t-test is not appropriate analysis; a Mann-Whitney U test was done instead. Each week, the NI group had higher English grades for the writing activities than the WI group. These English grades showed no significant difference until the third week, whether the NI group had an average of 12.2 and the WI group had an average of 10.8 ($\rho = 0.011$). These results are shown in Fig. 5.1. It appears from our data that the writing instruction did not have a positive impact on the WI group. In fact, it is possible the effect was negative. Possible explanations for this observation are discussed later in this chapter.

In contrast, the opposite trend is seen in the physics grades for the writing assignments. Although the WI group started with significantly lower essay scores in physics, by the third essay, their scores were nearly identical. The physics scores were significantly different for the first essay, where the WI group had 5.9 and the NI group had 8.5, but not significantly different in the later weeks. This can be seen in the Fig. 5.2. It thus appears the WI group improved more in physics than the NI group, even though they never did as well any given week.
Which of the Three Circuits Labs

Figure 5.1: English grades for writing activities.
Which of the Three Circuits Labs

Figure 5.2: Physics grades for writing activities.
The correlations between the English and physics grades were not quite as strong as observed in the summer project. For the first essay, the Pearson correlation coefficient between the English and physics grades was 0.492, it was 0.740 for the second essay, and 0.391 for the third essay. The correlations are summarized in Table 5.2. In particular, the correlations for the second writing activity are encouraging. (The lower correlations in the third week are not surprising due to the nature of that assignment.) Not only is the coefficient high between the physics and English grades that week, but the grades don’t correlate well with the grades from the other essays. Once concern is that the correlations are due only to the fact that better physics students may be better at writing. The fact the correlations are weak among different essays indicates that the correlations are due to more than just some innate ability of the students. This is further supported by the fact that the essay grades do not correlate well with midterm or final exam grades.

<table>
<thead>
<tr>
<th>Physics</th>
<th>English 1</th>
<th>English 2</th>
<th>English 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physics 1</td>
<td>0.492</td>
<td>0.251</td>
<td>0.385</td>
</tr>
<tr>
<td>Physics 2</td>
<td>0.537</td>
<td><strong>0.740</strong></td>
<td>0.079</td>
</tr>
<tr>
<td>Physics 3</td>
<td>-0.073</td>
<td>0.129</td>
<td><strong>0.391</strong></td>
</tr>
</tbody>
</table>

Table 5.2: Correlations between English and physics grades for selected essays.

5.4.2 Lab quizzes

The same lab quizzes were given to all students in the course; one quiz is turned in by each group of students who work on the lab together; they all obtain one grade. Significant differences were seen for the lab quizzes for each of the three circuits
labs. Interestingly, the C group performed significantly better on the lab quiz prior to the circuits labs ($\rho < 0.001$ as compared to the NI and WI groups combined). However, in the second and third circuits labs, the C group did significantly worse than the WI and NI groups ($\rho < 0.0001$ and $\rho < 10^{-12}$, respectively). This indicates that the writing activities has an larger impact than the traditional lab activities on physics understanding as measured by the lab quizzes. This is an important measurement, because the traditional lab activities are already well designed, and based on physics education research; the fact that new, untested, writing activities showed an improvement on physics understanding is very encouraging.

It was noted above that correlations between writing activity grades and course midterms and finals was not strong. The average correlation coefficient between the writing grades and the exams was $-0.071$. For the first circuits lab quiz, the correlation was also slightly negative, at $-0.076$, but for the second and third quizzes the correlation became stronger, with values of 0.308 and 0.275, respectively. This is additional indication that the writing activities may have had a positive impact on the lab quizzes, since those who did well on the writing also did well on the quizzes. This is not due to innate ability because no such differences were measured on course exams.

However, there is nothing encouraging to be seen for the effects of the writing instruction. At the end of the first circuits lab, the WI group did significantly better than the NI group ($\rho < 0.05$). However, in the second circuits lab, the NI group did better on the lab quiz than the WI group ($\rho < 0.05$). Although the NI group also did better on the quiz at the end of the third circuits lab, this result was not significant. This result is not as consistent as the differences between the students who wrote and
the C group. It does not seem as if the English instruction had a strong positive or negative effect as compared to the NI group.

I examined each of the lab quizzes in detail, to see if there are specific problems on which the groups performed differently. For the first quiz, the control group performed significantly better than the WI and NI groups combined on one problem ($\rho < 0.05$). Interestingly, this was a question asking students to explain their reasoning to the previous question. A similar explanation question had higher scores in the NI group than the other groups, but the only significant difference was between the NI and WI groups ($\rho < 0.02$). Another question with subparts showed no significant difference
between any of the groups for the accompanying explanations. In two other quiz questions requiring explanations, the WI group did significantly better than the NI group ($\rho < 0.01$ for each). In total, there were eleven questions on the quiz (including sub-questions); the WI group did best on 3 of them, the NI group did better on 4, and the C group did better on 3. All groups had the same score on 2 of the questions. (This sums to twelve because the NI and C groups tied for the best score on one question.) It does not appear that there are any clear trends on this first quiz, which could have been expected since it is doubtful the writing activities could have had an impact after one lab.

In the second lab quiz, either the NI or the WI groups had the highest average for each of the six problems. The NI group did significantly better than the C group for all questions but one ($\rho < 0.02$ for each). For the two questions that the WI group scored highest on, their scores were also significantly higher than the C group ($\rho < 0.01$ and $\rho < 0.05$ for those two questions). An example of a problem for which the NI and WI groups combined scored significantly higher than the C group is the following. “For the circuit above, explain the charge accumulation on both plates of each capacitor from immediately after the circuit is connected to a long time later.” This concept was covered in the lab for all groups, and not explicitly needed in the writing activity. However, in the writing activity, students made use of the charge accumulation to determine if the current flow is increasing or decreasing with time near the capacitors. A similar explanation question, however, had lower scores for the WI group than the C group, though this result is not significant. It appears that the students who completed the writing activities did do better overall relating the concepts to slightly different questions, but these results are still not consistent.
The quiz at the end of the third lab contained five questions. On four of these, the WI group did significantly better than both the NI and C groups ($\rho < 0.05$ for one of them, and $\rho < 0.01$ for the other three). For all but one of these, the NI group also did better than the C group, but not significantly. On the fifth question, the C group did significantly better than both the NI and WI groups ($\rho < 0.001$). This fifth question asked students to explicitly solve for currents in a given circuit, using Kirchhoff’s laws. This circuit ended up being fairly difficult to solve, and only one group out of all fifteen lab sections completed this problem. Because students took extra time to test the Kirchhoff’s laws instructions written by other groups, the time for solving the lab quiz was limited. I strongly suspect the difference in grades on this problem is only due to the lack of time the writing groups had to spend on the quiz. This is further supported by the fact that the previous two questions asked students to write the junction and loop equations, and the WI group did best on these. The other two questions, for which the WI and NI groups did better than the C group, were purely explanations of Kirchhoff’s laws. These had to be articulated in the lab by all groups, but this was done again in the writing activity. This is the only indication that the writing instruction had a positive impact; it seems there is overall little effect of writing instruction measured in this study.

5.4.3 CSEM and DIRECT

As previously reported, there were no significant differences on the CSEM pretest for the three groups; the same is true for the posttests. However, the normalized gain [41] (defined as the number of points improved divided by the number of points that could have been improved) is significantly different between the NI and WI group.
The gain for the three groups are 0.37 for WI, 0.26 for NI and 0.29 for C. It is not clear why the WI group would have improved more on the CSEM since there are no circuits problems on this test. There was another study being conducted in the same class the same quarter: a clicker study run by Dr. Reay. As a compromise, he did not give any extra emphasis to circuits problems in his study, in the hope that the effects of the two studies would not mix. This study involved different treatments between the two lecture sections. The students in the groups for my project were mixed fairly equally between the two lecture sections. Though it is unclear why, it could be possible this other study is the cause for differences in the CSEM.

The DIRECT only covers circuits, and was chosen specifically for testing if there were different levels of conceptual understanding between the three groups after the writing activities and instruction. This was not given as a pretest, because it was felt giving a diagnostic test four times was excessive, and the CSEM was already scheduled. There was no significant difference seen in the DIRECT test though the same score order was seen as with the CSEM gain: 51.6 for WI, 48.6 for NI, and 50.4 for C.

5.4.4 Recitation Quizzes and Course Exams

Although it is encouraging that the writing students eventually did better on the lab quizzes, it would be even more interesting if the students did better on measures in other aspects of the course. The control group did significantly better than the WI group and the average of the NI and WI groups (both $\rho < 0.05$) on this first midterm; this was before the circuits labs. There were no significant differences on the second midterm or the final. Of all the three recitation quizzes, scores were only
significantly different for the first quiz, before the circuits labs. On this quiz, the WI and NI groups did significantly better than the C group ($\rho < 0.05$). There was no difference on the recitation quiz that covered circuits.

I looked more closely at specific questions on the second midterm and the final exam. There were no problems on the second midterm that were significantly different among any of the groups for either lecture section. This midterm covered circuits questions, along with other content. However, there were some problems that were significantly different on the final exams. From Dr. Barrett’s lecture section, the only difference was on the first multiple choice question. The NI group did significantly better than C group ($\rho < 0.05$). However, this result is not interesting, because this question involved forces between stationary charges.

For Dr. Reay’s lecture section, however, there were significant differences on several questions. The C group did significantly better than the WI group (and better, but not significantly, than the NI group) ($\rho < 0.05$) on a problem that involved finding the charge on a capacitor in a circuit with only capacitors and a battery. Students did study capacitor circuits in lab, but all of them also contained bulbs. The WI group did significantly better than NI group, and better than the C group (but not significantly) ($\rho < 0.01$ for two and 0.05 for the third) on three subparts (out of five) for a question on fields and potentials between charged plates. This is not directly related to activities in the capacitor circuits lab, but is somewhat related to discussing the charge flow near capacitors, as they did in lab. The NI did significantly worse than the C group for one of these subparts, but not the others ($\rho < 0.05$). The last question for which significant differences were found was also not related to the circuits lab; it involved finding the net magnetic field due to current carrying wires. The control
group did significantly better than the NI group on this question ($\rho < 0.05$) (and also better, but not significantly, than the WI group).

5.5 Discussion and Conclusions

Although the writing activities seemed to be at least as successful as the research-based lab activities, there is no clear evidence of a difference between the physics knowledge gained by the different groups, as shown outside the circuit labs themselves. It is not too surprising that conceptual gains did not transfer to other portions of the course, as in this course structure the labs are viewed as fairly separate from the rest of the course. Students often report that the labs are not helpful to the course, and it is likely that many students would not consciously reflect on material covered in lab when doing recitation quizzes or course exams. It is also possible that the content knowledge gained in the lecture and recitations dominated any overall effects that would be observed from the lab activities.

There is also no indication that the writing instruction had either a clear positive or negative overall effect on the students who completed the writing activities. There are a few strong reasons why the results here are not consistent with the results from the summer project. Recall that in the summer, significant differences between the NI and WI group were not seen until the sixth writing activity (and fifth writing instruction). It is quite likely that three weeks of writing and instruction were not enough practice for the students to gain much from the writing instruction, or the writing activities. In addition, the students were not provided weekly feedback on their writing, which may have limited the effects of the writing activities and instruction. This provided students less opportunity to reflect on their writing. It is
also possible that the writing activities themselves did not promote as much reflection as the activities completed in the summer project. These activities also did not rely as heavily on a full argumentation, since students were allowed to list instructions - minimizing their need to tie the ideas together as they wrote.

It can not be forgotten that the lab activities that were replaced were tested, research-based activities, involving predictions and tests for various circuits. The fact that the students who wrote did not do worse than the control group does indicate that we managed to teach the physics concepts as well as a proven method. This is perhaps already a strong result, since we managed to teach conceptual knowledge and allow students to practice needed writing skills at the same time, with minimal extra effort and instructor time (discounting the writing instruction).

However, the bottom line is despite a large-scale controlled study, there is still little evidence that writing impacts conceptual understanding. Conducting a similar study with improvements to minimize these confounding factors would help determine if this is truly an overall null result.
CHAPTER 6

THE STUDY OF REVISIONS

According to Fitzgerald [34], when writers use revision to rework their thoughts and ideas, it may affect the writers’ knowledge. The revision process helps allow new lines of arguments to form from a reorganization of ideas, discovering new ideas in the process. Fitzgerald summarizes a view of revision based on problem solving, that is at the core of the belief that learning takes place during revision. The first is when the readers identify discrepancies between their intentions and the actual text. They must be able to judge their own text and read it from an observer’s perspective. They diagnose what changes or alternatives are needed to fix the discrepancies. Finally, they carry out those changes.

Although revising is viewed as a central part of writing, it has not been studied extensively. Until the 1970s, writing was viewed as a linear process of pre-writing, writing and post-writing (sometimes called prevision, vision, and revision). It was later pointed out that revision can occur during each of these steps, leading to a more complex and nuanced view of how people write. Revisions were also categorized in two fashions: internal (“everything writers do to discover and develop what they have to say”), and external (“what writers do to communicate what they have found”) [26]. Updated cognitive models of writing are described in Hayes and Flower [36]. Bereiter
Scardamalia [20] drew notice to the fact that changes a writer makes before writing should be addressed just as much as those made while writing. It was also pointed out that reviewing a text can lead to revisions.

Typical research on revisions involves coding student work for various hierarchies of revision. Most of this type of research shows what revisions are made, and when in the writing process, but do not show why the writer chose to make those revisions. Early coding schemes were not rooted in a cognitive process theory of writing, and often included categories that were not mutually exclusive. In addition, different types of changes were not clearly distinguished and some kinds of revisions were neglected. Later papers showed more detailed coding schemes. For example Faigley and Witte “devised the first taxonomy of revisions that would account for revisions related to the semantic structure of text, not just syntactic aspects. The taxonomy distinguished characteristics of changes such as surface and meaning and microstructure and macrostructure features [32].” Another coding scheme can be found in Bridwell, which breaks each type of change down by the level, from word level to sentence level to text level and between [23].

In order to track revisions, four main types of method were prevalent as of Fitzgerald’s 1987 paper [34]. The first is to use “think-aloud” method, where writers state out loud what they are thinking while they write. The second is questionnaires, the third is interviews, and the fourth is asking individuals to tape evaluations of their work after each draft. In interviews, a writer may be asked about reasons or goals behind a change that was made, or can be asked about changes that could or should be made to a specific location in their paper. More recently, keylogging software has been used to track the writing and revision process electronically [27].
Research involving instruction on the revision process has included methods in which a writer and an observer sit together, and the observer interacts with the writer while taking notes on the writing process. An instruction method called “simulation-by-intervention” was developed by Scardamalia and Bereiter [20]. In this technique the writers had statements to guide the writing process such as “People won’t see why this is important.” (This is also used in their paper on sixth-graders cited above.) Different types of instruction given in this technique include: procedural support, direct instruction, teacher or peer feedback, or giving directions. Each of these types of revision instruction has yielded visible improvements in student writing, except giving directions, which has yielded mixed results.

Results from research on revision processes are often mixed and depend on many factors. There are some very concrete reasons why the revision process is not simple to study. For one thing, experienced writers may be very good at reviewing and revising their texts to make them good, but they may also be very good at thinking about what to write before they write it, and choosing their wording carefully, thus putting down good text the first time needing little revision. This is evident from my own interviews with students. Also, the diagnosis of problems in writing and revising is complex as there can be many reasons for them. Writers may not clearly know their intentions before sitting down to write. Writers may not have the instruction or skill to know how to edit their text. Writers may lack the ability to review their text keeping in mind the viewpoint of a third party [17, 35]. Writers may lack knowledge about the subject itself, or simply have trouble recalling the details, or may not have the skills to put it into words. Studies suggest that the revision process may be blocked when students have difficulty juggling goals of how to express their ideas and
what content is necessary. One study showed when writers have a lot of presentation-related constraints to focus on they made fewer content revisions [40]. It also seems writers are more reluctant to edit their own texts than the texts of others, even at the level of elementary children [17]. In addition, is is believed that the process of managing all the operations that are necessary for good writing can be difficult even if a writer has the individual skills.

Typical studies involve how much revision occurs, when it occurs, and what kinds of revisions are made. Many of these studies do not try to correlate the revisions with the quality of the text, or whether the revisions are “good” choices on the part of the writer. Faigley and Witte [32] found that expert and inexperienced adult writers agreed little on how to edit first drafts, but Scardamalia and Bereiter [74] found that fourth-graders’ choices did not match those of experts, sixth graders’ choices matched some, and eighth graders’ choices matched well (though justifications of why they made those revisions did not match with those given by experts). Differences in revision processes do emerge clearly between novice and expert writers. Children focus on very specific details when looking for changes needed in text, while experts deal with issues on the level of the entire text [74]. While, on the whole, most revisions people make are on the word level, or on the surface level (“revision as proofreading”), it has been verified in many studies that experts make more text-level, or idea-level revisions while novices focus on details and word-level changes. This has been studied at all levels, including the high school level and beyond. Faigley and Witte [32] found that expert professional writers made one meaning-related revision for every surface change, advanced college student writers made one for every three, and inexperienced college writers made one for every seven. Improvements in these
traits have been measured between fourth and eighth grade [74]. In addition, more experienced writers are more likely to improve their text by revision, which is not always the case for revision by children. Even in unskilled college students’ revisions have been shown to result in worse drafts [67].

Revision is content dependent. Students revise social science papers more than science papers, but less than English papers. Also, students revise more if a topic requires theorizing or analyzing than if a topic requires reporting or summarizing [24]. Some evidence shows that more revision is done if students write on computers than with pen and paper, though such additional revisions may be primarily surface-level.

Fitzgerald [34] suggests that more effort should be put into understanding what students are thinking while making revisions. It is difficult to analyze results because one action may involve a number of different cognitive processes. For example, a pause while writing to re-read can facilitate the generation of new ideas, or facilitate revisions [21]. Fitzgerald also suggests more effort be put to understanding the pre-writing and between writing revisions students make in their minds. She also suggests more research be done to correlate the quality of the text and the quality of the revisions with the factors already studied. She also recommends coming up with new ways of studying writing to help understand when the revisions take place and how they break down. She also suggests more studies of revision across multiple samples of writing from a single writer.

6.1 Specific Studies on Revisions

The following give details from three projects aimed at studying revision processes.
6.1.1 Huub Van Den Bergh and Gert Rijlaarsdam [21]

These authors pose the questions “how are cognitive activities distributed over the writing process; which activities are more dominant at the beginning of the process, which are the end?” “do writers differ in the way the cognitive activities are distributed over the writing process?”, and “what is the impact of these distributions on text quality.” They aim to use information about when revisions occur to help understand the cognitive processes. They studied 36 ninth graders who wrote an essay in an argumentative style (the topic was “living alone: yes or no”). The students wrote using the “think-aloud” technique and the sessions were independently coded. They broke the types of revisions into two main categories: “formulating” and “text representation”. Formulating activities include things such as thinking aloud about their own ideas or the assignment. Text-representation activities include thinking aloud about how to phrase an idea. Formulating activities occurred far more often. Most students had more task-representing activities initially, then had almost none by the end of writing, where formulating activities increased then decreased again by the end of writing. These trends correlated with the quality of text, as well (students who deviated from these temporal curves had poorer text). They claim this lends support to the old adage “think first, then write”, and suggest that prompting task-representation activities in the beginning of writing may be an effective teaching strategy.

6.1.2 Lillian Bridwell [23]

Bridwell’s research questions are: “what do twelfth graders do when they revise, are there any differences between the patterns of more successful and less successful
twelfth-grade writers, and do these patterns reveal evidence that might be useful for an evolving theory of composing processes?” She took almost all students from the twelfth grade in a high school, and randomly selected 100 of them to study. The students were given an explanatory task of telling about something they were familiar with (for example, a place). Many students argued why the thing was good, and did not just explain it, so this writing style is called “transactional” involving both “explanatory discourse” and argumentation. The students worked on their writing for three days, the first day doing one draft, the second day marking the draft for edits, and the third day writing a final draft. Markings on each day were done in different color, and independent coders coded them. She chose to consider the revisions on a linguistic structure. The first level is surface features: spelling, punctuation, singular vs. plural, and other minor edits. The second is lexical, or word-level: additions, deletions, substitutions, or switching two words. The third, fourth, and fifth are phrase, clause, and sentence level, respectively. These include additions, deletions, substitutions/alterations, order shifts, expansions and reductions (for example from a word to a phrase or vice versa), and marginal notations. The sixth level is multi-sentence, and includes the previous items and indentation choices. The highest level (seventh) is text-level, including changing the function of the essay, changing the audience, the content, or near total rewrite. She did not use the seventh level, and she saw very little at the fourth (clause) level.

She found that the quality of second drafts were significantly better than first drafts. She found that level 1 revisions were significantly positively correlated with level 2 revisions, but not with any other level. This shows that a propensity to attend to surface-level matters is not associated with attending to larger units of
language. Revisions at levels 2-5 were significantly positively correlated with each other. Most multiple-sentence level revisions occurred in the third sitting, suggesting that the longer students had to explore their ideas, the more likely they were to make higher-level changes. She also found that surface level changes were made more when students were writing in the first and third sitting than when they were reviewing and marking in the second sentence. This is in contrast to their having been taught that they should write their ideas and fix the details while revising. It was observed that students had so many technical problems that they had to fix them while writing to get their ideas out. Students also spent a great deal of time trying to find the appropriate words to express their thoughts. “Many students would write a word on the page, delete it, add it again, substitute a similar word, and so on, almost though the recurring lament ‘I know what I want to say, but I don’t know how to say it’ were revealing itself in the revising process.”

The best predictor she found for “second draft mechanics” were sentence-level revisions. This suggests that students who could manipulate large units of text have better mechanical skills. In confirmation of some of the difficulties with revisions mentioned above, she found that students who produced good quality text were spread from those who revised extensively to those who hardly revised. However, of those who revised very little, more than half were below average in quality. Papers that were extensively revised ranged from good to bad, however.

6.1.3 Eva Lindgren and Kirk P.H. Sullivan [58]

This paper discusses a way to layer graphs and gain information from the output of keylogging programs. They warn that the log files contain no information about
the writer’s underlying intention. They suggest pairing keylogging with a think-aloud protocol to gain both bits of information. They also suggest graphing based on the total numbers of characters typed so non-linear changes are seen more clearly. Their graph has the total number of characters typed on one axis, and time on another, so that information on the temporal spacing of revisions is not lost. They plot the total number of characters typed, the number of characters retained, the time of each revision, and the distribution of editing activity within the text. They also layer information from the “think-aloud” on top of this. They show some examples of these graphs, and show how they help the researcher see more into how the writing progressed, and thus into the nature of the revisions. They admit it does not tell why writing is interrupted, or the nature of individual revisions and what the writer was thinking.

6.1.4 David Wallace and John Hayes [80]

This paper discusses a study in which half of freshman college students were given 8 minutes of instruction on global revisions, and all were asked to revise a text given to them. Wallace and Hayes claim that there is a lack of skills in students to detect and fix both local and global problems, and suggest that this may be the biggest problem students face when doing revisions. They also state that even if students have the skills, they may not have the ability to coordinate them simultaneously to adequately address a text. For example, “students might fail to attend to the interests of their audience because they are completely absorbed in dealing with problems of grammar or word choice.” The third difficulty they discuss is inappropriate task definition:
“the writer’s understanding of what he or she is supposed to do when facing a writing task such as revision.”

Wallace and Hayes state many researchers have found that “inexperienced writers treat revision as a local task, that is, a task of changing words and sentences rather than of modifying the goals or organization of the text to meet criteria of the rhetorical situation.” Their idea behind 8 minutes of instruction was to cue students on task definitions, and help them cue abilities they already possess. They gave students a text about water processing at a treatment plant because they state that “procedural texts require global coherence.” The text included local and global problems. The instruction they gave showed how an expert and novice revise similar texts. Differences were illustrated using overheads. Both the procedures and approaches of the two revisers were discussed; the expert reading the whole text then revising, while the novice makes changes immediately while searching the text. The instruction emphasized the types and amount of changes, noting the global issues the expert addressed.

To analyze their results, the 38 texts were rank-ordered based on final quality and separately ranked based on global revisions. Local revisions they noticed were spelling, vocabulary improvements, making wording less awkward, and adding temporal cues. Global changes included adding or revising sections of text, including purpose statements, reordering text, and adding transitions. The students with instruction did significantly better on the global revisions ($\rho = 0.008$ level from a Mann-Whitney test). Since revisions are not always improvements, it is not obvious that more revisions will lead to better text; however, the correlation between final text quality and revisions was 0.805 (found using a Spearman rank correlation). The students who received instruction had better quality final texts ($\rho = 0.012$ from
Mann-Whitney). They also checked to make sure that the bulk of the students who received instruction did better, not just the best of the writers in the treatment group. The groups did not have a significant difference on the number of local revisions made. They question the effect of having students revise someone else’s text and wonder how different things would look if they were revising their own work.

6.1.5 Brian Monahan [63]

This paper discusses the study of 8 twelfth grade students, split into two groups: basic and competent writers. These students wrote two essays of similar topic, but one was written for the audience of a teacher and the other for peers. Research suggests that the audience will affect writing decision making, such as how students approach revisions. Expert writers exhibit more audience awareness [37], as well as more successful revisions. He quotes a study by Pianko [68] stating that better writers spent more time pre-writing and rescaning, and they pause more often to think. This paper suggests “that those pauses may have given them a clearer conception of the content of their essays and helped them reflect on what was being written.” Flower and Hayes [37] report that better adult writers generate a majority of their ideas based on larger rhetorical problems such as the audience and goals, and less based on the topic itself, while the ratio is switched for poor writers.

In this paper Monahan analyzes the writing based on the point, level, type and purpose of the revision. Points describe what part of the writing stage (pre-writing, second draft and so forth). Levels are surface, word, phrase, clause, sentence, paragraph, and discourse. Types include addition, deletion, substitution, reordering, and embedding. Purpose includes cosmetic, mechanical, transitional, informational, and
stylistic. For example, a revision might be coded as: “discourse level substitution during the final draft for cosmetic purposes.” He considers one basic and one competent writer for case studies, and also compares all texts. The competent writer embedded a series of revisions while going back and rereading. She was also aware that changes she was making were to make things more clear for the audience. She paused less often, but longer than the basic writer. The basic writer was often pausing and considering individual sentences at a time, while she was looking at larger chunks. Her revisions were also extended episodes, where revisions cued additional changes, while the basic writer did not do that. Both sets of students exhibited the same types of revisions, but not with the same percentage or frequency.

6.1.6 Lester Faigley and Stephen Witte [32]

This paper discusses a new coding for types of revision, and applies it to various types of writers. Faigley and Witte mention that most papers consider the revisions based on how much text they affect, but not on how the changes affect the meaning of the text. They define surface changes as those that don’t change the information in the text and meaning changes as those that do. Within surface changes, they discuss formal ones (such as spelling, tense, and format) and meaning-preserving ones (such as rephrasings or permutations). Their primary mode for seeing the difference was to see if the revision affected the understanding of other parts of the text. They remind the reader that inexperienced writers seldom revise during composition except to correct errors, while experts review to improve and also to generate new content. Since some experts make changes in their mind, expert adults they studied made
fewer revisions than the advanced student writers. Also, 65% of changes by skilled adults were macrostructure changes.

6.2 OSU’s Studies of Revisions

One idea that stands out from the information above is that the concept of revision is very broad. Here is a list of some definitions of revision from the literature:

1. “Revision is bringing the writing into line with the writer’s intentions [77]”

2. “The goal of revision is to improve the text so that the writing task is solved more appropriately [58]”

3. “Revision can be seen as the detection and correction of, for example, spelling and grammar mistakes, errors of fact, incoherence, disorganization, faulty text structure, inappropriate tone, and ambiguities of reference [45]”

4. “Revision means making any changes at any point in the writing process. It involves identifying discrepancies between intended and instantiated text, deciding what could or should be changed in the text and how to make desired changes, and operating, that is, making the desired changes. Changes may or may not affect meaning of the text, and they may be major or minor. Also, changes may be made in the writer’s mind before instantiated in written text, at the time text is first written, and/or after text is first written [34]”

5. “What the writer does after a draft is completed to understand and communicate what has begun to appear on the page [26]”

Flower and Hayes [44] make a clear distinction between reviewing and editing. They define editing as an automatically triggered event, which may occur in brief
episodes, interrupting other processes. On the other hand, reviewing occurs when
the writer decides to devote time to systematic examination and improvement of the
text. They claim this typically occurs after an episode of writing, rather than as an
interruption to the writing.

For the purposes of this study, revisions are defined as modifications or pauses in
the writing or editing process that are non-linear, and/or do not appear at the end
of the text already written. For example, if a student is writing and jumps back to
address previously written text, that is a revision event. Although this is a limited
view of revisions, we have no process for studying any pre-writing, nor the ideas that
would be captured using think-aloud protocols. The details of our data for studying
writing and revising are given in the next several chapters. In the context of what
has been studied, here are the types of things we aim to study specific to revision
behavior.

With our data on revisions we have three modes of study:

1. Study revisions by a single student across time and many writing samples.

2. Compare revisions across time (different essays through the quarter).

3. Compare revisions by students across different quality/grade levels.

We can consider additional revision factors:

1. Percentage and number of revisions.

2. Types of revisions (in specific, the percentage of revisions that are content-rich
   versus surface-level).
3. Interval between jumps when students revise (do students look at one clump at a time, or revise much of the text at once?).

4. How linear student writing is (how often the revisions are out of sequential order within the text).

When looking at the second drafts, we can consider:

1. What did students cut?

2. Do students rephrase things they don’t cut to make the context flow?

3. Do students shorten things they want to keep, or just cut large chunks to make the word count?

Additional factors we can consider:

1. The quality of the text on both drafts and how this correlates with revision behavior.

2. Where students pause, and how often they pause.

3. How long it takes the students to write the entire text.
CHAPTER 7

WINTER 2006: WRITING IN PHYSICS BY INQUIRY

During Winter quarter, 2006, we implemented essay writing in all three sections of the Physics by Inquiry (PbI) course at the Ohio State University main campus, as well as one section of PbI at the Marion campus. This particular quarter, the course covered balance, measurement and error, density of inhomogeneous objects, and concentration. Instead of focusing on whether writing affects content knowledge, we decided to take a step back and look more carefully at the student writing. We focus on the following questions:

1. Can we quantify behaviors in student writing?

2. Can we observe differences between these behaviors in different types of students (for example, high-scoring versus low-scoring)?

3. Do these behaviors change with practice as the quarter progresses?

We think it is important to understand student behaviors while writing science content before we can best utilize writing to positively impact learning.
7.1 Implementation Details

Due to the aforementioned problems of student resistance to writing (as seen with the difficulties in the Summer and Autumn quarters), and the disjoint nature of the lab and lecture setting in the traditional introductory course, the PbI classroom seemed like a more natural place to study student writing. This course is primarily for education majors, to instruct them in inquiry-based science teaching methods. Many students are upper-level undergraduates, though some are masters level and others are in-service teachers returning for coursework toward additional teaching certifications. In this course, students always work in a lab setting with no lecture; there are no separate recitations or labs. Standard assignments include traditional problems and weekly journal writing; in addition, students take notes on the experiments they perform and the conclusions they draw from them. Students are expected to write detailed explanations on their exams; in fact, the exams essentially are essays. Students understand writing to be a regular part of the course. Therefore, writing essays in this course is consistent with the course goals and may assist students in organizing their ideas for the exams.

7.1.1 Writing Activities

The writing activities given to the PbI students were modeled after assignments being given simultaneously by our collaborators at the Rochester Institute of Technology (RIT). The essays were intended to prompt students to broadly explain content developed in the course and to bridge that content to an application related to daily life. For example, one assignment stated: “Other than simply using “trial and error,” explain how to balance two objects of unequal mass in a way that would be clear
for an intelligent but uninformed person. Also explain how this applies to real-world examples such as using scales or children on see-saws. Refer explicitly to in class observations and/or activities.”

Four of the five essay topics were to be completed in two drafts, over the course of two homework assignments. The first draft was required to consist of between 225 and 275 words in length. The second draft was required to be no more than 125 words. Having two drafts was crucial for this study because we wanted to study the choices students made when revising their work. The shorter length of the second assignment served two purposes: first, to force students to make changes from their first draft, and second, to force students to focus on what information is most important to include. In the words of Scott Franklin at RIT, “it is very difficult to cut your essay in half without doing some major restructuring in your head, deciding what is important and what is not.” When completing their second draft, students were requested to use Microsoft Word’s “Track Changes” option, which is discussed in more detail below.

There were nine homework assignments given during the quarter; eight of these were first and second drafts of four essay topics. The ninth assignment could not be given in two drafts; therefore, the students were asked to address the essay topic directly in 125 words or fewer. This gave us a way to determine if students were able to complete a good quality “final” draft successfully on the first attempt, having practiced such assignments all quarter.
7.1.2 Grading

Each week, the homework was worth ten points, with two of those coming from the essay grade. The first draft of the essays were given the full two points if completed within the given range for the number of words. Fewer than 25 words from the maximum or minimum word limit received 1.5/2 points, while students who deviated more than that received 1/2 points. Comments were returned for all students to help them correct their work before submitting the second draft. However, it was important not to give traditional writing feedback, such as circling or otherwise pointing out specific parts of their essays for editing. It was important that we minimally influence the choices students made when revising. However, if some important aspect was missing or unclear, a comment such as “how will the mass be determined?” was written at the bottom of the essay. If all information was present and clear in the first draft, students received a simple comment such as “good”.

For the second draft, students were given 2/2 points for a complete, correct, and clear essay. They were given 1.5/2 points if something was unclear, or if some small point was missing. They were given 1 out of 2 points if there were larger errors. Few students received fewer points unless a half or full point was deducted for going over the word limit. If students exceeded the word limit by a few words nothing was deducted; if exceeded by between five and 25 words, a half a point was deducted; and if exceeded by more than 25 words, a full point was deducted. Feedback was given to indicate what was good in the essays and, if applicable, why points were deducted.
7.2 Difficulties Experienced

Although the students in PbI were more accepting of the essay assignments than the engineering students had been, the implementation still encountered some small problems. Although among the four sections, 135 student completed the course but only 64 students completed the final essay assignment. Even fewer students completed all nine essay assignments. Since the essays were worth only one-fifth of their homework grade, some students felt they took too much time and were not worth their point value. However, many of the students who were not submitting essay assignments were also skipping other assignments such as journals.

Initially many students complained that it was not possible to answer all parts of the essay assignment within the word limit and complained that points should not be taken off for exceeding the word limit if that was necessary to respond to the question in full. In order to alleviate this complaint we provided students with a good essay written by one of their classmates that responded to all parts of the essay within the word limit. A small discussion was included as to how the student achieved concise, clear points in the essay.

Students also complained that the essays were graded too harshly, feeling that 1.5/2 was only 75%, and thus too low a grade for a good (but not great) essay. It was pointed out that between their first and second draft the grade was actually 3.5/4 points (88%), and thus accurate for work with is good but not great. It seemed some of the instructors were not in agreement with this grading and had supported the student complaints, resulting in some students deciding not to complete the essays. This is unfortunate, since the grading was approved by the main course instructors, and I was not notified of any problems until harsh complaints arrived from the students.
Some students cut bulk from their first draft to achieve the lower word limit without much regard to the final content in their second draft. Often important points would be missing from the second draft that had been present in the first draft. Once students realized each essay would be cut in half for the second draft, other students started to add filler to their first draft so their second draft could be completed easily by removing those parts. Although this somewhat undermines the purpose of the second draft, it is not considered a bad behavior on the part of the students. Many students initially had difficulty responding to the essay in full within the word limit, so their ability to achieve the second draft word limit on the first draft and needed to add filler to reach the first-draft word limit shows the students were better able to write concisely as the quarter progressed.

7.3 Data Obtained

Of the students who completed most assignments using “Track Changes” and gave consent to participate in this research study, thirty were chosen for further analysis. Ten students were chosen from the top third of the class based on their final exam grade, ten were chosen from the middle third, and ten from the bottom third. The final exam grade was well representative of their final course points, with these students falling into similar tiers based on them as well. Although it is possible that their course points may not correlate at all with their writing ability, they should correlate with student understanding of the content knowledge and thus be a reasonable way to separate the students into categories for further analysis of their writing.
For these thirty students, each of the nine essays were graded in detail based on the clarity and correctness of the ideas necessary in order to answer the question. Each idea was given between zero and three points: 3 if included, clear and correct, 2 if included but not clear or not quite correct, 1 if alluded to, or included but incorrect, and zero if not included at all. They were also given scores between zero and three based on the overall correctness of the essay, the logical order of the points within the essay, and the introduction, conclusion, and transitions within the essay. Three to five important points were determined for each essay, giving the students a total maximum possible score of 18 to 24 points, depending on the essay topic. In addition to providing more information about the quality of each essay, this provides an accurate way to determine any change in quality between first and second drafts of each essay.

From Microsoft Word’s “Track Changes”, each revision students made when writing their second draft was coded in detail. “Track Changes” records the location and time of all additions and deletions to a document. The location of the revision within the final text was recorded based on the number of words prior to the revision, as well as the time of each revision. Each revision was coded based on those used in Ref. [23]. These codes record the size of the revision, as well as if it was an addition or deletion, or in the case of a change to a single word, it notes things such as a spelling or verb tense change. For example, 3.2 would mean a deletion of a phrase while 5.1 would mean the insertion of a sentence. The purpose of the change was also coded following Ref. [32]; with changes marked as being mechanical, transitional, informational, or stylistic. A general assessment of the quality of the revision was made, each was coded as being a good, neutral, or bad change. For some changes it
was also noted whether the text either contained irrelevant information or had been too wordy.

From this coding, the number of revisions was determined, as well as the number and percentage of additions, deletions, single word changes, single word additions/deletions, phrase level additions/deletions, sentence level additions/deletions, and multi-sentence level additions/deletions. The percentage of good versus bad changes was recorded, as well as the percentage of changes that were for each of the four specified purpose codes. The number and percentage of revisions made in the first 25% of the final essay text was recorded, as was the percentage in the second 25%, the third 25%, and the fourth 25% not including at the end of the essay, as well as the number and percentage of revisions made at the end of the essay.

7.4 Results

The data obtained provide us with many ways of analyzing student writing. We have two dimensions across which to study the data: based on the ranking of the students, and across time. This addresses the two major questions: do students with different levels of content knowledge approach the writing differently, and does the way a student approaches the writing change with practice/time. These questions can be addressed both from the standpoint of considering their detailed essay grades, and considering their revision habits as tracked in the second drafts. In addition, we can use this data to work toward characterizing revision behaviors.

7.4.1 Analysis of Detailed Essay Grades

It is interesting to see whether students who had better content knowledge (as measured by their final exam grade) performed differently on their essays, and whether
those differences were persistent throughout the quarter. Recall that the essays were graded in detail based on specific points needed to respond to the prompt, based on overall correctness, point ordering, and flow/transition. The total score for each essay based on the included points was calculated and named the content score. The total score based on the correctness, ordering and flow/transition was named the writing score. The sum of the content and writing score is the total score for the essay. From this total score, the change in the essay score from the first to the second draft was calculated.

Two types of tests were run on the data for each of the nine essay assignments. Both of these tests used are non-parametric because there are not a lot of data points for each assignment (there are at most 30, but not every student submitted each assignment). Since students were further grouped into high-, medium-, and low-scoring categories based on their final exams, these groupings contain no more than 10 sets of grades for any given assignment. The first test determines correlations between variables: it is the Spearman correlation, yielding a correlation coefficient comparable to the parametric Pearson’s correlation. A two-tailed version of this test was used to avoid prior assumptions about which group would have higher average scores. The second test determines if the groups are likely to have come from different populations: it is the Kruskal-Wallis test, the non-parametric version of ANOVA. All significant differences reported in this section are calculated using one of these two tests.

In the first assignment, it is interesting that the students final exam grades correlate with their essay content grade with $r = 0.420$ ($\rho = 0.029$) but not the writing grade with $r = 0.166$ ($\rho = 0.407$). This is a possible indication that students with
stronger content knowledge initially produced better content in their essays, but not better writing. This statement can not be made strong because the final exam grade may not be a good measure of content knowledge early in the quarter. Even if this is not true, this still indicates that students who could initially produce good content did not necessarily know how to approach writing the essays. In fact, the content and writing grades for the first essay did not correlate well, with $r = 0.133$ ($\rho = 0.501$).

All correlations are summarized in Table 7.1.

It can be seen from the table that the final exam grade correlated significantly with essay content grade for the first six essay assignments, but in assignment seven and eight the correlation is actually negative. It returns to being positive in the ninth week, but is not significant. The essay topic for weeks seven and eight involved predicting whether an object will sink or float in a liquid. This topic did not require extrapolating information from class to a slightly different situation, so was perhaps easier for the students with weaker content knowledge to manage. It is also possible that with essay writing practice, sheer course content knowledge is no longer a good predictor of how well a student will do on the essay content; the students may be gaining other writing skills that would be better predictors. The correlations between the final exam grade and the total essay grade show similar results.

The correlations between the final exam grade and the writing grade for each essay do not show a clear pattern. For essays one and two there is not a correlation, while for essays three and four there is a significant correlation. Essays five and six are almost and significantly correlated, respectively, while essays seven and eight are not. Essay nine has a negative (but not significant) correlation between the final exam
<table>
<thead>
<tr>
<th>Correlation Variables</th>
<th>Week 1</th>
<th>Week 2</th>
<th>Week 3</th>
<th>Week 4</th>
<th>Week 5</th>
<th>Week 6</th>
<th>Week 7</th>
<th>Week 8</th>
<th>Week 9</th>
</tr>
</thead>
<tbody>
<tr>
<td>Final vs. Content Grade</td>
<td>.420</td>
<td>.397</td>
<td>.514</td>
<td>.378</td>
<td>.457</td>
<td>.413</td>
<td>-.060</td>
<td>-.123</td>
<td>.203</td>
</tr>
<tr>
<td></td>
<td>(.029*)</td>
<td>(.041*)</td>
<td>(.006*)</td>
<td>(.043*)</td>
<td>(.014*)</td>
<td>(.036*)</td>
<td>(.776)</td>
<td>(.577)</td>
<td>(.451)</td>
</tr>
<tr>
<td>Final vs. Writing Grade</td>
<td>.166</td>
<td>.032</td>
<td>.521</td>
<td>.511</td>
<td>.335</td>
<td>.440</td>
<td>.085</td>
<td>.074</td>
<td>-.018</td>
</tr>
<tr>
<td></td>
<td>(.407)</td>
<td>(.873)</td>
<td>(.005*)</td>
<td>(.005*)</td>
<td>(.081)</td>
<td>(.025*)</td>
<td>(.686)</td>
<td>(.737)</td>
<td>(.947)</td>
</tr>
<tr>
<td>Final vs. Total Grade</td>
<td>.481</td>
<td>.341</td>
<td>.572</td>
<td>.444</td>
<td>.469</td>
<td>.515</td>
<td>.007</td>
<td>-.150</td>
<td>.148</td>
</tr>
<tr>
<td></td>
<td>(.011*)</td>
<td>(.082)</td>
<td>(.002*)</td>
<td>(.016*)</td>
<td>(.012*)</td>
<td>(.007*)</td>
<td>(.975)</td>
<td>(.495)</td>
<td>(.584)</td>
</tr>
<tr>
<td>Content vs. Writing Grade</td>
<td>.133</td>
<td>.379</td>
<td>.629</td>
<td>.571</td>
<td>.585</td>
<td>.348</td>
<td>.459</td>
<td>-.258</td>
<td>.709</td>
</tr>
<tr>
<td></td>
<td>(.501)</td>
<td>(.047*)</td>
<td>(.000*)</td>
<td>(.001*)</td>
<td>(.001*)</td>
<td>(.076)</td>
<td>(.018*)</td>
<td>(.224)</td>
<td>(.002*)</td>
</tr>
<tr>
<td>Change vs. Total Grade</td>
<td>.456</td>
<td>.030</td>
<td>-.385</td>
<td>-.134</td>
<td>.095</td>
<td>.153</td>
<td>.108</td>
<td>.478</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.002*)</td>
<td>(.880)</td>
<td>(.047*)</td>
<td>(.515)</td>
<td>(.644)</td>
<td>(.486)</td>
<td>(.633)</td>
<td>(.021*)</td>
<td></td>
</tr>
<tr>
<td>Change vs. Final</td>
<td>-.055</td>
<td>-.385</td>
<td>-.134</td>
<td>.095</td>
<td>.153</td>
<td>.108</td>
<td>.478</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.790)</td>
<td>(.047*)</td>
<td>(.515)</td>
<td>(.644)</td>
<td>(.486)</td>
<td>(.633)</td>
<td>(.021*)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change vs. Content Grade</td>
<td>.399</td>
<td>.054</td>
<td>.154</td>
<td>-.091</td>
<td>.647</td>
<td>-.037</td>
<td>-.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.038*)</td>
<td>(.783)</td>
<td>(.453)</td>
<td>(.680)</td>
<td>(.000*)</td>
<td>(.853)</td>
<td>(.940)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Change vs. Writing Grade</td>
<td>.647</td>
<td>-.037</td>
<td>-.015</td>
<td>.478</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(.000*)</td>
<td>(.853)</td>
<td>(.940)</td>
<td>(.021*)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7.1: Correlation coefficients for Winter 2006 essay grades. Significance levels are in parenthesis, and those marked with an asterisk are $\rho < 0.05$. 

110
grade and the writing grade. This does not support the idea that other skills that may correlate better with the essay writing are being gained.

In the initial week, the writing and content grades did not correlate strongly. However, in almost every other week they did correlate significantly. This is consistent with the results from Summer 2006 showing a strong correlation between physics and English grades. Week eight, however, stands out because the correlation is negative (though not significant). It is positive and significant for week seven, which is the first draft of that same essay topic. This is the same topic mentioned above about sinking and floating. For this second draft, the students who did better on the final exam had somewhat better writing but their content was not as good. This is consistent with the idea that the content in this topic was more approachable by all students.

Even more puzzling are the correlations involving the changes in grades from the first to the second drafts. For the first three of the four second drafts, students with higher final exam grades actually had less improvement in grades. This is not due to a ceiling effect where they could not improve more; in many cases students actually had lower scores on the second draft. Only for the second essay was there a significant difference between the grade change from the first draft and the final essay grade. This is true also for the correlation with the content grade. However, for both the second and eighth essay (two of the four second drafts) there was a significant correlation between the grade change and the writing grade on the second draft. For the fourth and sixth essays, however, this correlation was negative (but not significant). This will be explored further in the next section, under the hypothesis that the change in grade between the first and second draft is more related to the students revision habits.
The other test used was to determine if the high-, medium-, and low-scoring students did significantly differently on the essay grades. In the first essay, only the total essay grade showed significant differences ($\rho = 0.023$), with the students scoring better on the final doing better overall on the essays. For the second essay, only the content grade showed a significant difference ($\rho = 0.022$), again with those scoring higher on the final doing better on the content grades. For week three, content, writing, and total grades were all significantly different ($\rho = 0.010, 0.014$, and $0.007$, respectively), again with students scoring better on the final doing better in each of these categories.

For the fourth essay, the content grade, total grade and the change in grade from the first draft all showed significant differences between the groupings based on the final exam grade ($\rho = 0.050, 0.049$, and $0.044$, respectively). However, while the grades were higher for the students with higher final exam grades, the changes in grade from the first draft were lower for students who did better on the final. It does not appear that this is a ceiling effect. The average score for the students with high final exam grades was $17/24$, while the average change in the essay grades was only $1.48$. This is only about a fifth of the points they could have improved.

In the fifth essay assignment, similar results are seen except that the group of students who scored average on the final exam did significantly ($\rho = 0.011$) better on the writing grade than those who scored high on the final exam. In the sixth assignment, the content and total essay grades are significantly correlated (both $\rho = 0.001$) but the writing grade is not, nor is the change in grade from the first to the second draft. No aspects of the essay grade were significantly different in the seventh or eighth or ninth essay assignments. These results are summarized in Table 7.2.
7.4.2 Analysis of Revision Behaviors

Looking at essay grades only gave a hint of what is going on with the student writing. It may be possible to gain a better understanding of the differences by looking at the details of how students write. With “Track Changes,” we only have information on how students approached revisions, and thus can study the details of how students approached the second drafts in assignments two, four, six, and eight. The same two statistical tests are used: the Spearman correlation and the Kruskal-Wallis test.

First, we look to see if any of the tracked revision behaviors correlate with final exam grades, essay content, writing, and total grades, and changes in grades from the first draft. In the second essay assignment, the percentage of additions negatively correlated with both the final exam grade \( r = -0.407, \rho = 0.039 \) and the essay content grade \( r = -0.330, \rho = 0.092 \), while the percentage of deletions positively correlated with these \( r = 0.444, \rho = 0.023 \) and \( r = 0.453, \rho = 0.018 \), respectively. This means that students who did better on the final exam, and, in addition had better content in their second essay, had a lower percentage of additions when revising their first draft, and a higher percentage of deletions. This makes sense because the first drafts for this assignment had better content already, so new content did not need to be added to make the essay correct. The percentage of revisions coded as good correlated strongly with student final exam grade \( r = 0.524, \rho = 0.006 \), their essay content grade \( r = 0.481, \rho = 0.011 \), their essay writing grade \( r = 0.374, \rho = 0.055 \), their total essay grade \( r = 0.568, \rho = 0.002 \) and the change in their grade from the first draft \( r = 0.475, \rho = 0.014 \). Similarly, the percentage of negative comments had significant negative correlations with these variables. Students who
had improvements in their essay grade from the first to the second draft had a lower percentage of sentence level revisions ($r = -0.430, \rho = 0.028$) and a higher percentage of multiple-sentence level revisions ($r = 0.460, \rho = 0.018$). If students initially had major problems with their first draft it would be consistent that they needed to make larger changes to make big improvements. However, this is inconsistent with the model that smaller revisions are indicative of more reflection.

In the fourth assignment, it was found that students who had high final exam grades had a more revision events ($r = 0.390, \rho = 0.036$), while surprisingly, the change from the first to the second draft grade negatively correlated with the number of events (though not significantly, with $r = -0.260, \rho = 0.181$). This time no strong correlations appeared with the percentage of revisions coded as good, though students with better writing and total essay grades had fewer revisions marked as bad ($r = -0.543, \rho = 0.002$, and $r = -0.407, \rho = 0.026$, respectively).

In the sixth essay assignment, it is again found that students had higher essay content, writing, and total grades if they had a higher percentage of revisions coded as good ($r = 0.572, \rho = 0.002$, $r = 418, \rho = 0.030$, and $r = 0.633, \rho = 0.000$, respectively). However the percent of “good” revisions did not correlate strongly with their final exam grade ($r = 0.275, \rho = 0.175$), suggesting that some writing behaviors are more strongly related to the essay grade than their overall course content knowledge.

In the eighth assignment, it was found that students with good final exam grades had a higher percentage of additions to their essays ($r = 0.423, \rho = 0.050$) but this actually negatively correlated (though not significantly) with their total essay grade ($r = -0.248, \rho = 0.254$). This sheds some light on the results reported in the previous section. Perhaps students with good course content knowledge felt changes needed to
be made to their first drafts, which were not needed. Since this content was probably more approachable, perhaps these students did not realize that their content was already adequate in their first drafts, and their changes actually did not help for their second draft. Interestingly, students with higher content and final exam grades had a higher percentage of revisions that were marked as bad ($r = 0.614, \rho = 0.002$ and $r = 0.485, \rho = 0.019$, respectively), also suggesting that their first drafts were good and they over-worked revisions for the second draft. Students with better content and total essay grades had a higher percentage of sentence level changes as well ($r = 0.536, \rho = 0.008$ and $r = 0.469, \rho = 0.024$, respectively). Students whose second draft score improved more from their first draft had a lower percentage of phrase level revisions ($r = -0.544, \rho = 0.007$), a higher percentage of revisions coded as good ($r = 0.540, \rho = 0.008$), and a higher percentage of revisions in the first quartile of their draft ($r = 0.488, \rho = 0.018$). This suggests that those students who needed to improve on their first draft made bulk changes early on, and that improved their grade.

In no week did the time spent editing correlate with the grades. Some students did not revise their draft in one sitting. This can cause a large false reading for the amount of time students spent on the draft. For example two half-hour editing sessions would look like 20 hours of editing if the student worked on it a second night. Even more surprisingly, in no week did the number of revision events correlate with the essay grades. In addition, even though students who did better on the final exam had more revision events for the first two revised drafts, they did not for the later two. These results are summarized in Table 7.3.
Looking at the three subgroups of students based on their final exam grade allows us to see if different types of students approached the revisions differently. In the second essay assignment, the percentage of revisions that were additions was significantly different between the groups ($\rho = 0.045$), with the low scoring students having more additions. As mentioned above, I suspect this is due to the fact those students did worse on the first essay assignment, and they needed to add new content to correct their first draft. Students with higher final exam grades had a higher percentage of revisions coded as good ($\rho = 0.025$), also consistent with the information found from the correlations. No other behavior was found to be significantly different among the three groups, suggesting that their revision behavior is only somewhat different.

In the fourth essay assignment, the number of revision events is almost significantly different among the three groups ($\rho = 0.056$). The students who scored low on the final exam averaged only 17.2 events, while those who scored in the medium range averaged 30.3 events and those in the high range averaged 32 events. A Mann-Whitney U test (the non-parametric equivalent of the t-test) shows a significant difference between the high and low scoring groups ($\rho = 0.022$). This may simply mean that students who work harder in the course realized a need to put more effort into the second drafts, but it is also consistent with the model that more advanced students will do more revising. There was also a significant difference between the groups on the percentage of revisions that were transitional ($\rho = 0.015$) and informational ($\rho = 0.030$). However, for neither of these was there a clear pattern, with medium-scoring students having more transitional revisions (and low-scoring having the least), and low-scoring students having more informational revisions (with medium-scoring having the least). This last fact is consistent with my hypothesis above that students
who initially did badly on the first draft needed to add a lot of content to fix their second drafts.

In the eighth essay assignment, only the percentage of revisions coded as good was significantly different between the three groups based on final exam grade ($\rho = 0.040$). Interestingly, this time the medium-scoring had the highest percentage of good changes, with 93%, while the high-scoring students had 80% and the low-scoring students had 72%. This is the same assignment mentioned above, in which the high-scoring students actually decreased their grade from the first to the second draft. It seems something about the essay content or perhaps the feedback given led these students to make some bad changes to their first drafts, which were on average fairly good. These results are summarized in Table 7.4.

### 7.4.3 Looking for Behavior Patterns

In addition to running statistical tests across all variables, it is useful to make hypotheses about certain behaviors and graphically look for trends based on them. One hypothesis is that students could be grouped based on the time it took them to write and the number of revision events. For example, students who do well in the course may have a longer time and more events, while those who don’t do well may have less time and fewer events. There may also be students who make a lot of revisions in a short time; those students probably make superficial revisions and are unlikely to have high essay grades. Other students may take a long time to do few revisions; those may be well thought out, and may have caused good changes from their first drafts.
To test these hypotheses, several plots were made. The first, Fig. 7.1, tests to see if there are any patterns in the number of events versus time for students who had different levels of course content knowledge as measured by the final exam. The plot shown is from the fourth essay assignment, and shows the clearest trends of the four revised essays. For each essay, the medium- and higher-scoring students had some data points above (more events) and to the right (more time) from the low-scoring students. In addition, those students who took a long time on their revisions had more events. (It should be noted that I cut off the x-axis at 1 h, because of the students who edited over the course of more than one sitting and had very long times such as 17 h.) Further, there is a clump of higher-scoring students who made more revisions in the same amount of time as the low-scoring students. In each plot, it is difficult to distinguish between the medium- and high-scoring students. To see these trends over all data, Fig. 7.2 and Fig. 7.3 show the number of events and the time on task (respectively) plotted against each second draft essay assignment. This also shows that these behaviors remained fairly steady throughout the quarter.

In addition to looking at students based on their final exam grade, we made a similar plot based on their essay grade, as well as one based on their change in grade from the first to the second draft. In order to compare students based on their essay grade across all essays, it was necessary to first normalize the essay grades so they were all out of the same total point value. Then students were ranked based on their detailed essay grade and divided into three categories: high, medium, and low scoring. These results are plotted in Fig. 7.4. Surprisingly, no clear trends appear in these data, indicating that the students’ overall course content knowledge has more to do with their revision habits than their essay grade. It is possible that combining data
<table>
<thead>
<tr>
<th>Essay #</th>
<th>Which Grouping</th>
<th>Significant Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essay 1</td>
<td>Final Exam Grade</td>
<td>Total Essay Grade</td>
</tr>
<tr>
<td>Essay 2</td>
<td>Final Exam Grade</td>
<td>Essay Content Grade</td>
</tr>
<tr>
<td>Essay 3</td>
<td>Final Exam Grade</td>
<td>Essay Content Grade, Essay Writing Grade, Total Essay Grade</td>
</tr>
<tr>
<td>Essay 4</td>
<td>Final Exam Grade</td>
<td>Essay Content Grade, Total Essay Grade, Change in Grade from First Draft</td>
</tr>
<tr>
<td>Essay 5</td>
<td>Final Exam Grade</td>
<td>Essay Writing Grade</td>
</tr>
<tr>
<td>Essay 6</td>
<td>Final Exam Grade</td>
<td>Essay Content Grade, Total Essay Grade</td>
</tr>
</tbody>
</table>

Table 7.2: Results from Revision Analysis on First Drafts for Spring quarter, 2006.

Figure 7.1: This shows the number of events versus time for students who scored high, medium, and low on their final exams.
<table>
<thead>
<tr>
<th>Essay #</th>
<th>Type of Result</th>
<th>Significant Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essay 2</td>
<td>Positive Correlations</td>
<td>Percentage of Deletions &amp; Final Exam Grade, and the Essay Content Grade; Percentage of “Good” Revisions &amp; Final Exam Grade, Essay Content Grade, Essay Writing Grade, Total Essay Grade, and the Change in Grade from the First Draft; Percentage of Sentence and Multiple-Sentence Revisions &amp; Change in Essay Grade</td>
</tr>
<tr>
<td>Essay 2</td>
<td>Negative Correlations</td>
<td>Percentage of Additions &amp; Final Exam Grade, and the Essay Content Grade; Percentage of “Bad” Revisions &amp; All Grades</td>
</tr>
<tr>
<td>Essay 4</td>
<td>Positive Correlations</td>
<td>Number of Revision Events &amp; Final Exam Grade</td>
</tr>
<tr>
<td>Essay 4</td>
<td>Negative Correlations</td>
<td>Number of Revision Events &amp; Change in Essay Grade; Percentage of “Bad” Revisions &amp; Essay Writing Grade, Total Essay Grade</td>
</tr>
<tr>
<td>Essay 6</td>
<td>Positive Correlations</td>
<td>Percentage of “Good” Revisions &amp; Essay Content Grade, Essay Writing Grade, and the Total Essay Grade</td>
</tr>
<tr>
<td>Essay 8</td>
<td>Positive Correlations</td>
<td>Percentage of Additions &amp; Final Exam Grade; Percentage of “Bad” Revisions &amp; Essay Content Grade, and the Final Exam Grade; Sentence Level Revisions &amp; Essay Content Grade, and the Total Essay Grade</td>
</tr>
<tr>
<td>Change in Grade</td>
<td>Positive Correlations</td>
<td>Percentage of “Good” Revisions; Percentage of Revisions 0-25% Into the Text</td>
</tr>
<tr>
<td>Change in Grade</td>
<td>Negative Correlations</td>
<td>Percentage of Phrase Level Revisions</td>
</tr>
</tbody>
</table>

Table 7.3: Results from Revision Analysis on First Drafts for Spring quarter, 2006.
<table>
<thead>
<tr>
<th>Essay #</th>
<th>Which Grouping</th>
<th>Significant Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Essay 2</td>
<td>Final Exam Grade</td>
<td>Percentage of Additions; Percentage of “Good” Revisions</td>
</tr>
<tr>
<td>Essay 4</td>
<td>2 Groupings Based on Final Exam Grade</td>
<td>Number of Revisions</td>
</tr>
<tr>
<td>Essay 4</td>
<td>Final Exam Grades</td>
<td>Percentage of Transitional and Informational Revisions</td>
</tr>
<tr>
<td>Essay 8</td>
<td>Final Exam Grade</td>
<td>Percentage of “Good” Revisions</td>
</tr>
</tbody>
</table>

Table 7.4: Results from Revision Analysis on First Drafts for Spring quarter, 2006.

![Number of Revisions Over Time](image)

Figure 7.2: This shows the number of events over each second draft for students who had high, medium, and low scores on their final exam.
Figure 7.3: This shows the time on task over each second draft for students who had high, medium, and low scores on their final exam.
Figure 7.4: This shows the number of events versus time for students who scored high, medium, and low on their essay.

from different assignments is causing us to lose any visual trends, so I also plotted the same thing for only assignment four (see Fig. 7.5). Although this graph is easier to read, it still does not show the clear trends that were seen when looking at these behaviors in terms of the students’ final exam grades.

Finally, we look for patterns of these behaviors in terms of the change in the students’ grade from the first to the second draft. Looking at Fig. 7.6, there is no clear trend, so we again look at only assignment four. Fig. 7.7 shows these results. It is interesting that it seems that many students who had good changes seemed to have fewer events and spent less time on their revisions than some of the students
Figure 7.5: This shows the number of events versus time for students who scored high, medium, and low on their fourth essay assignment.
Figure 7.6: This shows the number of events versus time for students who had high, medium, and low changes in their grade from the first to the second draft.

who did not make good changes from their first to their second draft. This is even more surprising than seeing no pattern at all.

Finally, we consider differences in the size of the revisions for various groupings of students. We hypothesize that all students will have typos, which result in changes to an existing word. These things were lumped into a broad category named grammar revisions; for example, if a verb tense was changed, an item was capitalized, or contractions were added or removed. If a single word was added or deleted, these are called 1 word changes. We hypothesize that all students will have a lot of such changes, but students who are putting more effort into the revisions will have more changes than those who are putting in less effort. For example, to keep existing content but to make
Figure 7.7: This shows the number of events versus time for students who had high, medium, and low changes in their grade from the first to the second draft.
Figure 7.8: This shows the number of events of each size for students who had high, medium, and low scores on their final exam.

It shorter it is necessary to rephrase things, causing word, phrase, and sentence-level changes. Students from both categories will have few multi-sentence level changes, largely due to the fact these essays are short, so there is not much bulk to be edited. These trends are observed in Fig. 7.8, which shows the number of events versus size for students with high, medium, and low final exam scores. We do not see an increase in sentence level revisions for the better scoring students, but these other predictions are confirmed.

Graphing this same information but grouping the students in terms of their (normalized) essay grade does not show the same trend. Fig. 7.9 shows that the students
who scored lower on the essays actually had more mid-size edits on average. This is another indicator that the students overall course content knowledge, as measured by their final exam grade, is a better predictor for their writing behavior than the actual quality of their essays. As was hinted above, this may be due to an effort level; although this class does have challenging content for many students, those who put in effort are likely to do well, while those who do not will not do well.

Also interesting is looking to see whether these behaviors change over time, so the number of events is again plotted versus the size of events, but this time they are grouped based on which essay assignment it is. This is shown in Fig. 7.10. The behaviors in essay assignment two are similar to number eight, but number four is
Figure 7.10: This shows the number of events of each size for each of the four second draft essay assignments.

noticeably different. In essay four, the behavior is similar to what was predicted for students who put in more effort. My guess here is that students initially did not know how to deal with the second drafts, and once they knew getting full points would not be easy they put more effort into their next second draft: essay assignment four. This difference in behavior might have been lost by week eight either due to students not wanting to put in so much time as the quarter progressed, or due to the frustrations students had with the essays, which were mentioned previously.
7.5 Conclusions

This quarter we had rough measures of student writing behaviors based only on second drafts. This allowed us to begin to test predictions for how students approached the writing, and which behaviors correlated with student conceptual understanding and student writing quality. While almost none of our predictions were strongly confirmed, we do see some suggestive trends in our data. For most weeks, students who had more conceptual understanding, and who had higher essay grades, made revisions that were coded as good changes. It would be shocking if this were not true, but at least this gives us some confidence that our coding is accurate. For many parameters we tested, there were stronger correlations with students final exam grade than with the detailed essay grades. This indicates that student behaviors may have more to do with the amount of effort they were putting in overall than with their actual writing quality. This is further supported by the fact that essay content correlated more with the final exam grades than the grade given to the aspects of the essay related to writing: overall correctness, ordering, and flow/transitions.

We see evidence that the writing can be thrown off by certain essay topics. In previous quarters, we found that if the content was difficult, the writing suffered. Here, however, we found that the easier essay topic threw off the students who did better on the first draft. It is as if they over-corrected when revising for their second draft. They felt that content had to be changed, and not just shortened, and often made poor revision choices. It is possible that some feedback I gave was confusing to them and prompted these mistakes. Although I can not recall what that may have been, when essays were initially very good I sometimes pushed them to consider a
more advanced idea, which may have been misleading (indeed, I can remember one instance where this was the case in Spring quarter).

It does not appear that any writing behaviors improved consistently through the quarter, as measured by the types of revisions made and the correlations with the content and writing portions of the essay grades. Because no writing instruction was given, besides minimal feedback on the essays, this is possibly not surprising. Some behaviors seemed to improve initially, which may be due to writing knowledge being activated (or students getting used to the assignment) but they either did not change, or reverted back to initial behaviors by the end of the quarter. It is hard to say how much the student resentment toward the essays affected this.

We did find that students with higher final exam scores on average spent more time on their revisions, and had more revision events. They also had more mid-sized revision events. If the final exam score is some measure of academic quality, this would support the idea that these students are behaving more like expert writers than novice writers. However, I think the more accurate statement is that these students are the ones who are willing to put more effort into their course work.

Overall it is hard to see strong trends in these data. It is possible that the revisions themselves don’t tell us as much as how the students wrote the first drafts. In the next study, we devised a method for capturing these details as well, to continue to look for ways to characterize and analyze student writing.
CHAPTER 8

A PROGRAM TO TRACK WRITING

For the study of writing, computer programs have been developed that track what students do while they write [6, 7, 27]. These programs make use of keystroke capture technology; each typed letter is captured, so words that do not make it into the final draft can still be seen. This provides an additional window into the writing process. JEdit is an early tracking program developed by Severinson-Eklundh, Kollberg, Cederlund, Kim, and Nilsson [6]. Though similar to our tracking needs, this program runs only on old Macintosh computer systems. A companion program, Trace-it, converts the output of JEdit into a readable archive of revisions. These programs were used to track differences in writing habits of people on four different genres [27], looking for differences such as how many edits, pauses, and the total time spent on essays. A more recent tracking program, Inputlog [7, 57] was written to be compatible with Microsoft Word. This program also meets many of our needs, but has limitations that require it to be used in a controlled setting. For example, students must come in to use the software on a computer whose settings are fixed, which not only limits the studies that can be done, but having students write in a controlled setting may affect the way they write. In addition, these tracking programs, when not used in a controlled setting, can cause unintended text to be captured; for example,
S-notation:
• I am writing a {short} text. It will [probably] be revised [somewhat] later. Now [I am] it is finished.

Final text:
• I am writing a short text. It will be revised later. Now it is finished.

Figure 8.1: Example of S-notation.

if a student is instant messaging a friend while writing the assignment, the personal messages may be captured.

The programs mentioned above are capable of producing S-notation (developed by Severinson-Eklundh and Kollberg) as an output. This notation shows insertions and deletions embedded in the final text, along with markers showing where in the text a student was writing when those revisions were made. This allows for the text to be read fairly fluently while seeing changes that were made as it was being written. S-notation uses the following symbols:

1. Breaks in writing are marked by a horizontal line and a sequential index

2. Insertions are marked by curly brackets and a sequential index

3. Deletions are marked by square brackets and a sequential index

A short example of S-notation is shown in Fig. 8.1.
At the Ohio State University, Dr. Lei Bao and the Physics Education Research Group has created its own set of programs for tracking the writing process. The first program, EditNav, is a text editor similar to notepad. There are two versions of this program. One can be downloaded and run directly from the hard drive of any PC. The other is kept on a server, and can be accessed by students online. This allows for a much easier, less invasive method of studying student writing. Students can log in to the program anytime to write their assignments in the environment in which they usually work, with no need to do the extra work of downloading anything. There is a database that is kept on the server allowing students to save files for later editing. In addition, neither program tracks unintended typing students might do in other programs open at the same time. Currently the downloadable version is slightly more stable; the cursor tracking has occasional bugs online. The web version of this program is shown in Fig. 8.2. In addition to providing a notepad for students to type their assignments, the program shows a running word count, which was useful for our assignments, and requests students to input their name, email address, their section, and which assignment is being submitted. This information is saved at the top of each text file.

When students submit their writing using the online program (or save their file using the downloaded version) two files are produced. One is the text file, as seen by the student. The other is a log file that contains a snapshot of the text each time a student pauses, backspaces, deletes, or moves the cursor. This file indexes each event, gives the time, what type of activity the student is doing, the text snapshot, and the cursor location. An example from a log file is:

1. 10:57:027 AM: Typing A circuit is all (CU)
Figure 8.2: Snapshot of the online version of EditNav.
The log file can then be converted either to S-notation, or to a data file that is useful for directly running statistical analysis on the revisions made. The S-notation we produce from this is modified somewhat from that mentioned above. Since we also track the pause times, we include the symbol (P#), where the number given is the length of the pause in seconds. S-notation is useful for reading the student writing and seeing the edits in context, but for the purpose of determining quickly where and what type of edits are made, the log file is more convenient. It includes an index of each event, the time, the number of words added or deleted (or the pause time), at which word in the essay the event took place, the total length (in number of words) of the essay up to that edit, as well as what text was added or deleted. Because the file is comma delimited, it can be easily imported into a program such as Microsoft Excel or SPSS for analysis. In the example below, the student deletes the e, adds more text, then fixes a typo.

1. 3:24:042 AM, A, 2, 68, 69, other e,

2. 3:24:042 AM, D, 1, 69, 68, , e

3. 3:24:049 AM, A, 5, 69, 73, terminal of the bulb (the boot,

4. 3:24:049 AM, D, 1, 74, 73, , ot
With this data set, it is easy to determine whether edits were done on recently typed text, for example, if a typo is immediately corrected, versus whether students are going back to previously written text and modifying it. It is also easy to see how often students are writing, editing, and pausing. In some assignments we gave, word limits were crucial. It is also easy to see from this output how much editing students do after they reach the word limit. It will soon also be possible to search for edits around keywords, for example, to see if students spend more time or do more revisions near certain content.
CHAPTER 9

SPRING 2006: TRACKING WRITING IN PHYSICS BY INQUIRY

During Spring quarter, 2006, we did an implementation of essays in Physics by Inquiry (PbI) similar to that in Winter quarter. There were again students participating from all three sections of PbI on the Ohio State University main campus, and the one section at the Marion campus. In Spring quarter, the course covered topics in electric circuits. The main difference between the writing activities Spring quarter compared to those from Winter quarter is that students wrote their essays using the program to track writing that was described in Chapter 8. This allows us to find ways to quantify behaviors in student writing and to look for differences in behaviors of different types of students and over time / practice. However, we are now able to look at how students wrote their first drafts; we are not limited to studying students’ second drafts.

9.1 Implementation Details

The student population is the same as described in the previous chapter with some students from both campuses. The course structure is also the same, with other writing in the form of journals and an emphasis on making explanations. Apart from
the use of the new tracking program, two factors aided our implementation in Spring quarter. First, the course content is more unified, making it easier to keep the essay assignments similar depth and difficulty through the quarter. Second, we were able to use our experiences from the previous quarter to minimize previous problems. For instance, we were able to address student concerns about how the essays were graded at the beginning, and minimize resistance to the grading. In addition, I had helped instruct the course Winter quarter, so I was better able to understand the students’ needs and expectations. We had also tested some of these essay topics in a Winter quarter section of this course at the Marion campus covering this same content. (Those data were not reported due to the small number of students completing those essays.)

9.1.1 Writing Activities

The writing activities were of the same structure as the Winter quarter implementation reported in Chapter 8. An example of an essay given Spring quarter is: “Explain clearly to an intelligent but uninformed friend why a circuit diagram of the wires in your house may be more useful than a diagram showing the layout of the wires. Include in your essay a description of how the process of redrawing the circuit can clarify the relationship between or among networks and might be related to the distinction between series and parallel circuits. Refer explicitly to in-class observations and/or activities if necessary. Pay close attention to writing a meaningful introductory and concluding sentence, as well as the transitions between sentences.”
Each essay prompt started with the same eight words establishing a uniform audience, and ended with the same two sentences to remind students of how to approach the essays.

The essays were again given in two drafts each, with the first needing to be close to 250 words and the second needing to be no more than 125 words. Because the quarter was only ten weeks long, there were only nine homework assignments, so only nine essay assignments were given. As a result the last essay consisted of only one draft, with students directly writing no more than 125 words.

9.1.2 Grading

The grading was done the same way it had been in Winter quarter, with two small exceptions. First, the students were given a sheet with their first assignment to explain the grading scheme, minimizing student frustration with grading and emphasizing the word limit importance from the start. Second, since the essays were submitted online, feedback was returned electronically, prompting me to give more feedback (I find typing easier than writing). I gave the same feedback as Winter quarter, avoiding pointing out specific locations or otherwise prompting students to know exactly what needed to be changed. In addition, I feel my general comments were more clear when typed.

9.2 Difficulties Experienced

During the first two weeks of the quarter, the word count function of the new tracking program was not working. Students resorted to writing their essays in another program (such as Microsoft Word) then copying the text and pasting it into the online tracking program to submit their essays. I did not penalize them for this
because the word count was so crucial and I did not want to hassle the students. However, even after this was fixed, some students refused to trust the online program, and we did not get tracking for them. Also, because there was not yet a database in place where students could save their essays, some preferred to write them elsewhere so they knew the essays were saved multiple times while writing. Also, without the database, the first drafts had to be emailed back to the students, then pasted back into the program for making the second draft. Due to these initial program limitations, we have a limited pool of tracked essays on which to base our study.

Just as we had seen in the Winter quarter implementation, some students started filling their first drafts with fluff, knowing that they had to shorten the essay for the second draft. Other students took notes before writing their essays, which would not yield the same writing behaviors as if they had started writing directly. As before, some students took multiple sittings to work on their essays, or took breaks, or were otherwise distracted such as watching TV while writing, placing limitations on any conclusions we might make.

9.3 Data Obtained

Of the students who completed most assignments using “Track Changes” and gave consent to participate in this research study, thirty were chosen for further analysis. Of the 114 students who completed the course, we chose to further study those who had written at least five of the essays assignments directly in the webpage, and had signed consent to participate in the research project; there were 37 students in this group. From these students, between 8 and 30 essays were available for further study for any specific assignment.
As in Winter quarter, students were categorized based on their final exam scores. These categories were also representative of students final course grades. For Spring quarter we chose to use four categories, allowing us to do more flexible analysis. Nine students were in the very low category, with the lowest final exam scores, nine were in the low category with slightly below average scores, ten were in the high category and nine were in the very high category. Because of the low number of data points for some weeks, however, some analysis was based on only two groups: above or below average on the final exam.

As in Winter quarter, for this subset of students, each of the nine essays was graded in detail based on the clarity and correctness of the items necessary in order to answer the question. Each item was given between zero and three points. Students were also given scores between zero and three based on the overall correctness of the essay, the logical order of the items include within the essay, and the introduction, conclusion, and transitions within the essay. This time, five important points were determined for each of the essays, eliminating the need to normalize essay scores for comparison. This gives a total maximum possible score of 24 points, depending on the essay topic.

Using the data from the new tracking program, each first and second draft was coded based on information captured by the log file. This included the location and time of all additions and deletions to a document. The location of the revision within the final text was recorded based on the number of words prior to the revision, as well as the time of each revision. The data included details about where and for how long students paused while writing. The word count of each revision was sorted into 5 categories: 1 word, 2-6 words, 7-12 words, 13-20 words, and greater than 20 words.
A revision of length 2-6 words is a small phrase, while a revision length of 7-12 words is a large phrase or short sentence. A revision length of 13-20 words is one to two sentences, and a revision of more than 20 words is considered a bulk edit.

The locations of the revisions is more interesting to consider for first drafts than second drafts. If students are writing new content at the end of their essays, it shows that they are not going back through their essay to review/revise them before the first submission. When students are cued by some new content, they are likely to go back and see whether that content flows with the rest of their essay, prompting earlier revisions. This type of writing behavior is consistent with the cognitive models of how students may learn from writing. The ability to track first draft writing gives us a valuable new dimension from which to analyze student behavior.

9.4 Results

As with the Winter quarter data, we have two dimensions across which to study the data: based on student rank, and across time. We again consider these data from the standpoint of correlations using Spearman’s correlation for non-parametric data due to the low numbers of essays available for coding. We will also again use the Kurskal-Wallis test and the Mann-Whitney U test (the non-parametric ANOVA and t-test, respectively) to check to see whether students from groupings based on the final exam exhibit statistically significant behaviors. We will also plot various combinations of data to look for differences in the chosen groupings, across the weeks, and see if clusters naturally form without assuming any previous groupings.

Under the knowledge that some students worked on their essays in more than one sitting, and some worked with distractions such as the TV and/or friends and family,
when analyzing pause times and total work times, I did not consider pause times more than three minutes or total work times of more than one hour. This will allow for more reasonable analysis based on these variables, since it will minimize times that were are long for reasons independent of writing behavior.

Another difference from the Winter quarter analysis is the ability to determine how much the students wrote after reaching the word limit. Because our tracking software captured a screen shot with each event, we know the number of words typed at each event. We therefore created a variable called “percentage word limit” that tells at what percentage (as measured by the number of events) the student reaches the word limit. For example, if this variable is 100%, that means the student’s last revision resulted in him or her reaching the word limit, and he or she made no further revisions; 50% would mean the student was only half way through the revision process when the word limit was reached as so spent considerably more effort after that point.

My first hypothesis was that students with better quality second drafts would have a lower percentage word limit. However, upon looking at students’ work, it became apparent that some students cut bulk out of their essays then went way below the word limit and had missing content, so that they then needed to do quite a bit more work to finish their essays. Other people made smaller changes, slowly converging on the word limit, but were essentially finished with their essay once they reached the word limit. It is therefore likely that a higher percentage word limit is actually indicative of more careful work.
9.4.1 Analysis of Detailed Essay Grades

For all first drafts (essay assignments 1, 3, 5, 7, and 9), the students’ final exam scores significantly correlated with their essay grades based on content ($r = 0.242, \rho = 0.009$), based on the writing ($r = 0.227, \rho = 0.015$), and based on the total essay grade ($r = 0.282, \rho = 0.002$). However, overall the content scores did not have a significant correlation with the writing scores ($r = 0.160, \rho = 0.088$). For the second drafts (essay assignments 2, 4, 6, and 8), there are again significant correlations between students final exam scores and their essay grade based on content ($r = 0.223, \rho = 0.040$) and their total essay grade ($r = 0.298, \rho = 0.005$). However, there is not a significant correlation with their grade based on writing ($r = 0.187, \rho = 0.084$). Similar to what was observed the previous quarter, the correlation between the change in grade from the first to the second draft and the student’s final exam score is actually negative, but not significant ($r = -0.088, \rho = 0.491$). Once again the grade based on writing and the grade based on content do not correlate ($r = 0.045, \rho = 0.681$).

Now we look at the individual essay assignments more closely. For the first assignment, we see results that differ from the previous quarter. Neither the essay grade based on content or writing correlate with the students final exam scores ($r = -0.060, \rho = 0.888$, and $r = 0.147, \rho = 0.728$, respectively). Nor do the writing and content scores correlate with each other significantly ($r = 0.574, \rho = 0.137$). Note that here there is a strong positive correlation between the writing and content scores, but it is not significant, probably due to the low number of data points; only eight essays from the first assignment were available for analysis due to the problems mentioned above.
For the second drafts of this essay (the second essay assignment) there were also only eight available essays for analysis, so it is again difficult to have significant results. Both the essay grade based on content and writing correlated negatively with the student’s final exam grade \( r = -0.224, \rho = 0.718, \) and \( r = -0.123, \rho = 0.816, \) respectively), but the total essay grade has a small positive correlation \( r = 0.166, \rho = 0.827, \) which perhaps an example of the randomness associated with such small numbers of variables. The grades based on content and on writing showed a large positive correlation that was close to significant \( r = 0.825, \rho = 0.086. \) Correlations between the change in score from the first to the second draft and the final or essay grades are meaningless since there were only two essay data points.

For the third essay assignment, which is the first draft of the second essay topic, nearly significant positive correlations were seen between the grade based on content and the final exam grade \( r = 0.422, \rho = 0.056, \) but no correlation was seen between the final and the writing grade \( r = 0.021, \rho = 0.929. \) There was also no correlation between the content and the writing grades \( r = 0.152, \rho = 0.511. \) For the fourth essay (the second draft of this same topic), there was no correlation between the final exam score and either the content or writing grade \( r = 0.035, \rho = 0.861, \) and \( r = 0.090, \rho = 0.654, \) respectively) nor between the content total and the writing total \( r = 0.049, \rho = 0.808. \) There were positive, but not significant, correlations between the change in scores from the first to the second draft and the final exam score \( r = 0.231, \rho = 0.407 \) and the writing score \( r = 0.319, \rho = 0.246, \) but not the content score \( r = -0.020, \rho = 0.945. \)

For the fifth essay assignment (the first draft of the third essay topic), a significant positive correlation was seen between the final exam score and the essay grade based
on writing ($r = 0.512, \rho = 0.005$) but not that based on content ($r = 0.232, \rho = 0.225$). The content and writing scores also had a strong positive correlation ($r = 0.548, \rho = 0.002$). For the second drafts of this topic (the sixth assignment), there was again no correlation between the final exam score and either the content score or the writing score ($r = 0.156, \rho = 0.456$, and $r = 0.088, \rho = 0.674$, respectively). The content and writing scores did positively correlate, but not significantly ($r = 0.294, \rho = 0.153$). The change in score from the first to the second draft negatively correlated with the final exam score ($r = -0.223, \rho = 0.319$), but had a strong positive correlation with the content score ($r = 0.497, \rho = 0.019$) and a positive correlation with the writing score ($r = 0.237, \rho = 0.288$).

For the seventh essay assignment (which was the first draft of the fourth essay topic) there were significant positive correlations between the final exam grade and the essay grade based on content ($r = 0.410, \rho = 0.024$) and positive but not significant between the final and the writing grade ($r = 0.309, \rho = 0.097$). There were strong positive correlations between the writing and the content grade ($r = 0.629, 0.000$). For the second draft of this essay (the eighth assignment) there were also strong significant positive correlations between the final and the content grade ($r = 0.319, \rho = 0.040$) and the final and the writing grade ($r = 0.383, \rho = 0.044$). There were positive, but not significant, correlations between the writing and content grades ($r = 0.239, \rho = 0.221$). There were no correlations between the change in grade from the first to the second draft and the final or any of the essay grades.

For the ninth essay assignment (the only draft of the fifth essay topic), there were no significant correlations between the final exam grade and either the content or
writing grade \( r = 0.024, \rho = 0.908, \) and \( r = 0.269, \rho = 0.184, \) respectively). However, the content and writing grade were significantly correlated \( r = 0.581, \rho = 0.002).\)

The results from each of these tests is summarized in Table 9.1. One trend that can be seen that is consistent with both the Summer and Winter quarter data is that with the first drafts, the correlation between the content and the writing grade improved with practice. This time we saw few, and inconsistent, significant correlations between the final exam grade and either the content or writing grades (though overall there were positive correlations with both). This indicates that students who showed strong course content knowledge (based on their final exam) did not necessarily do better on either the content or writing of the essays, nor did there seem to be any steady change in that through the quarter.

For the second drafts, correlations between the final and both the content and writing grade improved through the quarter, suggesting that either the students with stronger content knowledge did better on the second drafts with practice. Almost no correlations were found between the final or any of the essay grades and the change from the first to the second draft, suggesting that the quality of the first draft, or the students overall content knowledge did not determine how well they were able to improve their work in the second draft. However, averaged over the whole quarter, those with higher writing and total essay grades did improve their drafts more, but not those with higher content grades. This is a suggestive result, which will hopefully be more clear with analysis of revision behaviors.

The other way to look at this data is using the Mann-Whitney U and Kurskal-Wallis tests. For the first drafts, averaged over the entire quarter, it is interesting that when grouping the students into above and below average on the final exam, no
Table 9.1: Correlation coefficients for Spring 2006 essay grades. Significance levels are in parenthesis, and those marked with an asterisk are ρ < 0.05. Note that no significant differences are given for correlations with the change in grade from assignment two because only two data points were available for that analysis.
significant differences are found between the essay content, writing, or total grades ($\rho = 0.067, 0.305, \text{ and } 0.074$, respectively). However, when grouping them into four categories, there are significant differences ($\rho = 0.013, 0.068, \text{ and } 0.004$, respectively).

When looking at the average scores for these groupings it becomes apparent why this is true. In many cases, the students who did a little below average on the final had similar grades to those who did a little above average, however those who did well below average did much worse than the other groups. For example, the total essay grade average for the very low group was 16.24, but it was 18.61 for the low group, 19.09 for the high group, and 19.14 for the very high group. No significant differences are found with either grouping for the individual first draft assignments, possibly in some cases due to the low number of data points.

For the second drafts, the only significant difference is seen with four groupings based on the final exam score, when looking at the average total essay grades through the quarter ($\rho = 0.032$). These averages are very similar to those of the first drafts, with the very low final exam students averaging 16.83, the low-scoring students averaging 18.85, the high-scoring students averaging 18.75, and the very-high-scoring students averaging 19.60. This again shows a marked difference between the very low scoring students and the other students. In the eighth essay assignment, there is also a significant difference between the four groupings based on final exam grade on their total essay grade ($\rho = 0.034$). This time the very low scoring students averaged 18, the low scoring students averaged 19.2, the high scoring students averaged 18.83, and the very high scoring students averaged 20. Again the students scoring just below and above average did similarly, this time with those scoring very high and very low standing out as different, though not as markedly as before. No significant differences
were seen between any groupings based on final exam and the change in scores from the first and second draft, as could be expected from the results given above. These results are shown in Table 9.2.

<table>
<thead>
<tr>
<th>Essay #</th>
<th>Which Grouping</th>
<th>Significant Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>All First Drafts</td>
<td>4 Groups Based on Final Exam</td>
<td>Essay grade based on content, Essay grade based on writing, Total essay grade</td>
</tr>
<tr>
<td>All Second Drafts</td>
<td>4 Groups Based on Final Exam</td>
<td>Total essay grade</td>
</tr>
<tr>
<td>Essay 8</td>
<td>4 Groups Based on Final Exam</td>
<td>Total essay grade</td>
</tr>
</tbody>
</table>


9.4.2 Analysis of Revision Behaviors

Now we analyze the results of these tests on revision behaviors. For all of the first drafts, the number of events positively correlated with both the essay grade based on content and the final essay grade ($r = 0.308, \rho = 0.001$, and $r = 0.279, \rho = 0.003$, respectively). These grades also significantly correlated with the number of pauses ($r = 0.202, \rho = 0.031$ and $r = 0.207, \rho = 0.027$, respectively) and the percentage word limit ($r = 0.193, \rho = 0.050$ and $r = 0.205, \rho = 0.037$, respectively). Students’ time on task did not correlate with the final exam or the essay grades except for the grade based on content ($r = 0.221, \rho = 0.023$). The size and location of the revisions did not correlate with any of the grades.
Looking at the individual weeks will give us an idea whether any of these correlations changed through the quarter. In the first assignment, the final exam grade correlated positively with percentage of revisions 2-6 words in length \( (r = 0.833, \rho = 0.010) \), but negatively with percentage of revisions 7-12 words in length \( (r = -0.929, \rho = 0.001) \). The students with higher final exam grades made more smaller changes while writing their first drafts.

In the third essay assignment, however, the essay grade based on content correlated positively with the percentage of revisions 7-12 words in length \( (r = 0.452, \rho = 0.040) \) and the percentage of revisions 13-20 words in length \( (r = 0.458, \rho = 0.037) \). Students who did well on the essay were making large changes while they wrote the first draft. In the fifth essay assignment, the only significant correlations were in the percentage of revisions occurring between 50-75% of the way through the essays. This correlated positively with the essay grade based on content \( (r = 0.384, \rho = 0.040) \), the essay grade based on writing \( (r = 0.403, \rho = 0.030) \) and the total essay grade \( (r = 0.459, \rho = 0.012) \). This is indication that students who were looking back at what they had recently written and revising it while they write did better on the essay.

There were no significant correlations between the final exam grade or any of the essay grades and any of the revision behaviors for the seventh essay assignment. For the ninth essay assignment, the essay grade based on writing correlated positively with the number of pauses \( (r = 0.406, \rho = 0.040) \), and the essay grade based on content as well as the total essay grade correlated positively with the percentage word limit \( (r = 0.444, \rho = 0.030, \text{ and } r = 0.460, \rho = 0.024, \text{ respectively}) \). The essay grade based
on writing also correlated positively with the percentage of revisions made from 25-50% of the way through the essay \( r = 0.419, \rho = 0.033 \), again indicating that people who looked back at what they had already written during the composing process had better essays overall.

<table>
<thead>
<tr>
<th>Essay #</th>
<th>Type of Result</th>
<th>Significant Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>All First Drafts</td>
<td>Positive Correlations</td>
<td>Number of Revision Events &amp; Essay Grade Based on Content, and the Total Essay Grade; Number of Pauses &amp; Essay Content Grade, and the Total Essay Grade; Percentage Word Limit &amp; Essay Content Grade, and the Total Essay Grade; Time on Task &amp; Essay Content Grade</td>
</tr>
<tr>
<td>Essay 1</td>
<td>Positive Correlations</td>
<td>Percentage of Revisions 2-6 Words in Length &amp; Final Exam Grade;</td>
</tr>
<tr>
<td>Essay 1</td>
<td>Negative Correlations</td>
<td>Percentage of Revisions 7-12 Words in Length &amp; Final Exam Grade</td>
</tr>
<tr>
<td>Essay 3</td>
<td>Positive Correlations</td>
<td>Percentage of Revisions 7-12 Words in Length &amp; Essay Content Grade; Percentage of Revisions 13-20 Words in Length &amp; Essay Content Grade</td>
</tr>
<tr>
<td>Essay 5</td>
<td>Positive Correlations</td>
<td>Percentage of Revisions 50-75% Into the Text &amp; Essay Content Grade, Essay Writing Grade, and the Total Essay Grade</td>
</tr>
<tr>
<td>Essay 9</td>
<td>Positive Correlations</td>
<td>Number of Pauses &amp; Essay Writing Grade; Percentage Word Limit &amp; Essay Content Grade, and the Total Essay Grade; Percentage of Revisions 25-50% Into the Text &amp; Essay Writing Grade</td>
</tr>
</tbody>
</table>

Table 9.3: Results from Revision Analysis on First Drafts for Spring quarter, 2006.

The results for the first draft are summarized in Table 9.3. They are interesting because in general there were no significant correlations based on the size or location of the revisions. However, looking weekly, correlations between making revisions in previously written text and the quality of the essays began to appear. However,
correlations we would expect, such as with the number of revision events and pauses only appeared when looking at all essays, but not looking weekly. This may be because this result is weaker, so having more data made the results significant. The variable “percent word limit” would be lower if students continued to work on the essay after reaching the word limit. The hypothesis is that better students would go back and review/revise the existing text after reaching the word limit. However, a correlation between the percent word limit and the essay grades did not consistently appear in the data.

For the second drafts, the final exam grade, the essay grade based on content and the total essay grade had significant positive correlations with the time on task ($r = 0.333, \rho = 0.002, r = 0.353, \rho = 0.001$, and $r = 0.305, \rho = 0.005$, respectively), the number of events ($r = 0.390, \rho = 0.000, r = 0.347, \rho = 0.001$, and $r = 0.241, \rho = 0.025$, respectively), and the number of pauses ($r = 0.254, \rho = 0.018, r = 0.306, \rho = 0.004$, and $r = 0.222, \rho = 0.040$, respectively). These three variables also negatively correlate the percentage of edits larger than 20 words ($r = -0.345, \rho = 0.001, r = -0.391, \rho = 0.000$, and $r = -0.251, \rho = 0.020$, respectively). The percentage of revisions between two and six words in length positively correlated with the students’ final exam grades and with their essay grades based on content ($r = 0.316, \rho = 0.003$, and $r = 0.240, \rho = 0.027$, respectively). These results are consistent with predictions one might have made based on the models for how experienced students approach revisions. However, the essay grade based on the writing and the change in grade from the first to the second draft did not correlate significantly with any of these variables.
The essay grades based on writing correlated negatively with the percentage of revisions occurring between 50-75% of the way through the text \((r = -0.249, \rho = 0.021)\) and positively with the percentage of revisions made at the end of the text \((r = 0.247, \rho = 0.022)\). This means that students who were adding new content at the end instead of revising existing text had better writing grades, which is an unexpected result based on the cognitive theories. The total essay grade also positively correlated with percentage of revisions made at the end \((r = 0.226, \rho = 0.037)\).

In the second essay assignment there was a significant negative correlation between the percentage of changes on single words and the essay grade based on content \((r = -0.894, \rho = 0.041)\). The essay grade based on writing also negatively correlated, but not significantly. This means students with better second drafts made a lower percentage of word-choice, grammar, and spelling revisions. There were also negative correlations with essay grades and the percentage of revisions occurring between 25-50% of the way through the essay. However, this was only significant for the essay grade based on writing \((r = -0.926, \rho = 0.008)\). The essay grade based on content and the total essay grades correlated positively with the percentage of revisions occurring between 75-99% of the way through the essay \((r = 0.894, \rho = 0.041, \text{ and } r = 0.986, \rho = 0.000)\) and at the end of the essay \((r = 0.894, \rho = 0.041, \text{ and } r = 0.882, \rho = 0.020, \text{ respectively})\). The essay grade based on writing also positively correlated with these, but not significantly. This indicates that students who did better on their second drafts made more changes at the end of their essays.

For the fourth essay assignment, the percentage of revisions greater than 20 words in length correlated negatively with the essay grade based on content \((r = -0.446, \rho = 0.020)\) but positively on the essay grade based on writing \((r = 0.471, \rho = 0.013)\). This
indicates that students who made large changes may have gotten rid of important content from their first drafts, while at the same time improving the flow of the writing. The percentage of revisions occurring at the end of the the essay correlated positively with the grade based on writing \((r = 0.585, \rho = 0.001)\) and the total essay grade \((r = 0.490, \rho = 0.009)\). For the second drafts, revisions occurring at the end include getting rid of content and adding new content, so although this does not confirm the idea that students who make revisions throughout their essays will do better, it does indicate the content choices made by the students were good.

For the sixth essay assignment, the students final exam grades correlated negatively with the percentage of revisions greater than 20 words in length \((r = -0.399, \rho = 0.048)\) and positively with the percentage of revisions occurring between 25-50% of the way through the essays \((r = 0.505, \rho = 0.010)\). If students who do well on the final exam are more expert-like in their writing, these results would be consistent with the writing research. However, we do not know if that is true, and previous weeks do not confirm this result. In addition, as before, the total essay grade correlated positively with the percentage of revisions made at the end of the essay \((r = 0.397, \rho = 0.049)\).

In the eighth essay assignment, there were more significant correlations. The only significant result seen with the time on task appears here; positively correlating with the final exam grade \((r = 0.424, \rho = 0.027)\). The final exam grade, the essay grade based on content, and the total essay grades all correlated positively with the number of revision events \((r = 0.461, \rho = 0.013, r = 0.510, \rho = 0.006, \text{ and } r = 0.443, \rho = 0.018, \text{ respectively})\). The number of pauses also positively correlates with the total essay grade \((r = 0.398, \rho = 0.036)\). The percentage of revisions between 2-6 words in length positively correlated with the students final exam grade \((r = 0.538, \rho = 0.003)\),
and also with the essay grade based on content, but not significantly. In contrast, the percentage of revisions greater than 20 words in length negatively correlated with the students final exam grade, the essay grade based on content and the total essay grade ($r = -0.469, \rho = 0.012, r = -0.587, \rho = 0.001$, and $r = -0.561, \rho = 0.002$). These results are consistent with predictions based on writing research.

The results from the second drafts are summarized in Table 9.4. In none of the weeks did the change in grade from the first to the second week correlate significantly with any of the revision behaviors. In the last two second drafts, more behaviors that are consistent with the models from the writing literature are observed. It is not possible to state exactly why the results are different for these weeks compared to the situation for the first two second drafts. It could be due to the content; however, it could also be due to behavior changes as the quarter progressed.

Significant differences based on grouping the students by their final exam grades (as done in the previous section) can give us more information on differences in the revision behaviors of the students based on overall course content knowledge. For all first drafts, when students are grouped as being above or below average on the final, there is a significant difference between the percentage of revisions between 7-12 words in length ($\rho = 0.027$, with students below average having more) and the percentage of revisions occurring between 75-99% of the way through the text ($\rho = 0.031$, with students below average having more).

When grouping the students into four groups by the final exam grade, the time on task was significantly different ($\rho = 0.013$), the number of pauses ($\rho = 0.000$), the percentage of revisions between 50-75% through the text ($\rho = 0.017$), 75-99% through the text ($\rho = 0.004$), and occurring at the end of the text ($\rho = 0.018$). As
<table>
<thead>
<tr>
<th>Essay #</th>
<th>Type of Result</th>
<th>Significant Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Second Drafts</td>
<td>Positive Correlations</td>
<td>Time on Task, Number of Events, and the Number of Pauses &amp; Final Exam Grade, Essay Content Grade, and the Total Essay grade; Percentage of Revisions 2-6 Words in Length &amp; Final Exam Grade, and the Essay Content Grade; Percentage of Revisions at End of Text &amp; Essay Writing Grade, and the Total Essay Grade</td>
</tr>
<tr>
<td>All Second Drafts</td>
<td>Negative Correlations</td>
<td>Percentage of Edits Over 20 Words &amp; Final Exam Grade, Essay Content Grade, and the Total Essay Grade; Percentage of Revisions 50-75% Into the Text &amp; Essay Writing Grade</td>
</tr>
<tr>
<td>Essay 2</td>
<td>Positive Correlations</td>
<td>Percentage of Revisions 75-99% Into the Text, and Percent of Revisions at the End of the Text &amp; Essay Content Grade, and the Total Essay Grade</td>
</tr>
<tr>
<td>Essay 2</td>
<td>Negative Correlations</td>
<td>Percentage of Revisions 1 Word in Length &amp; Essay Content Grade; Percentage of Revisions 25-50% Into the Text &amp; Total Essay Grade, and the Essay Writing Grade</td>
</tr>
<tr>
<td>Essay 4</td>
<td>Positive Correlations</td>
<td>Percentage of Revisions More than 20 Words &amp; Essay Writing Grade; Percentage of Revisions at the End of the Text &amp; Essay Writing Grade, and the Total Essay Grade</td>
</tr>
<tr>
<td>Essay 4</td>
<td>Negative Correlations</td>
<td>Percentage of Revisions More than 20 Words &amp; Essay Content Grade</td>
</tr>
<tr>
<td>Essay 6</td>
<td>Positive Correlations</td>
<td>Percentage of Revisions 25-50% Into the Text &amp; Final Exam Grade; Percentage of Revisions at the End of the Text &amp; Total Essay Grade</td>
</tr>
<tr>
<td>Essay 6</td>
<td>Negative Correlations</td>
<td>Percentage of Revisions More than 20 Words &amp; Final Exam Grade</td>
</tr>
<tr>
<td>Essay 8</td>
<td>Positive Correlations</td>
<td>Time on Task &amp; Final Exam Grade; Number of Revision Events &amp; Final Exam Grade, Essay Content Grade, and the Total Essay Grade; Number of Pauses &amp; Total Essay Grade; Percentage of Revisions 2-6 Words in Length &amp; Final Exam Grade</td>
</tr>
<tr>
<td>Essay 8</td>
<td>Negative Correlations</td>
<td>Percentage of Revisions More than 20 Words &amp; Final Exam Grade, Essay Content Grade, and the Total Essay Grade</td>
</tr>
<tr>
<td>Change</td>
<td>Correlations</td>
<td>None</td>
</tr>
</tbody>
</table>

Table 9.4: Results from Revision Analysis on Second Drafts for Spring quarter, 2006.
in the previous section, it is seen that students just below and just above the average final exam score exhibit similar behavior, which is probably why these did not show up as significantly different when looking only at two groupings based on the final.

For instance, with time on task, students well below the final exam average spent 12.4 minutes on their first drafts, and students well above average spent 15.1 minutes. However, students just below average spent 20.6 minutes and students just above average spent 20.5 minutes. This shows that students who did very well on the final did spend longer on the first drafts, just not as long as those who were near average. A possible explanation for shorter time on task is that students who did very poorly on the average may put less effort into the course. Those who are near average may be trying hard but struggling, motivating them to spend longer on the essays. The results from the number of pauses is similar (very low averaging 5.0, low averaging 12.0, high averaging 9.8, and very high averaging 6.5). Looking at where the revisions occur within the text does not show as clear behavior as the time and number of pauses.

For the first essay assignment, due to the low number of essays it was only possible to evaluate the students based on their being above or below the final exam. The percentage of revisions between 2-6 words in length and between 7-12 words in length were significantly different ($\rho = 0.029$ for both). Students with higher essay grades had more revisions between 2-6 words in length and fewer between 7-12 words in length.

For the third essay assignment, grouping the students into two groups based on the final exam yielded significant differences on the number of revision events ($\rho = 0.036$). Students with lower essay grades had an average of 163.8 events while those with
above average grades had 120.6 events. There were no significant differences found compared to grouping the students into four groups based on the final exam grade. For the fifth essay assignment, no significant differences were found with either way of grouping the students by their final exam score.

For the seventh essay assignment, the percentage of revisions one word in length was found to be significantly different based on grouping students into above and below average on the final ($\rho = 0.038$). Students with higher essay grades had more than those with lower essay grades. When grouping students into four categories based on the final, the percentage of revisions occurring between 50-75% of the way through the text and the percentage of revisions occurring at the end of the text were significantly different ($\rho = 0.039$, and $\rho = 0.026$, respectively). Again, there is no consistency in the four groups on these results from the averages of these variables. In the ninth essay assignment, no significant differences were found with either way of grouping students by their final exam.

When grouping students into two categories based on the final exam score, the second drafts are significantly different for time on task ($\rho = 0.029$, with higher scoring students spending more time), the number of events ($\rho = 0.007$, with higher scoring students having more), the percentage of revisions between 2-6 words in length ($\rho = 0.002$, with higher scoring students having more), the percentage of revisions greater than 20 words in length ($\rho = 0.009$, with lower scoring students having more), the percentage of revisions between 25-50% of the way through the text ($\rho = 0.030$, with higher scoring students having more) and almost significant for the percentage word limit (0.056, with higher scoring students having a higher percentage). These results are consistent with those found above based on correlations.
When grouping the students into four categories based on the final exam, significant differences were found for the time on task ($\rho = 0.016$), the number of events ($\rho = 0.001$), the number of pauses ($\rho = 0.043$), and the percentage of revisions between 2-6 words in length ($\rho = 0.020$). For time on task, students above and below average again did similarly, but this time students who had very high final exams spent the most time (average 11.8 minutes for very low scoring students, 15.5 for low scoring, 13.1 for high scoring, and 21.2 for very high scoring students.) For the number of events, very high scoring students again had far more than the other groups (very low averaging 80.0, low averaging 63.2, high averaging 80.4, and very high averaging 120.6). No clear trends were seen in the number of pauses or the percentage of changes 2-6 words in length.

For the second essay assignment, no significant differences were found (possibly due to the lower number of data points). When grouping the students as above or below average on the fourth essay assignment, significant differences were found for the percentage of revisions between 2-6 words in length ($\rho = 0.016$), with higher scoring students having more. When put into four groups based on the final exam, no significant differences were found.

For the sixth essay assignment, the percentage of revisions made from 25-75% of the way through the text was significantly different when the students were grouped as above or below average ($\rho = 0.046$) with higher scoring students having more. When grouped into four groups, no significant differences were found. For the eighth essay assignment, the percentage of revisions from 2-6 words in length was found to be significantly different when students were grouped as above or below average ($\rho = 0.010$), with higher scoring students having more. When grouped into four
<table>
<thead>
<tr>
<th>Essay #</th>
<th>Which Grouping</th>
<th>Significant Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>All First Drafts</td>
<td>2 Groups Based on Final Exam</td>
<td>Percentage of Revisions 7-12 Words in Length</td>
</tr>
<tr>
<td>All First Drafts</td>
<td>4 Groups Based on Final Exam</td>
<td>Time on Task, Number of Pauses, Percentage of Revisions 50-75% Into the Text, Percentage of Revisions at the End of the Text</td>
</tr>
<tr>
<td>Essay 1</td>
<td>2 Groups Based on Final Exam</td>
<td>Percentage of Revisions Between 2-6 Words in Length and Between 7-12 Words in Length</td>
</tr>
<tr>
<td>Essay 3</td>
<td>2 Groups Based on Final Exam</td>
<td>Number of Revisions</td>
</tr>
<tr>
<td>Essay 7</td>
<td>2 Groups Based on Final Exam</td>
<td>Percentage of Revisions One Word in Length</td>
</tr>
<tr>
<td>Essay 7</td>
<td>4 Groups Based on Final Exam</td>
<td>Percentage of Revisions 50-75% Into the Text and at the End of the Text</td>
</tr>
<tr>
<td>All Second Drafts</td>
<td>2 Groups Based on Final Exam</td>
<td>Time on Task, Number of Events, Percentage of Revisions 2-6 Words in Length and More than 20 Words, Percentage of Revisions Between 25-50% Into the Text</td>
</tr>
<tr>
<td>All Second Drafts</td>
<td>4 Groups Based on Final Exam</td>
<td>Time on Task, Number of Events, Number of Pauses, Percentage of Revisions 2-6 Words in Length</td>
</tr>
<tr>
<td>Essay 4</td>
<td>2 Groups Based on Essay Grade</td>
<td>Percentage of Revisions 2-6 Words in Length</td>
</tr>
<tr>
<td>Essay 6</td>
<td>2 Groups Based on Final Exam</td>
<td>Percentage of Revisions 35-50% Into the Text</td>
</tr>
<tr>
<td>Essay 8</td>
<td>2 Groups Based on Final Exam</td>
<td>Percentage of Revisions 2-6 Words in Length</td>
</tr>
<tr>
<td>Essay 8</td>
<td>4 Groups Based on Final Exam</td>
<td>Number of Events</td>
</tr>
</tbody>
</table>

Table 9.5: Results from Revision Analysis on First Drafts for Spring quarter, 2006.
categories based on the finals, the number of events is significantly different ($\rho = 0.036$), with very low scoring students averaging 76.0, low scoring students averaging 53.8, high scoring students averaging 95.3, and very high scoring students averaging 138.4. The time on task is not quite significantly different ($\rho = 0.067$) but interesting, with very low students averaging 11.5 minutes, low scoring students averaging 11.2 minutes, high scoring students averaging 13.5 minutes and very high scoring students averaging 24.2 minutes. These results are summarized in Table 9.5.

### 9.4.3 Looking for Behavior Patterns

Instead of looking to see if behaviors that can be predicted based on the writing literature are present in our data, we now shift toward looking graphically at our data to see if we find behavior patterns emerging, as in the previous chapter. First, we look for trends in the number of events versus the time on task for the first drafts. Fig. 9.1 shows these data for all first drafts. Notice that a considerable number of average students are spending a long time on task, with students just below average final exam grades having fewer events than those who scored just above average. Also notice that there are no very low scoring students who spent a long time on task. The other interesting thing is the predominance of above average students who had more than 200 revision events when writing their first drafts. Almost no below average students appear above that line. These data were also graphed individually for each first draft essay assignment. There are no additional trends when looking these individual graphs.

As with the Winter quarter data, we plot the number of events versus the time through the quarter (based on which essay assignment it is) in Fig. 9.2, as well
Figure 9.1: This shows the number of events versus the time on task for all first drafts. The data are grouped based on the students’ final exam scores.
as looking at the time on task versus time in Fig. 9.3. These two plots confirm what the previous plot showed, and also indicate that these trends did not vary from week to week, nor change through the quarter. The students scoring very low on the final exam are at the lower range for both the number of events and the time on task for each essay assignment. In addition, each week it is the mid range scoring students who have the most time on task. The only trend that is not as steady as that above is that the higher-scoring students do not dominate the number of events for assignments three and five. It is also interesting to note that the number of events for week nine is less than the other weeks; recall this is the assignment for which they were to have no more than 125 words on their first draft, so it was a shorter essay than the others. The time on task is also lower for this essay.

Now we consider the second drafts. Fig. 9.4 shows the number of events versus time on task for all second drafts, again with the students grouped by their final exam scores. This shows similar trends to the first draft data, with most students clumped at shorter times and numbers of events, and a roughly linear correlation between those two variables. Again there are more higher-scoring students with more numbers of revision events, however, the average-scoring students do not dominate the region indicating more time spent on the essays. Few very low-scoring students spend long time on the essays. We see fewer students here spending a short time but having a lot of revision events, as compared to the first drafts. Correcting typos is also considered a revision event, so it is possible that when writing first drafts, students fix a lot of things such as spelling corrections as they type, whereas these immediate events occur less frequently for second drafts, because students are generating less new text.
Figure 9.2: This shows the number of revision events for each of week of the first draft essay assignments. The data are grouped based on the students final exam scores.
Figure 9.3: This shows how much time on task was spent for each of week of the first draft essay assignments. The data are grouped based on the students’ final exam scores.
Figure 9.4: This shows the number of events versus the time on task for all second drafts. The data are grouped based on the students’ final exam scores.
These data are plotted separately, showing the number of events on each second draft essay assignment through the quarter, and the time on task through the quarter shown similarly. Fig. 9.2 shows the number of events across the quarter. It is interesting that in weeks four and eight higher-scoring students have more events, but this result is not as strong for weeks two and six. Very low scoring students almost always dominate the lower number of events, however. It is also interesting that the number of events increased after the second assignment. Perhaps this indicates students realized the need to take the second drafts more seriously after receiving their first feedback. Fig. 9.3 shows the time on task for each of these second drafts. Average-scoring students spent more time on essays four and six, but higher-scoring students spent more time on week eight. The weekly behavior for the second drafts is not as consistent as it is for the first drafts.

The first and second drafts are again graphed looking for behavior patterns between the number of events and the time on task. This time the students are grouped based on their scores on the essays, and based on their change in score from the first to the second draft. Fig. 9.7 shows all of the first drafts. Interestingly, students with medium essay scores seem not to have spent a long time on their essays. For those who did spend a long time, higher-scoring students had more events than lower-scoring students. This could mean those students who had higher-scoring essays were more focused while they wrote (had fewer distractions slowing them down), or they were more efficient at making revisions (getting more in during the same time), or perhaps they were putting less thought into each revision (therefore doing them faster).

The second drafts are plotted grouped on the scores of the essays in Fig. 9.8. Here the trends are not as clear, but it appears a handful of medium-scoring students spent
Figure 9.5: This shows the number of revision events for each of week of the second draft essay assignments. The data are grouped based on the students’ final exam scores.
Time on Task vs. Time Through the Quarter (all second drafts)

Figure 9.6: This shows how much time on task was spent for each of week of the second draft essay assignments. The data are grouped based on the students’ final exam scores.
Figure 9.7: This shows the number of events versus the time on task for all first drafts. The data are grouped based on the students’ essay grades.
Figure 9.8: This shows the number of events versus the time on task for all second drafts. The data are grouped based on the students’ essay grades.

more time and had more revisions than the other students. They were apparently putting in more effort than was reflected in their final essay grade. Again, most of the very-low-scoring students are clumped at the low end in time and number of revisions. Fig. 9.9 shows that the students with very low changes (on average negative) in score from their first to their second draft did not spend as long as some of the students with medium and higher changes in score.

As with the Winter quarter data, we look for trends in the size of revisions. This time we also look for trends in the location of revisions, since that is meaningful for first drafts; edits only at the end of the text indicate a lack of reviewing/revising during the writing process. We track trends in these data based on the students’ final
Figure 9.9: This shows the number of events versus the time on task for all second drafts. The data are grouped based on the change in the essay grades from the first to the second draft.
exam grade, their essay grade, and time through the quarter. For the first drafts, it is not surprising that students make many more smaller changes than larger changes. This also indicates that they don’t write very large pieces of text without making typos. Notice in Fig. 9.10, as was found before, the highest-scoring students have the most revisions, the lowest-scoring students have by far the least, and those just above and below average have similar behavior. The size of the revision events yield very similar results when students are plotted in terms of their essay grade, as shown in Fig. 9.11. It is somewhat surprising that high-scoring students should have so many more 1-word changes since those also involve superficial changes.

The shape of this graph and the number of revisions of each size does not change much from week to week, except in the final essay, which was half the length of the others. It is not surprising students would have fewer revision events that week. This is shown in Fig. 9.12.

The location of revisions is similar for students across final exam score, as seen in Fig. 9.13. Better-scoring students had more revisions at the end of their texts, while students scoring in the middle range of the final exam had slightly more between 75-99% of the way through the text. These students were looking at recently written text and revising it as they write their first drafts. The general trend of far more revisions occurring in the latter part of the text is seen in Fig. 9.14 to be similar when the students are grouped by their essay grades. The difference is that here the students with high-scoring essay grades have more revisions across each category than the other students.

Again, the major difference seen when these data are grouped based on week is that in week nine fewer revisions were made. These results are shown in Fig. 9.15. It
Figure 9.10: This shows the average size of each revision event for all first drafts. A 1-word revision includes when students fix typos while writing. The data are grouped based on the students’ final exam scores.
Figure 9.11: This shows the average size of each revision event for all first drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ total essay grades.
Figure 9.12: This shows the average size of each revision event for each first draft. A 1-word revision includes students fixing typos while writing. The data are grouped based on the assignment number, which also represents progression of time through the quarter.
Figure 9.13: This shows the average location of each revision event for all first drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ final exam scores.
Figure 9.14: This shows the average location of each revision event for all first drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ total essay grades.
Figure 9.15: This shows the average location of each revision event for each first draft. A 1-word revision includes students fixing typos while writing. The data are grouped based on the assignment number, which also represents progression of time through the quarter.

It is also interesting to see that fewer revisions were made in the first week. Similarly to what was concluded above, it is possible that students realized the need to put more effort into the essays after the feedback from the first assignments.

We do the same analysis for the second drafts, with the addition of looking at the size of revisions and the location of revisions in terms of the change in essay grade from the first to the second draft. The location of revisions is harder to interpret for
second drafts, as when editing a number of factors influence where a student makes changes.

Fig. 9.16 shows the size of the edits, with students grouped by their final exam grade. This graph is very similar to the graph for the first drafts, with a few exceptions. Note the slight increase in revisions longer than 20 words; this represents the students who made large bulk changes to their drafts, and is not surprising to see in second drafts versus first drafts. In addition, the lower-scoring students have very similar averages for each category, while the higher-scoring students have more smaller revisions with the highest-scoring students having by far the most. Fig. 9.17 shows similar trends by grouping the students based on their essay grade. The notable difference is that the medium-scoring students now appear more like the high-scoring students, while the low-scoring students stand out as having far fewer small revisions.

Second draft writing behavior as determined by the size of the revisions did not change much from week to week, as can be seen from Fig. 9.18. It is again seen that there are fewer week 2 events, which may be due to a lack of realization about what this type of assignment would entail, as discussed previously. While behavior in week 4 and week 8 are almost identical, week 6 also has an average fewer small revisions. A possible explanation is that this may be due to essay content, or something outside the realm of writing, such as correspondence with a course midterm that decreased the amount of focus students could put on the essays. It is no surprise that no distinctions appear when looking at these data based on the change in grades from the first to the second draft. This can be seen in Fig. 9.19. In fact, almost no parameters correlated with the change in essay grade.
Figure 9.16: This shows the average size of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ final exam scores.
Figure 9.17: This shows the average size of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ total essay grades.
Figure 9.18: This shows the average size of each revision event for each second draft. A 1-word revision includes students fixing typos while writing. The data are grouped based on the assignment number, which also represents progression of time through the quarter.
Figure 9.19: This shows the average size of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the change in grade from the first to the second draft.
Looking at the location of the revisions we do not expect the clear shape we saw with first drafts, with most revisions occurring near the end of the essay. Looking at Fig. 9.20 shows that this is true. There is overall a steady trend of fewer edits as one gets further into the essay text. Sometimes people added new content at the beginning and sometimes at the end, while often smaller revisions consisted of changes within the essay. It is interesting that the students with very high final exam scores had a more steady slope than the other students, but there is no obvious reason why that might be true. Grouping these students based on their essay grade shows a similar downward trend as the students progressed through the length of their essay. This can be seen in Fig. 9.21. This time the lower-scoring students had the more steady decrease.

Looking at the location of revisions from week to week could be more interesting for a few reasons. First, as students get better at writing we would expect fewer bulk changes at the beginning and end of their essays. Second, we know some students started adding filler to their first drafts they expected to cut out in the second draft. These two effects would counteract each other, however. No clear trend is seen in these data from Fig. 9.22. There also could be interesting trends when looking at this from the perspective of the change in essay grade as perhaps the people who had to make large content changes (often occurring at the ends of essays) made more improvements. Fig. 9.23, however, again shows the same downward trend, but no real patterns in the data. It is interesting that the students who had worse changes in their grade had more revisions early on and fewer later on.
Figure 9.20: This shows the average location of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ final exam scores.
Figure 9.21: This shows the average location of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the students’ total essay grades.
Figure 9.22: This shows the average location of each revision event for each second draft. A 1-word revision includes students fixing typos while writing. The data are grouped based on the assignment number, which also represents progression of time through the quarter.
Figure 9.23: This shows the average location of each revision event for all second drafts. A 1-word revision includes students fixing typos while writing. The data are grouped based on the change in grade from the first to the second draft.
9.5 Conclusions

Overall, the results found from Spring quarter are more consistent with the writing and revision behaviors reported in the literature than the data from Winter quarter. There are several reasons why this could be true. They include the previously mentioned difficulties in the Winter quarter, with students being unsatisfied with the essay grading, and the lack of consistency each week on the essay prompts. It could also be due to the fact that using Microsoft Word’s “Track Changes” shows students their revision process as they write, which may artificially affect the way they write. It is possible that a major benefit of the new tracking program that the tracking is not visible to the students. In addition, we have gathered more data from the new tracking program, allowing us to study first drafts as well as second drafts.

For all drafts, the final exam scores were an indicator for the quality of the content in the essays and the overall essay grades, but not the quality of the writing. Correlations with between the content and writing grade did not appear until later in the quarter, which is possibly an effect of the writing practice. There were no correlations with the change in essay grade from the first to the second draft.

We found that the number of revisions typically increased with the time on task, but very-low-scoring students (both on the final exam and on the essay) tended to have fewer events and a shorter time spent. These events did not seem to change much with practice, though students had fewer events and spent less time on the shorter ninth essay.

On both the first and second drafts students had more revisions that were smaller, and fewer that were larger. However, they had slightly more very long (more than 20 word) revisions on the second drafts, consistent with cutting bulk to shorten them.
Trends were similar for high- and low-scoring students (except high-scoring students had more events overall) based both on the final exam grade and based on the essay grade.

For the first drafts, students made most of their revisions later in the essays, indicating that when students write first drafts they make most changes at the end of the text and to the most recently written text. They do not go back much to change text written early on. This is perhaps not surprising for such short essay assignments. This is consistent for groupings based on grades and based on the final. It is also consistent through the quarter. On the other hand, for the second drafts, the revisions were spaced fairly uniformly in location within the text.

Overall, we observe that student writing and revising behaviors are fairly consistent with expectations, and do not seem to change through the quarter. We also do not observe that predicted behaviors (such as time on task) correlate with the amount of improvement made in revising the first draft into the second draft. We also observe that the students’ final exam scores are better indicators of the essay grades than the revision behaviors.
CHAPTER 10

CASE STUDIES FROM PHYSICS BY INQUIRY

Thirteen students from the Spring quarter Physics by Inquiry (PbI) class volunteered to participate in extensive interviews about their views on writing, approaches to writing, and about specific behaviors captured through the tracking program. Students were given a $20 Barnes and Noble gift card to compensate them for their time. The interviews served several purposes. They helped us gather information on how the essays in PbI and more generally writing in physics is viewed by a sample of the students. Most importantly, they also allowed us to see if the interpretations we made of the tracking data were consistent with what the students were actually doing. Finally, the interviews provided us with a basis for looking at certain individuals more closely as case studies in how students in physics courses approach writing.

Some of the things we learned through the interviews are that the pause times are not a reliable measure of writing behavior. Some students reported writing in front of the TV, while others reported having family members in the background who sometimes demanded their attention. In the future, we will put a button on the webpage that allows students to “pause” the tracking program while they take a break, which should at least help with this limitation. In addition, we learned that some students take notes before writing their essays. We have no way of capturing
the revisions that occur during pre-writing, so if a substantial amount of changes are made at that stage we do not get an accurate representation of the students’ writing habits. In future studies, it would help to combine the tracking with some of the other research practices for studying writing, such as think-aloud protocols to find out what students do when preparing to write.

As was mentioned in Chapters eight and nine, we found out that some students started adding “filler” to their first drafts, knowing that the second drafts would need to be shorter. Although this means the first drafts became somewhat artificial, it also means the students were having an easier time responding to the essay prompt with fewer words. Many students reported initially having difficulty staying under 275 words for their first drafts, but said that they found it easier to stay within the word limit as the quarter progressed. It seemed some students were learning to write in a more concise scientific manner with practice.

Despite these limitations, we were able to confirm many of the student behaviors that were evident from the tracking data. It was clear which students had the practice of rewriting things they felt needed changing from their first drafts, versus those who tried to modify and rephrase what was there without cutting it altogether. It was also obvious which students went back and reviewed their drafts before submitting them. In addition, with the second drafts, students self-reporting on how they approached deciding what to change were consistent with our inferences. For example, some students started reading through and revising as they went (sometimes making several passes through the essay), while others homed in on specific content areas for editing, and skipped around the essay as they worked.
In terms of specific types of changes people make when working on second drafts, often students cut statements that they consider less necessary, such as wordy transitions or secondary examples. Other strategies included rephrasing to combine two content areas into one sentence, especially when things are being compared. In first drafts, most revisions included fixing spelling and grammar errors, and changing specific word choices. Sometimes people start to write a word or phrase, then erase it and start with a somewhat different approach.

From all the students in our case studies who completed the interviews, we identified two groups. One group of students submitted most of their essays using the tracking software. The other group was composed of students who did not submit most essays through the software. (Recall that some essays could not be used because some students wrote their essays in other programs, or did not always submit homework.) Six students composed this group. Of these, four students were chosen: one to represent each of the four groupings based on their final exam grade: very high, high, low, and very low. For the groupings that contained more than one possible student, we choose the students who had more consistent essay grades through the quarter.

10.1 Case Study 1: Very Low Scoring

The first case study will focus on a student who did very poorly on the final exam. On average, students in this category had far fewer revision events than other students, and typically spent less time on their essays. Our volunteer in this category was a senior majoring in communication and political science. We call her C1. She took PbI to fulfill a science requirement. She had taken a different topic of PbI
previously and had found it easy, which is largely why she chose to take it again with a new topic. She had not taken it the previous quarter, so had not done the essays in her other PbI course. She had not had physics since her junior year of high school.

She felt the purpose of the course was to learn how to learn without focusing on memorization. She mentioned critical thinking, and learning by doing when discussing the class. She found the essays to be relevant to the course, and thought they were easy, but did not become easier with practice. She thought the essays reinforced concepts from the class, but was not convinced they actually help you understand the concepts better. She felt the essays helped her put the concepts into words, but other parts of the class helped more to prepare her to explain concepts on the exams.

C1 spent 15-20 minutes on the essays, about 10 minutes on the journals, and 1-2 hours on the standard portion of the homework assignment.

In general, C1 felt that writing definitely helps learning. She said when you articulate what you are learning it helps you make sure you know what you are talking about. She thinks writing should be part of the process of learning physics, and added that explanations should also be part of math courses. When writing the essays for PbI, C1 did not do any planning before beginning to type. She tried to address each portion of the essay prompt as she typed, and she looked through her notes if she reached a part she could not address offhand. In her case, analyzing the location of her pauses may be some measure of what content she struggled with. She noted that in other courses she typically planned a lot before writing, but here no research was needed, and the essays were short so she was not as worried about language issues. She did not feel that writing directly in a webpage changed the way
she would normally approach the assignment, since she would normally write directly in another computer program.

Before writing the second drafts, C1 reviewed the essay prompt and looked at my feedback, then started reading the first draft and revising it as she went. She first removed things she felt were redundant; sometimes she knew when writing the first draft that some information was relevant but not necessary so she cut that as well. Once she revised down to 125 words she stopped working, as long as she knew the needed content was present.

I looked at her seventh and eighth essays in detail with her as part of the interview. For her seventh essay, she attributed pauses to reviewing what she had just written, looking up content, and answering the phone, though she could not identify individual pauses from the tracking data. There was an episode about two-thirds of the way through her text at a point where she stopped writing in a forward fashion and revised a chunk of her text. She remembered considering whether her structure was good: should she mention the voltmeter early on or wait and address all aspects of the question with respect to the voltmeter later on? She feels that when writing you usually think about the idea you want to write, then just write it without looking down or spending much time thinking about the language choices. She recalls normally reviewing the essay before submitting it, though there is no evidence of that in this particular assignment.

Fig. 10.1 shows samples of the types of writing we expect to see in first drafts. The black line represents a student who only writes linearly with all text added at the end. This student does not go back and alter any text once it is written. Because the revisions are plotted as the percent of the way through the text that existed at
that time, additions early in the text appear to occur farther back than they really are. For example, if the student has only written 3 words and changes the last word due to a typo, that change would be 67% of the way through the existing text. The red line represents a student who mostly adds text to the end, but also goes back and makes edits to previously written text. Here a large early revision episode is shown around event 25, and a small episode in recently written text is shown around 90.

Figure 10.1: Sample of approaches to first draft writing.

Fig. 10.2 shows the sequence of revisions C1 made while writing the seventh essay. Edits are a combination of an addition and deletion that was captured as one event by
Figure 10.2: This graph shows the location of the revisions (as the percent of the way through the text written up to that point) in order of the sequence of revision events for C1. The size of the dots represents the number of words that were added or deleted during that revision event.

The tracking program. It is not clear why occasionally these events were not captured separately. The most striking feature of this graph is that most of the data points are at the top, meaning most revisions occurred toward the end of the existing text. Almost all deletions are at the end, and are smaller than additions. Many of these are most likely typos that were being corrected as C1 typed. It is also interesting that there is a region about two-thirds of the way through her writing where she made a lot of changes to earlier text. This the episode she recalled in her interview.
In her eighth essay, the second draft of the same topic as the seventh assignment, C1 first looked at the text to rephrase ideas and shorten content that needed to be kept. She then looked at what bulk could be cut, sometimes choosing to cut content she had recently rephrased. Many of her revisions occurred in one particular sentence. She mentioned finding it hard to discuss the idea concisely without changing the meaning. She remembered working on that sentence and found it hard to phrase it the “right way.” This particular sentence discussed the importance of matching positive terminals of ammeters and voltmeters to positive battery terminals when hooking them up to circuits. She paused at the end before submitting the essay; she remembered reading it again to make sure it was okay, and also rereading the prompt to make sure all parts of it were answered.

Fig. 10.3 shows samples of the types of writing we expect to see in second drafts. The black line represents a student who makes three passes through the first draft while revising it. The revisions shown are fairly steady, indicating that the text is uniformly altered everywhere. The red line shows two passes of revisions, this time focused around specific regions of text. For example, the first half of the red line could be rewriting of one or two sentences only.

Fig. 10.4 shows the sequence of revisions C1 made while writing the eighth essay. This graph is strikingly different from the graph for the seventh essay. It appears that most of her bulk cuts were early on in the revision process, contrary to what she recalled as her usual habit during the interview. There are small edits interspersed with the large ones, however. There are also clumps of dots near each other indicating that successive revisions were on the same portion of text. Notice that those clumps also slope upward, indicating that as C1 edited a portion of text, she moved forward
through that area and made successive revisions. The single dots at the end of the revisions occur very spaced out in the essay and seem to correspond with re-reading the essay before submitting it and making final corrections.

Her revision habits seem to be well thought out; however, her detailed essay grade for the eighth assignment was only 19/24. For the seventh assignment it was 20/24, so she decreased the quality of her draft somewhat with her revisions. The difference was a decrease in one of the content areas needed to respond to the prompt; she decreased from a 3 to a 2, meaning the content was still there but not explained as clearly as before. The average score for all essays chosen for detailed study was 18.8 for the
Figure 10.4: This graph shows the location of the revisions (as the percent of the way through the text written up to that point) in order of the sequence of revision events for C1. The size of the dots represents the number of words that were added or deleted during that revision event.
seventh assignment and 18.6 for the eighth assignment. It is perhaps noteworthy that although she did very poorly on the final exam, her essays were above average, possibly indicating that her writing and revising strategies were somewhat successful.

C1 spent 15 minutes on the seventh assignment and had 180 revision events. She had the largest number of revisions on this essay of all students who were in the very low-scoring group, though her time on task was about average. She spent 10 minutes on the eighth assignment and had 87 revision events. No more than 10 of her events were pauses in either assignment. She had more events than the average person in her group for this assignment, and again her time on task was near average.

For the seventh assignment she had completed 93% of her revisions when she reached the word limit, so she did not do much with her essay after that. In contrast, she reached the 125 word limit for the second draft (the eighth assignment) 64% of the way through her revisions, indicating she did a lot more work after that point. This was lower than the percentages on her other second drafts, which is partly due to that one sentence she struggled with rewriting. In the second draft, she had several revision events larger than 20 words in length, though for both assignments most of her revisions were small. In her first draft, most of her revisions occurred at the end of her text, while for the second draft they were fairly evenly spaced throughout the text.

Overall, C1’s behaviors are typical of students in her group, with the exception that she made a greater number of revision events. Her revision strategies seemed well conceived, as well. It is possible these factors contribute to the fact that her essay grades were somewhat higher than you would expect simply by looking at her final exam grade. We also look to see if C1’s writing behavior changed with time. We show
in Fig. 10.5 her revision behaviors for each of her three second draft assignments; she did not complete the second assignment in the tracking program. Her revision patterns for each essay are fairly similar. She seems to start at the beginning of the essay, and work forward making revisions as she goes, then repeats this process to complete three forward passes through the essay. In each week she has clumps of revisions around portions of the text within her forward pass; her third pass is also shorter than the rest, indicating she made most of the changes she wanted in the first two.
In week four C1’s essay grade was slightly below average at 18/24, while in weeks four and eight they were slightly above average with 19/24 and 20/24 respectively. In week four her change in essay grade from the first draft was -2, while in week six it was +1 and in week eight it was -1. In week four she lost her essay points from the ordering of ideas, while in week six that is where she gained her point. In week eight one of her content areas was slightly less clear in the second draft, causing her to lose a point. Overall she seems to make good content choices when revising. This is not a trivial task; recall students must determine what content is important and how to fit it into half the length of the first draft. She had 64 revision events in week 4, only 38 in week 6 and 87 in week 8. Interestingly, the week she had the least revisions was the only week she improved her grade from the first draft. Her average pause lengths and the average sizes of her revisions were fairly similar each week.

10.2 Case Study 2: Low Scoring

The second case study will focus on a student who did below average (but not very poorly) on the final exam. On average, students in this category had similar behaviors to students who scored just above average. We call this volunteer C2. He was a fourth year middle childhood education major who took PbI to fulfill an education requirement. He had never taken any physics before this class. He told me he enjoyed the course, and found it similar to a required inquiry-based math class he had also enjoyed. He felt the purpose of the course was to learn to be a better teacher.

He thought the essays were good, and they helped him develop his knowledge further. He felt it hard to have less than 275 words in the first assignment, but found
that easier as the quarter progressed. He felt they became easier through practice and because he better understood the expectations. However, he also felt that the essays required too much work as the quarter progressed and he thought that the essays got more difficult. C2 spent about 1-1.5 hours on the essay assignments, and said he spent similar time on both the journals and the homework.

C2 felt that writing helps learning because it helps you remember what you are writing about, and since it is a “hands-on” activity it helps to make a more “lasting” impression. He felt this was true for the essays in PBI as well; along with the journals, he felt they also helped him review his understanding. He felt the essays also helped with writing explanations on the exam questions. He also stated that writing should be part of the process of learning physics, mainly because it is “hands-on.”

When C2 sat down to write the first drafts, he typically thought about what major points he would need to include, then wrote the essay, then reviewed what he wrote before turning it in. When asked if he felt the way he approached writing in PBI was similar or different from writing in other classes, he immediately stated it was similar to what he wrote in his math class. Then he added that it was also similar to other types of writing, and even the way he approached fiction; first thinking about what to include, then writing, then revising. Notice that he exhibits expert-like writing behavior with pre-writing, writing, and revising. C2 did not feel that doing the writing in the webpage affected his normal writing strategies.

When writing his second drafts, C2 first re-read the first draft and thought about the feedback given. Then he worked on shortening the first draft. He first looked for things that were unnecessary or wordy, and cut them. He said his first drafts were normally correct so he did not need to change the content, but just shortened
it. Therefore, once he shortens the first draft to the second draft 125 word limit, he was usually ready to submit the essay. He noted that when he wrote the first draft he usually felt that everything was necessary and did not see what is wordy, but when he read it later for the second draft he found it easy to see what could be cut.

Again the seventh and eighth essays were considered for further analysis, and discussed in the interviews. C2 had a lot of pauses spaced fairly regularly throughout his writing. He said he was paranoid about the word limit, and did not trust the word count from the tracking program, so he kept copying his text from the webpage, pasting it into Microsoft Word and checking the word count there. He made some interesting word choice changes while writing his first draft (the seventh assignment), such as going back and adding the word measuring, to specify that the essay was not just discussing current and voltage, but how to measure it. He also changed words to make them “more professional,” such as changing the word “moves” to “travels,” and adding the words through and across. He stated that he revises while writing, then goes back and revises it again at the end. The last event recorded before submission was, in fact, a pause of nearly 1.5 minutes, likely corresponding to him giving it a final read through.

Fig. 10.6 shows the sequence of revisions C2 made while writing the seventh essay. Notice how strikingly different the graph of his revisions is from those of C1. Many of his revisions are not in the most-recently-written text. He seemed to go back about 25% of the way and modify the text while he wrote, and near the middle of his writing he jumped to a previous location to make more changes. C2 also has a string of revisions at the end of his writing that start at the beginning of the essay and work through to the end. These revisions seem more extensive than just re-reading and
Figure 10.6: This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C2. The size of the dots represents the number of words that were added or deleted during that revision event.

checking for errors. He seemed to be doing text modification, perhaps smoothing out ideas, or fixing content issues.

In his eighth essay, the second draft of this same topic, C2 scanned the essay, made some changes, then repeated the process. He usually cleaned up the text near an area where he makes revisions, while he looked at that part of the essay, and then skipped around to work on another area. In this essay he started making revisions about two-thirds of the way through the text, then made a bunch of revisions further up, then made the last ones near the beginning of the essay. He did this because
Figure 10.7: This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C2. The size of the dots represents the number of words that were added or deleted during that revision event.

the top portion of the essay seemed like important content and he was not sure how explicitly he needed to discuss it. He also cut an example from the end because he did not like it, and felt he needed a lot more space to explain it clearly.

This information came from discussing the S-notation with C2 during the interview. The graph in Fig. 10.7 shows the sequence of revisions C2 made while writing the eighth essay, and confirms the information revealed in the interview. We see sets of revisions chunked through the essay, and not necessarily done sequentially. There is also not clear evidence of a final pass to check the content (the later revisions he makes are fairly large), which is also consistent with his self-reporting.
C2’s writing habits seem consistent with expert behavior, but his revision habits seem a bit more random. His detailed essay grade for the seventh assignment was 21/24. Compare this to the average grade of 18.8; his first draft was fairly good. However, his grade for the eighth assignment was only 18/24, which is below the average of 18.6. The points he lost from the first to the second draft were on content and correctness. He removed (or made unclear) content that was in his first draft, indicating that his revisions were not successful. It is possible he had difficulty choosing what content was important to keep. Another possibility is that his revisions were too coarse, and he did not make sure his final content was clear after editing it.

C2 spent 33 minutes on the seventh assignment and had 126 revision events. His time on task was larger than most people in his group (based on the final exam grade) but his number of events was not. He spent 19 minutes on the eighth essay and had only 60 revision events, which is also a relatively long time and small number of events. For the seventh assignment he had completed 78% of his revisions when he reached the word limit, so he made several changes after that point. In contrast, he reached the 125 word limit for the second draft 97% of the way through his revisions. Consistent with his interview and the graph above, we see that he did not spend much time on his essay after reaching the word limit. This lack of final edits may be partly responsible for his low second draft grade. C2’s overall writing seems inconsistent, since his first draft behavior is successful while his second draft behavior is not.

We can also look to see if C2’s writing behavior changed with time. We show in Fig. 10.8 his revision behaviors for each of his three second draft assignments; he did not complete the second essay in the tracking program. In his fourth and sixth assignment, most of his revisions are at the end of his text. This indicates that
although he started by revising his first draft, much of what he did was add new text instead of modifying the text he had in his first draft. However, toward the end of his revisions he did go back in the text somewhat, not just making all changes at the end. In his eighth assignment his revision pattern was very different. Here he appears to have made two major passes through the essay and spent almost no time making revisions at the end.

In weeks four and six his essay grades were above average (21/24 and 19/24 respectively) and in week eight his grade was slightly below average (18/24). In week four his change in grade from the first draft was 0, where in weeks six and eight they
were -2 and -3 respectively. It seems that his strategy of mostly re-writing instead of revising was successful in week four but not in week eight. This is deceptive, however, because in week four he had 392 words in his essay when the word limit was 125. Another interesting note is that in week four he increased his writing grade by two points but decreased his content grade by two points, so even though he added text he decreased the amount of needed content in his essay. He seems to consistently make bad choices each week about what content to keep; his content grades account for most of his decrease in points each week. In week six he reaches the word limit 58% of the way through his revision process; since most of his edits are at the end it is likely he cut a lot of text, then added more. In week eight he reached the word limit 97% of the way through his revisions. He has 120 revision events in week four, 97 in week six and only 60 in week eight. The average sizes of his revisions are fairly similar each week.

10.3 Case Study 3: High Scoring

The third case study focuses on a student who did above average (but not very high) on the final exam. The behavior of students in this category were overall similar to those who scored just below average. We’ll call this volunteer C3. She was a second year middle-school education major who took PbI to fulfill an education requirement. She had high school physics, and found it pointless; she could just memorize things said in class and do well without understanding it. She thought the purpose of PbI is to be able to explain yourself well, and also compared it to the inquiry-based math course.
She felt the essays fit with the purpose of the course but she did not like them; she found it hard to keep up with so much writing. C3’s view of the essays did not change through the quarter. She also did not feel they became any easier to write with practice. Since she felt the material in the essays was relevant she felt they had value even though she would rather not have written them. She said she spent about 15 minutes on the essays, while the journals took her about 10 minutes and the homework took 1 hour.

C3 thought that writing helps learning because you can see what you wrote and rewrite it if it does not make sense. She said she used writing in high school to help her remember things, for example rewriting her lecture notes. In PbI, she felt the essays helped about the same as the homework for understanding course concepts. She did not feel they helped prepare for the exams, however, because she felt the in-class work and the homework were sufficient for that. She felt that writing should be part of learning physics because it helps you remember things; but at the same time she said it would be tedious for the people who already understand the material.

When C3 sat down to write the first drafts, she first looked through the book to see how the topic was addressed in class, or she sometimes looked things up online. She then thought about what she needed to write, and had a general idea of the whole essay before writing. She said this is similar to what she does when writing for other courses. She did not like writing in the webpage; she preferred writing in a more traditional word processor. However, she does prefer typing to writing by hand. When writing the second drafts, C3 first looked at the original draft, the prompt and the feedback. When determining what to cut from the first draft, she first got rid of filler, then cut anything which is redundant. Then she saw if there are phrasings
that can be shortened, or simplified. She often cuts out the examples if she needed to remove more words, though not in the eighth essay because she felt the example was needed to make the content clear.

When looking at her seventh essay in more detail, we see that C3 has pauses spaced through the writing process. She often went back and read the whole text while she wrote. Sometimes she made revisions when she did that, but did not feel much need to do that with this topic because it was fresh in her mind from a recent in-class activity. In places she started to write something, then cut it out; often she said she realized she needed more explanation before starting on a new aspect of the response. She did not pause at the end of her writing, nor did she recall going back to re-read it before submitting the draft.

Fig. 10.9 shows the sequence of revisions C3 made while writing the seventh essay. We see her writing behavior is more similar to C1 than C2 in her first draft, with most revisions occurring in the very recently written text. We again see that most deletions were very small, indicating that they were mainly typos being corrected while she wrote. She did have an episode of revisions about halfway through the writing that are slightly deeper in the text, though she did not go very far back. This is consistent with the idea that she has most of the essay planned out in advance; she may not need to go back and revise or add content. It is possible that she is careful as she writes and would not need to do a final read of the text, but it is also likely that reading it again before submitting it would be beneficial.

In C3’s eighth essay, the second draft of this same topic, she initially made some edits that removed statements that in a longer writing might be needed for clarity but could be sacrificed in a shorter essay. She recalled jumping around a lot when
Figure 10.9: This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C3. The size of the dots represents the number of words that were added or deleted during that revision event.
revising the essay, as well as pausing to read what she had changed to make sure it made sense. She also paused to look ahead and see what to edit next. She sometimes cut content then realized it was needed later and added it back in. She made a lot of small edits where she reordered and rephrased content to condense it. She found that sometimes when she rephrased to make things clear the text got longer; then she needed to go back and cut other parts of the essay. She made several passes through the essay to edit it.

The graph in Fig. 10.10 shows the sequence of revisions C3 made while writing the eighth essay, and confirms the information revealed in the interview. It is very clear that she made multiple passes through the essay while revising it. There are also places in the text where she made several changes in one location before moving forward. There is also evidence that she made one final pass to make sure the essay was clear, as there are a few small edits spaced out at the end of the revision process.

Without knowing about the pre-planning process by C3 I would assume from the data that she was not exhibiting expert-like writing habits. However, her detailed essay grade for the seventh assignment was in fact 24/24; she produced an excellent essay. Her revision behaviors in contrast seemed very expert-like, though her detailed essay grade for the eighth essay was 21/24. This is still well above the average grade of 18.6. She lost one point for ordering; her second draft did not have as good a logical order as her first draft. She lost the other two points on two of the content areas; the points were still present but not as clear as they needed to be.

C3 spent 26 minutes on the seventh assignment and had 110 revision events. The time on task is fairly average for people in her group, but the number of events is fairly low. For her eighth assignment, she spent 16 minutes and had 87 revision
Figure 10.10: This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C3. The size of the dots represents the number of words that were added or deleted during that revision event.
events. Both the time on task and the number of revisions were fairly low compared to members of her group. This is not reflected in the quality of her essays; her grades were quite high on these. For her seventh essay she reached the word limit 65% of the way through the writing. In her eighth essay she reached the word limit 98% of the way through the writing. For the second draft this is not surprising based on the writing habits she explained, since she makes small careful edits to reach 125 words. Based on the tracking data, C3’s writing behavior did not seem expert-like on the first draft. However, her interview explained her pre-writing activities, changing this viewpoint. Her revision behavior did seem expert-like from the tracking data. Her habits for both are overall successful, also, despite spending less time and fewer revisions on each.

We can also look to see if C3’s writing behavior changed with time. We show in Fig. 10.11 her revision behaviors for each of her three second draft assignments; she did not complete the second assignment using the tracking program. In her fourth assignment, she appeared to make four forward linear passes through her essay while making revisions. She seemed to progress forward fairly steadily, typically not lingering around any one area. Occasionally students would hit enter at the ends of the lines when using the webpage, not realizing the text would wrap on its own. Those hard returns would show up in the saved text, so students had to delete them. It is likely the first pass for revisions in week 4 and week 6 are simply the student deleting those hard enter symbols that are evenly spaced in her text. In her sixth assignment, she made some changes in her first pass of deleting the hard returns, but most of her changes are made in a second pass, where she spent some time lingering especially at 1/4 and 3/4 the way through the essay. In her eighth assignment, she appeared to
Figure 10.11: This graph shows how C3 revised her essays when writing the second drafts.

In weeks four, six, and eight, C3 had higher than average essay grades (20/24, 20/24 and 21/24 respectively). However, in week four her change in grade from the first draft was -1, in week six it was +1 and in week eight it was -3. There was no indication from her revision behavior as to why her revisions were less successful in week eight. C3 had more events and longer pauses in week 2; these parameters were
very similar in weeks four and six. In week four she was slightly over the word limit; she did not manage to get as low as 125 words; this may be why she made so many revision passes; she may have been struggling to get the word count down. In week six she reached the word limit 87% of the way through her revisions, but in week eight she reached in 98% of the way through indicating she didn’t keep working on the essay once she reached the word limit. The average size of her edits each week were fairly similar.

10.4 Case Study 4: Very High Scoring

The fourth case study focuses on a student who did very well on the final exam. Students in this category on average had somewhat higher amounts of revisions than students in the other groups. We call this volunteer C4. She was a third year middle childhood education major, and took PbI as a requirement. She also thought the course would be good for learning to explain yourself; which she felt was the major purpose of the course. She also took the previous quarter PbI course covering different content, so had been exposed to the essays previously. She also had advanced placement physics in high school.

She felt that the essays Spring quarter were easier than Winter quarter, and she thought the topics were more uniform and progressed better with the course content, which was also more uniform. She also felt they were easier to write because of experience. She remembered being intimidated by the essays when she first encountered them Winter quarter, but then realized they were not that bad. She spends about 20 minutes on the essays, about the same on the journals, and 45 minutes to 1 hour on the homework.
C4 felt that writing helps learning, and wished there had been more writing in class. She felt it is helpful to express the material in different ways. She felt in the PBI class the essays helped to clear up her thoughts and solidify her understanding, and she felt better able to explain herself in class discussions if she had written about the topic. She also felt the second drafts helped her make her points more concisely, which was helpful for explaining herself during the exams. She felt that writing should be part of learning physics, and recalled using writing when practicing for the advanced placement test in high school.

When C4 sat down to write the first drafts, she looked at each part of the prompt and responded to them sequentially. When writing a paper for other classes she would usually start with an outline, but would still approach the writing piece by piece. She felt that writing in the webpage affected her writing, and she would have revised more if she were writing in a more standard word processor. She also felt forced to write in one sitting instead of being able to go back and work on her essay later.

When writing the second drafts she usually re-read the first draft, took out what she saw could obviously be removed such as extraneous clauses and extra information. If the comments were helpful for making revision choices (if the first draft is good there would be no indication of what to change) she uses those, and when example essays were sent as feedback to the whole class she used that as well. She felt that she has some idea of what can be cut from the first draft when she wrote it, but she did not consciously add filler.

Looking at C4’s seventh essay in more detail, she paused periodically, probably to check the next portion of the prompt that she needed to respond to. She made a lot of grammar changes while writing, for example, changing things so they would
not be in first person. At one point she purposely changed wording to be more vague because she was not clear about the details. She also remembered struggling with some of the content in other parts of this essay. She recalled referring to the book during some of these struggles, causing fairly long pauses. She had no pauses while discussing the examples, though; she remembered what was done in class and did not need to look it up. She did not pause at the end to reread the essay before submitting it.

Fig. 10.12 shows the sequence of revisions C4 made while writing the seventh essay. Her writing behavior is very similar to that of C1 and C3, and in fact most similar to C1. Most of her revisions are in the text she has just written, with many deletions again appearing to be correcting typos. However, she did have some larger deletes, which correspond to places where she changed her mind about how to word her text. There is no evidence from the revision data that she approached the writing in chunks.

In C4’s eighth essay, the second draft of this same topic, she read the essay, made changes, and read it again to make further changes. She also jumped back when she made changes and realized earlier wording should also be changed to be consistent. She also remembered pausing to really consider what could be removed because all the text seemed important with this particular essay. She spent a lot of time revising near the end of the essay because she had trouble finding a way to make the last sentence clear. She later took out big chunks because she needed to find a way to reach the word limit; in fact, she cut sentences that she had earlier spent time revising. She felt this essay took longer to revise than most. Once she reached the word limit, she felt she could not cut more because it was hard to find things that could be cut
Figure 10.12: This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C4. The size of the dots represents the number of words that were added or deleted during that revision event.
Figure 10.13: This graph shows the location of the revisions (as the percentage of the way through the text written up to that point) in order of the sequence of revision events for C4. The size of the dots represents the number of words that were added or deleted during that revision event.

at all. Normally she got to the word count faster and edits after that to make the essay clear.

The graph in Fig. 10.13 shows the sequence of revisions C4 made while writing the eighth essay, and confirms the information revealed in the interview. This behavior is not entirely obvious when looking at the graph, though it is evident that she made several changes to the text before moving on to the next sentence. It is evident that she did not go back and make any more small changes to the essay right before submitting it.
From the data it seems that C4’s writing process is very similar to C1. Like C3, C4 had 24/24 on the seventh essay and 21/24 on the eighth essay. These grades were well above average. An interesting side point is that C3’s essay grades were usually high, whereas a few of C4’s essays were not good. The major difference between C4 and the other people here are that she had a lot of revisions; 255 for the seventh essay and 197 for the eighth essay. She lost all her points for the second draft on content; with three items decreasing from 3 to 2. They were not explained as clearly after the revisions.

She spent 39 minutes on her seventh essay and 24 minutes on her eighth essay. Even though students who scored very high on their final tended to have more events, C4 had more than other people in her group. She also spent more time than most people on the seventh essay, though she spent a fairly average amount of time on her eighth essay. She felt that she spent much more time than usual on that essay, and she struggled with it, yet her time was not remarkably long.

She reached the word limit 85% of the way through her revisions of her seventh essay, and 77% of the way through her eighth essay. Although the second percentage does not seem consistent with her report that she stopped when she reached the word limit, a more careful look at the data shows that she had previously made a large cut putting her below the word limit, but she needed to put back in the content she had cut, so she was not done with the essay. Her writing habits do not seem as well thought out as those of C3. It is interesting that they have the same grades on the essay, yet C4 had better course content knowledge as measured by the final exam. Overall, however, her habits are successful and she seemed to put time and effort into the writing and revising.
We can also look to see if C4’s writing behavior changed with time. We show in Fig. 10.14 her revision behaviors for each of the four second draft assignments. In her second assignment (the first 2nd draft), she appeared to start revising at the end of the essay, and work her way backwards. It is likely this first backward pass is removing hard returns. She then makes two forward passes through the essay making further changes, with some jumping around. In her fourth assignment, she progressed more linearly and started at the beginning. She seemed to linger longer around certain text and leave other text alone, and make a few jumps backward. Essentially she has only made one or possibly two passes through the essay. In her sixth assignment she made two clear forward passes, making changes as she progressed and editing most parts of the essay both times. In her eighth assignment her behavior was a bit like the fourth, with mostly forward passes and some jumps and skips through the essay while she revised. The notable difference with the eighth assignment was she had almost twice as many revisions than she did for the other weeks.

Interestingly, in weeks four and eight, C4 has essay grades that were higher than average (19/24 and 21/24 respectively), while in week two she had a very low essay grade (15/24) and in week six she had a slightly below average essay grade (17/24). In weeks two and four she lost two points from her first draft grade when making the revisions, and in weeks six and eight she lost three points. In weeks two and weeks six she was also over the 125 word limit (by 28 words in week two and 10 words in week six); this did not factor into her detailed essay grade. In week four and week eight, a higher percentage of her revisions were at the phrase level than for the other weeks. This indicates she made more non-trivial changes those weeks. That seems to have been reflected in her final essay grade for those weeks, but not for her change in essay
Figure 10.14: This graph shows how C4 revised her essays when writing the second drafts.
grade. Finding a way to extract this data efficiently would allow further analysis such as this for multiple people, and may help us better characterize writing behavior.
CHAPTER 11

CONCLUSIONS

Many people proclaim the benefits of writing to learn. This includes our students as observed in the case studies from the Spring quarter Physics by Inquiry (PbI) course, literature from those who use writing in their classrooms [50, 54, 66], and a large portion of the academic community, which is actively promoting Writing Across the Curriculum (WAC), including the Ohio State University (OSU). People experience moments of clarity and increased understanding when they write, and presume this same effect must happen to their students as well. There is also the need to introduce students to discipline-specific writing, and to help students achieve the writing skills they will need for the job market.

While these are important reasons, they do not help us know whether writing actually helps learning, or if it does help, when does it help and what do we need to do to make it a learning activity. The benefits of writing are being heavily promoted in the literature and claims are being made that are not backed up by quantitative data or controlled studies. In order to address these questions we need to characterize writing behaviors and see what approaches are effective, and whether they change with time. We also need to understand something about the connection between writing
and content knowledge. The studies culminating in this thesis are an aggressive start
to addressing these complex questions.

While the existing literature is full of unsupported claims regarding Writing to
Learn, it is not without guidance; there are cognitive models for writing that show us
how writing can be a reflective process [51]. There are also studies that characterize
some writing and revision behaviors in terms of novice and expert practices [20, 44],
and in terms of the types of writing and revisions that are exhibited [23, 32, 63].
While some researchers have used methods to track the writing process with think-
aloud protocols [20] and even with computer tracking [27], these methods as they
have been carried out in the past have many limitations. Our tracking program is a
large step forward toward being able to extract large amounts of information easily
and, even more importantly, without interfering with the writing process.

This thesis establishes three main items. First, there is a need for quantitative
studies of Writing to Learn, and specifically of Writing to Learn within physics.
Second, we have also begun to characterize student behaviors in an effort to quantify
the study of writing. Third, we have developed a valuable new tool for quantifying
and analyzing student writing and revising.

11.1 Results of Our Studies

Our first project in summer 2005 focused on the impact of providing English
instruction and on looking at the relationship between English and physics content
when students write. Students in lab wrote essays for each of the six weeks they
met; some students had explicit writing instruction. We found that the writing
instruction had a positive impact on the quality of physics in the students’ writing;
those with writing instruction produced better physics and English with practice. This is not a surprising result considering the literature that discusses the need to instruct students and help them activate the appropriate skills to tackle a new type of writing. However, it is important that this result has been quantified. It is unclear whether they eventually understood the physics better than those who did not have English instruction, or whether they were better able to show what they knew more clearly because they were better able to write it. How much something is understood is also mixed with the ability to explain it clearly, so it is probably not possible to determine which effect is greater.

We did not test explicitly for differences in content knowledge, though students who wrote essays in lab and, specifically, those who had English instruction did do better on an explanation-based final exam question requiring content similar to one of the essays. The standard diagnostic test did not show any significant differences between the student groups. The essays were graded independently for physics and English by two graders. We found that the quality of physics and English correlated strongly each week, and that the placements and types of comments by the two graders were similar, indicating there is a connection between the physics and the English. A possible reason for this is rooted in argumentation structure, which seemed to the graders to be a strong common link between what would make either the English or the physics successful in this assignment. We have also provided quantitative evidence for the idea that writing quality is hindered when students struggle with discipline-specific content.

In the second project, autumn 2005, we did a similar study with students writing in lab, some with and some without explicit English instruction. This time we tied
the content more closely to the lab and to a specific diagnostic test. We saw students who wrote essays did better on many of their lab quiz questions, especially by the third (and final) writing activity. However, there were no meaningful differences in displayed content knowledge on recitation quizzes, course exams, or the diagnostic test. It seems any knowledge that was gained through the writing did not transfer to other parts of the course even though the content was relevant. The writing instruction also did not have an apparent effect, which may be due to the fact that there was less instruction and practice in this study, with only three weeks of writing practice instead of six. In addition, correlations between the English and physics were weaker with these assignments, which were more like writing instructions than writing more standard essays. One strong point of this study is that the students who wrote did no worse than those who did the existing good quality physics education research-based lab activities. The writing activities did no harm in the classroom, even when substituting for activities that are known to be helpful. This indicates that we may be able to teach discipline-specific writing needed by the students while still teaching as much content as physics education research-based lab activities.

We realized that testing for changes in content knowledge due to writing activities may be too much to ask before we have understood how our students are writing and if having them write repeatedly in physics is impacting their writing habits in any way. In addition, we realized the need for a systematic, quantitative way to analyze student writing behavior. We designed writing assignments to force revision, and we tracked the revisions in order to assess writing and quantify writing behaviors.

There were nine writing assignments given in Winter quarter, 2006, with the odd number assignments being first drafts and the even numbered ones second drafts of
the same topic. The second drafts were to be half the word length (125 words instead of 250) to force the students to be reflective when revising; their initial content had to be modified and condensed. We tested revision behavior using Microsoft Word’s “track changes” option as students were writing their second drafts. Using coding schemes from the literature, we coded and analyzed students’ revision behaviors. Revisions were also individually assessed to determine whether the changes to the original text were good or bad.

We looked to see whether different groups of students (for example, those who scored high on their final exams or their essays) had different revision behaviors, and to see whether revision behaviors changed with time. Students who did well on the final and on the essays had more revisions that were coded as good, indicating they were making better choices about what content was important and how to phrase it clearly than those who did not score as well. The quality of the content in students’ essays correlated better with their final exam grades than with their essay grades based on writing, indicating that overall content knowledge is more important for producing good physics than their ability to explain themselves clearly. Also, some writing behaviors, for instance, the number of revision events, correlated more strongly with the final exam grade than with the essay grades. This may just be an indicator that students who put in more effort have better grades on average.

We saw no steady changes in any of the tracked revision behaviors through the quarter, indicating that the writing practice did not impact the way students wrote. It is possible that they were not learning to write in this study; we do not know whether the lack of explicit writing instruction is a possible reason for this. However, there were some weekly variations that likely correlated to differences in the difficulty
of the essay content. This study gave us a preliminary look at how to characterize student writing, and showed us the limitations of current methods. It also helped us refine our writing implementation for our next study; for instance, we were able to make the essays more uniform to minimize fluctuations due to content.

In Spring quarter, 2006, we did a study similar to that in Winter quarter, with the main exception that we had students write directly in a web-based text-editor that has a built-in tracking program we developed. This allowed us to study writing as well as revisions, to automate the data collection process, and to have the tracking interfere far less with the way the students naturally write. The benefits of this program to the community studying writing behavior can not be overstated. Coding the revisions by hand from “Track Changes” took between 20 to 60 minutes per essay. In our Winter quarter study, over 200 essays were coded. This was a tremendous time commitment, and due to the repetition and grader fatigue is more likely to be error-prone than an automated program. The study from Winter quarter did not give enough data to characterize student behavior despite taking such a tremendous amount of time; the instructor time needed without the availability of our tracking program would be prohibitive for a larger study. This new program shows promise for enabling much needed rigorous quantitative studies in writing to learn. In addition to the other important benefits of the program mentioned previously, we have plans to expand it, allowing further analysis as will be explained in the next section.

From the Spring quarter data, we were able to extract a lot more information about student writing and revision habits. However, we still have seen no clear indications that student writing changed during the quarter. The one change we did see over time is that the essay grades based on writing correlated more with those based on content.
This could be an indicator that students were learning to explain themselves better, or better able to structure their physics content. Students with low final exam grades and low essay grades typically spent less time writing and had fewer revision events, while those who had high grades had more revision events. Most students wrote their first drafts without going back and editing previous text, only revising the text that they had just produced. For example, they could be fixing typos or making a different word choice for starting a new idea. When revising, however, students exhibited a range from small edits to large cuts. Overall, these behaviors are consistent with models for how students write and revise, but there are few indications that they are different for different subgroups of the population or that they change over time. It is unclear whether practicing the writing through the quarter had any impact on how the students wrote.

To get a better idea of the validity and breadth of the picture we obtain from the tracking program, we conducted student interviews and presented case studies from the spring class. We found that students who exhibit more expert-like writing habits (for example, having more non-trivial revisions) seem to have better essays than we would have expected. As one example, one student who scored very low on the final exam exhibited better writing behavior and her essays were above average in physics and writing quality. Also, one student who had a very high final exam score had better essay grades when she exhibited better writing behavior. Although generalizations usually cannot be made based on individual behaviors, we can confirm that many of the conclusions we can draw from the tracking are consistent with the behavior reported by the individual students interviewed. The inability to know how much planning students actually did before they began writing is the most serious
limitation of the present study. However, this problem will be characteristic of any other such study.

Overall, we have established the need to study the impact of writing by exhibiting the current lack of real data-driven statements supporting use of writing, we have shown that the link between writing and learning content is not obvious, and we have shown that students may not even be learning to write through practice in the context of physics. Our central contribution is the significant advances in terms of technology used for and practice in how to study writing, and how to characterize writing behavior. The novel program we have developed to track and analyze student writing has been shown to be a valuable research tool, which will be useful in many future studies. From the preliminary work we have done, we have also established a firm basis for future research into writing to learn.

11.2 Future Directions

Two pilot studies were conducted during the 2005-2006 academic year that were not included in this thesis. One of these involved tracking the writing of long (6 page) papers in the Physics 367 course. This study was problematic for several reasons, including a lack of access to a suitable tracking program (inputlog [7] was used for this study), and a last-minute change in course instructors, which had the effect of reducing access to the students. However, Christopher Manion, the OSU WAC Coordinator, was able to make generalizations about the students approaches to the paper based on the tracking. His assessment was compared to information obtained directly from the two main volunteers through interviews. This study would be useful to repeat under more favorable circumstances. Another study involved giving
feedback on video summaries made by students in Physics 104. The idea behind this study was to test whether feedback alone would impact student writing; students received points only for submitting the summaries, not for their quality. This was a small scale study and the results have yet to be analyzed.

In the future, advances to the tracking program and associated analysis techniques will help us to look more closely at student writing behavior. For example, we have added a database that allows students to access their essays in order for them to be able to write them in a more natural fashion, which will, we hope, eliminate the student discomfort experienced with the first version of the program. We may also add a place for students to pause the program explicitly (when they need to interact with others or pay attention to the TV, for example), eliminating false readings for student writing pauses. This will allow greater trust of our data on when and for how long students pause. This addition might alter student behavior, since it draws attention to the tracking. Students may also forget to use it, leading to the same limitations currently observed. It would have to be tested carefully.

In addition to those simple changes, we will work toward automating the process of analyzing the sequence of revision events in order to obtain information such as we graphed in the case studies. We would also like to work toward correlating revision events with essay content, which will require making an automated process for understanding text within our program. Programs for summarizing and analyzing text written by cognitive linguists have become quite sophisticated as shown in Refs. [64, 78]; this is possible beyond a first step of simply asking the program to look for key words. Eventually, it may even be possible to have judgments of the quality and depth of revisions be automated.
These changes would give us a fairly complete picture of student writing behavior with the exception of understanding pre-writing activities. Combining the tracking program with other research methods such as the think-aloud protocol could make a future study more complete. However, care would have to be taken to not interfere with the students’ natural writing habits. In addition to these improvements to our research methodology, we plan to do future studies including more explicit writing instruction to see whether this impacts writing behaviors. We would also like to test different types of writing and different populations of students to better understand how robust the characterized student behaviors are.

Once we have a more complete understanding of how to analyze student writing, we can focus on the broader question of whether writing impacts learning. The techniques can be combined with studies such as our earlier ones tying the writing to more traditional measures of content knowledge. Do students learn to write in more reflective fashions? Do students exhibit more expert-like behaviors with practice? Do the expert behaviors correlate with better content knowledge? Do students whose behaviors improve more show more improvement on diagnostic tests? The question of writing to learn is complex and there are many variables that limit the ability to reach strong conclusions. Our research methodologies will aid in attacking these broader questions with a firm scientific method.
ESSAY ASSIGNMENTS AND LESSON PLANS FOR SUMMER QUARTER

A.1 Essay and Prompt for the First Assignment

This assignment was not included in the analysis of this thesis due to the fact we had students work on different paragraphs within the same essay.

Instructions: The purpose of the following paper is to explain why people can’t walk through walls: that is, to explain the physics that prevent atoms from occupying the same space. It includes concepts you learned in class, concepts you used in lab, and observations from your lab experiments to explain this phenomenon. One of the paragraphs of the essay has been removed. This is marked within the text. Next to the cut paragraph is a prompt indicating what is missing from the essay. Your assignment is to read the essay and the prompt and write a paragraph to provide the information needed to complete the essay. Be sure to go back and see if your paragraph fits in with the overall context of the paper. The final essay, with your additions, needs to be coherent and unified, having smooth transitions between all the paragraphs. Feel free to ask for clarification on the physics from your TA.

Why One Cannot Walk Through Walls By Lars Schweidenback
In Bohr’s atomic model, electrons exist in quantized energy states in “shells” around the nucleus of the atom. The majority of the atom is filled with empty space; the nucleus of an atom is roughly \(10^{15}\) meters in diameter, and the electrons are much smaller, yet a typical atom has a radius of roughly \(10^{10}\) meters. With so much empty space it seems possible that atoms could share the same space. If two hydrogen atoms were completely overlapping, a quick calculation yields a probability for the nuclei of the atoms colliding to be around \(0.0000001\) percent. This seems to suggest that any atom could overlap with any other atom without difficulty. Therefore, since all matter is thought to be made of atoms, one should be able to push any piece of matter until it overlaps with any other piece of matter. This would allow people to have the ability to walk through walls. Anyone who has tried this realizes that this is not the case, so there must something that prevents atoms from overlapping. So why, then, is the ability to pass through walls left only for the "X-Men"?

**Prompt for paragraph 2:** Use observations from lab regarding Coulomb repulsion to explain why it is not possible to push atoms, which contain charged protons (+) and electrons (-), on top of each other.

Instructor’s response for prompt for paragraph 2: In lab, we observed that charges of the same type repel each other, for example positive charges repel other positive charges. The electrons in atoms are negatively charged, and in the Bohr model, the outer “shell” of each atom consists of electrons. Therefore, when one atom is brought near another atom, the electrons on the outer rims of the two atoms repel one another. The force between the outer electrons that pushes the atoms away from one another such that they will not overlap is called the Coulomb force. In fact the Coulomb force is so strong that if you take two electrons and put them “barely touching” so
the distance between them is the size of an atomic nucleus \(10^{15}\) meters, the repulsive force between them would be over 200 newtons.

**Prompt for paragraph 3:** Paragraph 2 suggests that the outer electrons in atoms are most responsible for interactions. Use observations from lab about the behavior of neutral conductors near charge objects to explain the consequences of this suggestion on how your body would behave near a wall.

Instructor’s response to prompt for paragraph 3: Another thing we learned in lab is that both positive and negative charges attract neutral objects by the process of polarization. In lab we also learned that conductors are materials with the property that electrons can move within them. This makes them easily polarized: in the presence of a positively (negatively) charged object, the electrons will move towards (away from) the end of the conductor closest to the charge. We observed that our own bodies attracted the charged pieces of tape and thus concluded our bodies are conductors. If the outer electrons of the wall are responsible for interactions with external objects, we, as conductors, should be attracted to the wall. However, outside of Velcro-enhanced carnival games, I have not observed people being attracted to walls, so our picture of the wall must be incomplete. So why aren’t people attracted to walls? This is because overall atoms are typically neutral, protons within the nucleus carry positive charge and the electrons carry negative charge. The total amount of charge within an atom is typically zero.

**Prompt for paragraph 4:** Paragraph 3 reminds us that atoms are typically neutral while the argument in paragraph 2 relied on the idea that only the electrons interact. This paragraph should address and resolve
this apparent contradiction by applying the inverse square proportionality between the Coulomb force and distance.

Instructor’s response to prompt for paragraph 4: The Coulomb force is inversely proportional to the square of the distance between the charged objects. This means that as the distance between the objects increases, the force between them strongly decreases. Consider two atoms near each other. The outer shell electron of one atom will be much closer to the outer shell electron of the other atom than to the other atom’s protons which are contained in the nucleus. Therefore, the repulsive force between the two electrons will be much stronger than the attractive force between the electron and the protons. The argument given for not being able to overlap one atom with another still holds, however the repulsive force would be a bit weaker than estimated due to the additional small attractive force previously neglected. However, when we are standing some distance away from a wall (we could even be very close, just so we are more than a few atomic distances away), the electrons and the protons of atoms at the surface of the wall appear to be essentially equal distance from our bodies. Electrons in our body would be pushed away from the wall due to the electrons in the wall; however the electrons in our body would be pulled towards the wall equally strongly by the force due to the protons in the wall. Since we are far from the wall, we see the overall property that the wall is neutral.

Our hypothetical wall is made of atoms which contain both positive and negative charges. From the Bohr model we have a model of the atom where the negatively charged electrons exist in “shells” around the positive protons in the nucleus. Since the electrons are more towards the outer edges of the atoms, the repulsive forces between the electrons in neighboring atoms prevents them from being pushed into
each other. However, the wall overall has no net charge. Even though the electrons are more towards the edges, the wall is neutral because the Coulomb force tells us that far from the wall, the difference in the distance between us and the electrons and the distance between us and the protons in the wall is negligible. Thus, it follows that force of attraction and repulsion will balance out and we will not feel any net attraction to the wall. This picture of the wall allows us to understand why we cannot walk through walls while being consistent with our other observation that we are not attracted to walls. Though an interesting fantasy, walking through walls will have to be left to fictional characters.

A.2 Essay and Prompt for the Second Assignment

Instructions: The purpose of the following paper is to explain why it is safe to be inside a car during lightning. It includes concepts you learned in class and concepts you used in lab. One of the paragraphs of the essay has been removed. This is marked within the text. Next to the cut paragraph is a prompt indicating what is missing from the essay. Your assignment is to read the essay and the prompt and write a paragraph to provide the information needed to complete the essay. Be sure to go back and see if your paragraph fits in with the overall context of the paper. The final essay, with your additions, needs to be coherent and unified, having smooth transitions between all the paragraphs. Feel free to ask for clarification on the physics from your TA.

Why are Car Passengers Safe During Lightning? By Lars Schweidenback

It is often said that one of the safest place to be when out in a thunderstorm is inside a car. It is known, however, that lightning, as a form of electricity, is attracted
to conducting objects, and cars are made of metal, which is highly conducting. Therefore, it would seem that lightning would be attracted to a car. This seems to be a hazardous situation for the car passengers, so why is it said that a car is safe? The answer can be found by examining Gauss’s law and the conductivity of metal.

First, one must grasp what lightning is in order to understand how it will behave when it strikes a car. During storms, air currents push positively charged ions towards the top portion of clouds, and negatively charged ions accumulate near the base of the clouds (the exact mechanism for this process is unknown). When the positive and negative charges are separated, there is an electric field between them. When this field is strong enough, it ionizes (splits into positive and negative parts) the air beneath it and a channel of excess electrons forms. Some of the excess electric charges then “ride” down this channel of electrons, completing a circuit to the ground and dissipating the excess charge. Simply put, lightning is just bolts of negatively charged electrons.

Since cars are primarily made of metal they are good conductors. The resistance of a metal conductor is much less than that of air, even when the air is highly ionized. Electricity prefers to follow the path of least resistance, and there will be less resistance if the lightning follows a path through a car instead of through the air. Therefore, the flow of charge in lightning will be more likely to pass through the car to the ground than the surrounding air to the ground. This same idea is the basis for lightning rods. They offer the electricity a path of lesser resistance to the ground than the walls of the house thus, protecting the house. So if lighting is more likely to hit a car, what prevents people inside from being unsafe during a storm?
If lightning strikes a car, the electrons will flow down the side of the car to the ground. This flow of electrons could be unsafe to passengers of a car. However, cars are conductors, this means that the many electrons already in the car are free to move. When an electric field is present near a conductor, these charges feel an electric force, and move until this force is minimized. This occurs when the charges are far apart, so the charges move to the surface of the conductor, where they can spread out. Another consequence of this redistribution of charge is that the electric field inside the conductor is zero. During a lightning strike these fields change very rapidly, and the car would have to be a perfect conductor (or the electrons move infinitely quickly) in order to completely shield the fields. Nevertheless, the electrons can move fast enough so that conducting shells are effective in canceling out most electric fields, even during lightning.

**Prompt:** We can also use Gauss’ Law to explain why the person inside the car is safe (safe = feels no electric field). Write a paragraph here that does so, describing the electric flux within conductors in general and within the car specifically.

The shielding effect of a conducting shell has many uses besides protecting people from lightning. Many highly sensitive experiments are done inside a metal room, called a Faraday cage, so that external electric fields will not affect the experiment. Similar shielding is also commonly used to protect electronic devices. Although in class, we only consider spherical and cylindrical shells due to the simplification of the mathematics involved, the idea of shielding is much more general and applies to a metal shell of any shape. Your car will act as a Faraday cage such that fields on the
outside of the “cage” are shielded from the inside, making a car’s passengers safe even during a lightning strike.

A.3 Essay and Prompt for the Third Assignment

Instructions: The purpose of the following paper is to explain how electrostatic precipitators clean exhaust gas. It includes concepts you learned in class, concepts you used in lab, and observations from your lab experiments to explain this phenomenon. One of the paragraphs of the essay has been removed. This is marked within the text. Next to the cut paragraph is a prompt indicating what is missing from the essay. Your assignment is to read the essay and the prompt and write a paragraph to provide the information needed to complete the essay. Be sure to go back and see if your paragraph fits in with the overall context of the paper. The final essay, with your additions, needs to be coherent and unified, having smooth transitions between all the paragraphs. Feel free to ask for clarification on the physics from your TA.

Applying Potential Differences to Clean Exhaust Gas (Adapted from [8], image from [9].)

Many industrial operations produce exhaust gases that contain dust, fly ash (unburned constituents from burning), fumes (fine elemental particles such as cadmium, sulfur and lead) and mist (such as coal tar), which the Environmental Protection Agency and scientists around the world feel are bad for both human health and the environment. One widely used method of removing such contaminants from exhaust gas is to use an electrostatic precipitator. This is a technology dating back to 1912, patented as a method of charging particles and then collecting them through electrostatic attraction (i.e. Coulomb forces between charged particles). Modern versions of
electrostatic precipitators can be up to 99% effective at removing particulate matter from exhaust gas.

In an electrostatic precipitator, the gas is forced to move through a passageway containing a series of electrodes, which either collect or emit electric charges. The precipitator contains both positively and negatively charged electrodes. The positively charged electrodes are called the discharge electrodes. The negatively charged electrodes are called the collector plates, since they are responsible for collecting the polluting particles. The collector plates are charged to several thousand volts (negative with respect to the discharge plate) with high direct voltage. The technology depends upon the polluting particles having outer-shell electrons that can be stripped off relatively easily by a large electric field, a process called ionization. An image of an electrostatic precipitator is given in Fig. A.1.

Prompt: Removing the particles is a multi-step process. First some outer electrons are removed. The now positively charged particle is then collected by the collector plate. This can be explained either by Coulomb’s law or invoking Electric fields. Write a paragraph explaining how the positively charged discharge plate strips off electrons and give TWO explanations for why the particle is drawn to the collector plate. Your explanation should include both general descriptions (for example, why positively charged particles in general might be attracted to the collector plate) as well as specifics (references to particulate contaminants).

The electrostatic precipitator removes contaminant particles from the gas flow as follows. When the particles in the gas flow pass near the discharge plate, their
Figure A.1: Electrostatic Precipitator

electrons experience a large attractive Coulomb force, since the electrons are negatively charged and the discharge plate is positively charged. Because some electrons are more loosely bound, this force can ionize the suspended particles, leaving them with a net positive charge. At this stage, the positively charged particles are present in a high potential difference between the collector and discharge plates. Positively charged particles will experience a force towards lower potential and move towards the collector plate. Another way to understand this is to consider the Coulomb forces these positively charged particles will feel. They will be repelled from the positively charged discharge plate and attracted to the negatively charged collector plate. Due to the attraction between the particles and the collector plate, the charged particles adhere to the collector plate and are thus removed from the exhaust gas.
An application of an electrostatic precipitator is for cleaning exhaust fumes in coke oven facilities. In this type of facility, coal is burned at high temperature to drive out impurities in order to make coke for producing iron. One of the byproducts of making coke is coke oven gas, which has coal tar particles trapped in it. To collect the coal tar particles, the gas is sent through an electrostatic precipitator, which charges the tar droplets so they are attracted to the collecting electrodes. Because the precipitated material is liquid, it collects and consolidates until there is enough weight for it to flow down the sides of the precipitator. Cleaner exhaust can be released into the atmosphere, and the liquid remains can be collected and disposed of.

A.4 Essay and Prompt for the Fourth Assignment

Instructions: The purpose of the following paper is to explain the wiring used in strands of mini-lights (common Christmas tree lights). It includes concepts you learned in class, concepts you used in lab, and observations from your lab experiments to explain this phenomenon. One of the paragraphs of the essay has been removed. This is marked within the text. Next to the cut paragraph is a prompt indicating what is missing from the essay. Your assignment is to read the essay and the prompt and write a paragraph to provide the information needed to complete the essay. Be sure to go back and see if your paragraph fits in with the overall context of the paper. The final essay, with your additions, needs to be coherent and unified, having smooth transitions between all the paragraphs. Feel free to ask for clarification on the physics from your TA.

The Circuitry of Mini-Light Strands
Mini-lights, introduced in the 1970s, dominate the market when it comes to strands of lights. However, one common drawback of mini-lights is the annoyance of many bulbs going out when there is a problem with just one light bulb. If the mini-lights were wired in parallel, this would not be a problem. If all of the bulbs in the strand were connected in parallel, each bulb would be directly wired to the voltage source, so if any one bulb were removed the other bulbs would remain lit. It seems the simple solution to having hassle-free light strands would be to wire the bulbs in parallel. However, mini-lights are 2.5 V incandescent bulbs (meaning that at higher voltages the filament will break), while household outlets produce 120 volts. Since there are no transformers at the start of the light strand to change the voltage, and each bulb in parallel would see the full voltage, we know the voltage across each bulb would be 120 V. The bulbs cannot be wired in parallel. However, in series, the voltage would be divided among each bulb, so they must be wired in series to keep the voltage low. In fact, we find that 120/2.5 = 48, so it is optimal to have 48 mini-lights to one strand. However, people like round numbers, so typically strands contain 50 lights, which only slightly dims the brightness of each bulb in the strand (each bulb receives 2.4 V).

Prompt: If all bulbs are in series, that means if one bulb burns out, all bulbs will go out. However, have you ever observed that if one bulb is removed from the strand, only a portion of the strand goes out? This paragraph should explain these effects, including the wiring scheme necessary to explain this second effect (consider the fact that consumers often want more than 50 bulbs, so bulbs are often sold in strands of 100, 150, 200, etc). Include a circuit diagram.
When mini-lights were first introduced, any bulb burning out would darken the entire strand. Today, the bulbs can burn out and the strand will stay lit, but if you pop one of the bulbs out of its socket, the whole strand will go dark. The difference in behavior occurs because bulbs now commonly sold in mini-light strands contain an internal shunt, as shown in Fig. A.2. The filament is the thin, bent wire connected between the two posts. If you look closely, you can see a thick, straight wire connected between the posts. This is the shunt wire; it contains a coating that gives it fairly high resistance. Because the resistance in the shunt is high, most of the current will flow through the filament and the shunt will not effect the operation of the bulb. However, when the filament burns out (breaks), all current will be directed through the shunt. The current in the strand remains un-interrupted. Because it is not advantageous to have a large resistor in the strand, the coating on the shunt is designed to quickly burn off from the heat caused by extra current flow after the filament breaks.

![Figure A.2: Mini-light Shunt](image)

Another interesting feature of modern mini-light strands is blinking lights. There are two different techniques that are used to create blinking lights. One method involves the installation of a special blinker bulb at any position in the strand. The
blinker bulb contains an extra strip of metal at the top which is bi-metallic (the strip is made of two different types of metal, one on each side). When the bulb is lit, current runs between the posts through the filament and the strip. As current flows, the strip gets hot. The two metals are chosen such that one will contract significantly more when heated, thus the strip will bend. The strip bends away from the post, breaking the current flow and extinguishing the bulb. As the strip cools it bends back, reconnects to the post and re-lights the filament, then the cycle repeats.

Whenever this blinker bulb is not lit, current in the series circuit is interrupted, so the rest of the bulbs will also not be lit; this simple mechanism keeps all of the bulbs blinking in unison. Obviously, these bulbs cannot have a shunt, for if they did, the rest of the strand would not blink. The other method for creating blinking lights is more sophisticated: light sets now come with 16-function controllers that can run the lights in all sorts of interesting patterns. In these systems, you typically find a controller box that is driving four separate strands of mini-bulbs. The four strands are interleaved rather than being one-after-the-other.

Information and image for this essay were taken from [10]

A.5 Essay and Prompt for the Fifth Assignment

Instructions: The purpose of the following paper is to explain how particles from the solar wind are trapped in the Earth’s magnetic field. It includes concepts you learned in class, concepts you used in lab, and observations from your lab experiments to explain this phenomenon. One of the paragraphs of the essay has been removed. This is marked within the text. Next to the cut paragraph is a prompt indicating what is missing from the essay. Your assignment is to read the essay and the prompt
and write a paragraph to provide the information needed to complete the essay. Be sure to go back and see if your paragraph fits in with the overall context of the paper. The final essay, with your additions, needs to be coherent and unified, having smooth transitions between all the paragraphs. Feel free to ask for clarification on the physics from your TA.

**Trapping the Solar Winds in the Earth’s Magnetic Field**

Reactions near the surface of the Sun produce particles which are radiated outwards; these particles are called the solar winds. Since many of these particles have a net charge, they can be trapped by the Earth’s magnetic field. Trapped particles exhibit three periodic motions which take place simultaneously: a fast gyration in the Earth’s magnetic field (typically thousands of times each second), a slower back-and-forth “mirroring” along the field (typically lasting 1/10 second), and a slow drift around the magnetic axis of the Earth, (typical time to circle the Earth is a few minutes).

**Prompt:** Explain the mechanism for trapping the solar winds into a path of fast gyrations in the earth’s magnetic field. Use the Lorentz force law (assume $E = 0$) to explain this motion (a diagram will help), and use the work-energy theorem to explain how these paths are sustained without inputting energy.

As the particles circle, the axis of their rotation generally shifts because the Earth does not have a constant magnetic field; this creates a spiral pattern. However, a subtle interaction causes the spiraling particle to be repelled from regions of stronger magnetic field, for instance near the Earth’s poles where the field lines converge. The regions of stronger field cause the particle to turn back or “mirror.” Without this
“mirroring,” ions and electrons would not be trapped in the Earth’s magnetosphere, but would instead follow field lines into the atmosphere, where they would be absorbed and lost. What happens instead is that every time a trapped particle approaches Earth, it is reflected back, and is thus confined. In addition to the rapid gyration and the “mirror” motion, trapped particles also undergo a slow “drift” across the magnetic field slightly rotated around the Earth’s magnetic axis. Viewed from the North Pole, a positive ion will gradually rotate clockwise, a negative electron counter-clockwise. These motions are depicted in the images below.

Because positive and negative ions drift in opposite directions, their motion creates a net electric current that circulates clockwise around the Earth when viewed from north. The current is called the ring current. The flow of the ring current (as with many currents in space) needs no energy input, but persists as long as charged particles are trapped in the magnetic field. Many aspects of such currents are hard to reproduce in the laboratory, and the magnetosphere is considered our best “natural laboratory” for studying the processes of distant space.

Information for this essay was taken from [11]

A.6 First Explicit Writing Instruction

Topic: What are the characteristics of a good paragraph?

1. Introduce myself

2. Discuss why are we talking about writing in a physics classroom

   (a) Writing helps people learn
(b) People in the scientific colleges are getting feedback from their alumni—wish they had more writing, specifically writing in their fields.

(c) People use writing. Dad and Pat: writes summaries of experiments, PowerPoint presentations, and emails to engineers who know what he is talking about and people in say marketing who might not.

3. How did the paragraphs go? What problems or challenges did you encounter with them? Do you have any questions about them?

4. Characteristics of a good paragraph

(a) Has all of the information that the prompt asks for

(b) Reread the prompt before your turn your paragraph in.

(c) Do some “prewriting” exercises to help organize your thoughts before you start writing: outlines, thought maps, or just jotting down ideas.

(d) Clear scientific language: Formal, Detailed

(e) Fits in the entire essay (flow)

(f) Make sure your introduction is appropriate—not too general for the, say, second paragraph

(g) Example: Atoms are made up of electrons, protons, and a nucleus. Not a good introduction because this was already stated in the first paragraph and it is super general (a lot of people already know this)

(h) Sentences fit together in the paragraph:

(i) Backs up generalization with specific examples
(j) For example: Human bodies are conductors.

(k) Pretend that the person reading this paragraph, understands a bit about the topic, but need more information.

5. Type up one of the bad paragraphs from another class and re-distribute the essays. Have them get in groups and come up with 3 things (based on this list) that the author could do to make this paragraph better.

A.7 Second Explicit Writing Instruction

Topic: Argumentation

Lecture: What is argument?

Ask students: Why does Dedra keep using the term physics argument to talk about these papers? Is the goal of the paper to explain something people already know? Or is it to prove a point about something in physics? (Mention too that people know this physics, but that when they get into the real world and work, they will be entering new territory and need to convince people of their ideas. You have to own your material).

The point of the paper is to prove a point. In writing, the more technical term for this is argument. So the purpose of all of these papers is to construct a physics argument, i.e., to prove a point about physics.

What is the purpose of the lightning paper? If it is an argument, what is it trying to argue? How does it do that?

The main argument of a paper is called a thesis. It is just a statement of argument. So for this paper, the thesis, what the author is trying to prove, is it is safe to be in cars when there is lightning.
All arguments need support. Each paragraph provides some sort of support for the argument. Go through the paragraphs and ask students what each individual paragraph is about.

1. Introduction

2. Describes lightening

3. Describes metal as a conductor of electricity

4. Describes the flow of lightning in a car

5. Uses Gauss’ law to describe the flow of lightning in a car

6. Conclusion

This is an outline of the paper. You can see the line of argument: what support the author is using to prove his/her point. So each paper has thesis, a statement about what it is arguing. So does each paragraph. This is called a topic sentence. It is the statement of argument for each paragraph.

For example, with paragraph 2: What is the argument? What is the author trying to get across? Argument–Lightning works a certain way.

1. Air currents push ions toward clouds

2. Negative ones go near the base

3. This creates an electric field

4. The field gets strong and ionizes (splits into positive and negative parts)

5. Excess charges ride down, completing the circuit
I can go through my paragraph here and connect each of these ideas to a sentence.

Activity: Hand the students back their paragraphs as well as a copy of the previous paper. In groups have them construct a discursive outline (one that uses fragments). The outline would start with their argument for that paragraph, i.e., their topic sentence, and then their support for that argument (items from lecture, lab, etc.). Then have them go through their old paragraphs and examine each one of their sentences. Each sentence should match up with a number on the outline.

If possible show them an example from last week’s lab.

The purpose of this exercise is to explain to students that this is an argument, not expository or explanatory writing where they are repeating facts. They are repeating things they learned to prove a point about something in physics.

A.8 Third Explicit Writing Instruction

(Be sure to emphasize to the students that the grading, which is independent, comes out very close. In almost all cases within one point. This shows that without clear language the physics is not clear: Cat’s input is important.)

Topic: Sentence level problems. Begin by telling students that the paragraphs in general have been getting much better. Is the exercise going better? Last week we worked on paragraph-wide problems, like organizations. Today we will be working on things you can do to make the individual sentences better.

Transitions: are words that serve as bridges or connectors between sentences, part of sentences or paragraphs. Why do you need to use transitions?: It is important to use transitions to let your reader see the relationships between different ideas. People will be able to understand your writing better if these relationship between ideas
are clearer. Good example: “...the particles are attracted to the positively charged electrode. Now the particles are positively charged.” The “now” really shows the cause and effect relationship between these two ideas. Other thing transitions can do is serve as a road map and let your reader know where you are taking them in the paper. For example, in this paper, you needed to give two explanations of the same processes. The paragraphs that were the easiest to read (for both me and Dedra) were the ones where the authors clearly marked in the text, using key words, that they had two different explanations for the same thing. Another paragraph contains the following phrases: There are two reasons for this./ The first... / This can also be explained by...” This immediately lets the reader know that a) he understood the prompt and b) there are two explanations for the same phenomenon. Equations: When you use an equation in a piece of scientific writing, you need to:

1) describe what the equation is.

2) Explain the relevant terms.

3) Apply it to the context at hand.

An example: “Coulomb’s law states $F = kq_1q_2/r^2$; $F$ is the force of attraction between two point charges $q_1$ and $q_2$ which in our case are the dust and [the charged plate].” He has described what the equation is, he has explained the relevant terms, and he has placed them in the context of the paragraph.

Concrete Language: Good writing, especially in scientific writing where you are doing a lot of observations and using a lot of jargon and technical language needs to be clear, concise, and correct. As far as conciseness goes, Dedra told me that when she is revising a paper she cuts out all unnecessary words. Last time she cut so many words out her 5 page paper turned into a 4 page paper. A good example of concise
writing: “... causing the particles to have a net positive charge. The collector plate is negatively charged, so the positively charged particle is attracted (by Coulomb’s law).” He described what was happening using formal, scientific language, with no extra words.

Observation / Explanation: Science is all about observing things and trying to explain them: observation and explanation. When you are writing these paragraphs, you need to describe what happens and then say why. That is what a lot of these paragraphs are asking you to do. Say what is happening and then say why it is happening. Good paragraphs need BOTH of these things. A good example of observation and explanation: “The positively charged particles then travel [near] the negatively charged collector plate and the plate collects the particles because positive and negative charges attract.” One thing to think about when writing these paragraphs is constructing your sentences very simply (as your writing skills develop you can write more complex things). As you are starting out, think about using a simple sentence structure of subject + verb. What is it? What is it doing?

Activity: Have each group of students work on a sentence that has a problem with something discussed above and make the sentence better.

Transitions:

There are two reasons for this. The first... This can also be explained by...

...the particles are attracted to the positively charged electrode. Now the particles are positively charged.

Equations:

Coulomb’s law states $F = kq_1q_2/r^2$, $F$ is the force of attraction between two point charges $q_1$ and $q_2$, which in our case are the dust and [the charged plate]
Concrete Language:

... causing the particles to have a net positive charge. The collector plate is negatively charged, so the positively charged particle is attracted (by Coulomb’s law).

Observation / Explanation:

The positively charged particles then travel through the negatively charged collector plate and the plate collects the particles because positive and negative charges attract.

A.9 Fourth Explicit Writing Instruction

Topic: Word choice.

Begin by mentioning the paragraphs are still improving. Overall the language is clearer. Ask students if the exercise is going better. Two weeks ago we worked on making your whole paragraphs better. Last week, we looked at things you can do to make the individual sentences better. Today we will be working on word choice. Even individual word choices can have a huge impact on your reader; confusing, inaccurate, or incorrect word choices can misconstrue the meaning of your sentences. As we did last time, the topics we are going to cover will be on the sheet and we will break into small groups to do an activity.

Word Choice

Use the verbs, adverbs, and prepositions of physics terminology: (The teaching assistant will cover this at the start of class to clarify the physics). It is important to use these terms because they are the language of physics, the way physicists have decided to describe the world around them and in physics class you need to use this language when you write. If you don’t use these words or groups of words, you will
appear not to know what you are talking about. Physics terminology: “current flows through,” “voltage applied across.” Scientific language: “as if,” “magically,” “out of nowhere,” versus better choices: “better,” “is,” “seems,” “appears.”

Use neutral (not emotive [bringing out emotions] or anthropomorphic [giving objects the characteristics of people]) words: Pick neutral and accurate words to describe situations. Also all of the things you are talking about in class are inanimate objects, not people. Don’t talk about them like they were people. For example: “burnt bulbs are disastrous” can be “burnt bulbs are an irritation,” “Strong charges enjoy interacting more than weak charges” can be “the force of interaction is directly dependent on the strength of the charge.”

Know when to use definite or indefinite language. Definite language should be used to describe physical processes, especially when there is no question about what is happening. As you advance in your studies, you might need to use indefinite language to talk about processes, but I don’t think any of the things described in the paragraphs are processes in gray areas of knowledge. “When one of the bulbs in a series circuit burns out or is removed, the entire strand might go out” can be “When one of the bulbs in a series circuit burns out or is removed, the entire strand will not light.” It is better to use indefinite language when voicing your opinions or drawing conclusions. It should also be used for some observations: “our data prove” is better as “our data indicate,” and “We know that the atoms interact with the surface purely by the Van der Waals interaction” is better as “The dominant interaction between the atoms and the surface is the Van der Waals interaction.”

Class Activity

1. Physics terminology: Fill in the blank with the correct physics term.
a) flux ______ through
b) electric field ______ from
c) charge ______ with
d) magnetic field is ______

2. Neutral words: Rewrite the sentences. (Two groups can do this)

   a) The cart wants to roll down hill.

   b) If the electrons cannot go through the entire pathway they don’t even bother trying at all.

3. Definite/Indefinite Language: What word sets off a red flag? Fix the sentence to get rid of that word.

   “The ball always lands 4 m from the launching pad”
APPENDIX B

ESSAY ASSIGNMENTS AND LESSON PLANS FOR AUTUMN QUARTER

B.1 Essay and Prompt for the First Assignment

For the two circuits shown below (see Fig. B.1), explain in clear and general terms how to determine the relative brightness of the bulbs within a given circuit, as well as how the brightness changes when the switch is closed (for circuit 1). To do this, you will need to explicitly mention what circuit parameter(s) determine the brightness of a bulb, and explain what factors affect that (those) parameter(s). You should use the concepts developed in the lab regarding current conservation, and current flow and resistance in parallel and series circuits. Your explanations should be general and clear enough that someone else could use them to determine the relative brightness of bulbs in a similar circuit. Apply your general explanations to the circuits below explicitly determining the relative brightness of each labeled bulb. For example: “any bulb connected directly across the battery will be the brightest; therefore, in circuit 1 bulb 1 will be the brightest.” Make use of the labeled names for the bulbs and junctions so your references are clear to the reader.
B.2 Essay and Prompt for the Second Assignment

For the two circuits shown below (see Fig. B.2), explain in clear and general terms how to qualitatively determine the voltage drop across each element in the circuit, and use this information to determine the change in brightness of bulbs over time in simple bulb and capacitor circuits. (You can discuss the voltage in terms of the color coding scheme from lab if you want to, but it is not required.) You should use the concepts developed in the lab regarding how to determine voltage drop across circuit elements, and current flow and potential while charging and discharging capacitors. Your explanations should be general and clear enough that someone else could use them to determine the voltage drop and bulb brightness in similar circuits. Apply your general explanations to the circuits below qualitatively indicating the voltage drop across each element and explicitly determining the brightness of each bulb at each labeled time. (Consider all wires in the circuits to be ideal, meaning there will be no voltage drop across them.)

Figure B.1: Two circuits for first writing prompt, Autumn quarter.
Figure B.2: Two circuits for second writing prompt, Autumn quarter.
B.3 Essay and Prompt for the Third Assignment

Write a general set of instructions for applying Kirchhoff’s rules. Include a definition and explanation for each of the two laws. The instructions should be clear enough that other people can use them to solve for all currents in any simple battery and resistor circuit. You need to explain specifics, such as what is the voltage drop across a resistor (hint: going either way around the loop!) Also include explanations for why the steps work.

B.4 First Explicit Writing Instruction

Lesson 1: Generalities and Specifics

Introduce myself. Introduction: why are we talking about writing in a physics classroom? Writing helps people learn: writing about content knowledge provides another chance for the brain to strengthen neural connections, writing makes use of different learning styles, strong correlation between writing and understanding; students who don’t understand a concept usually cannot write clearly or logically. People in the scientific colleges are getting feedback from their alumni, wishing they had more writing, specifically writing in their fields. People use writing: papers on experiments, presentations, grant proposals, journal articles, summaries of experiments, PowerPoint presentations, emails to disseminate information, for instance, to marketing departments.

We’ll be thinking about writing like an upside down triangle: we’ll start off talking about general concepts you’ll need to know for this assignment, then move on to specific techniques for working on an essay, and lastly talk about sentence level techniques: things like word choice, transitions, equations, etc.
For this assignment you’ll need to understand the relationship between generalization and specifics. You’ll need to look at specific pieces of information, in order to make a generalization (a large claim). But if you make a generalization, you’ll need to back it up with specific information. (Examples were discussed.)

Writing Instruction Activity - Circuits I

Purpose: practice building generalities from specifics and using specifics to back up generalities.

For example: General statement: Newton’s third law states that if one object exerts a force on another object, the other object will exert an equal and opposite force on the first object. Specific statement: When I push on the wall with my hand, the wall exerts an equal and opposite push on my hand.

Given a specific example, site a general rule:

1. In lab, we observed that a charge of +2 \( \mu \text{C} \) repelled a charge of +4 \( \mu \text{C} \), and in contrast, it attracted a charge of -4 \( \mu \text{C} \).

2. 5 m from a point charge the measured electric field has a strength of 8 N/C, but 10 m from the charge the electric field has a strength of 2 N/C.

Given a general rule, write a sentence citing a specific example:

1. If the net charge inside a closed surface is zero, the flux through the surface will also be zero.

2. If a charged particle is placed in an electric field it experiences a force equal to the product of its charge and the electric field strength.
B.5 Second Explicit Writing Instruction

Lesson Plan 2: Organization

Last class we talked about the general concepts you’ll need to complete these exercises. Today we are going to give you some suggestions to improve on the body of your essays. We looked over a lot of the essays, and for the most part they are pretty good. But they all can be improved. Here is one of the essays from last week. Is it clear? Go to each table and ask the students what could be done to make it clearer.

Most of the papers were better than this. But there are two main things the student needs to do improve the paper. First of all, it needs to be written in paragraph form; that should be one of our goals for today—to get through the physics AND to write a paragraph instead of just an outline. Secondly, the ideas are not organized well: the student presented many specific examples and some general physics concepts, but didn’t create any relationships between the two. If you weren’t doing this lab, you wouldn’t understand what he/she was talking about.

Keep in mind: there are places in this essay where you can’t understand what he said. There are probably places in your paragraph where an outside reader like me and Dedra can’t understand what you are saying. Write so that other people can understand your ideas.

We are going to give you a general scheme for thinking about how to organize these paragraphs. Then we are going to do an exercise to work that practices this scheme and works on making a clear relationship between ideas.

This builds on what we talked about last week about generalizations and specifics.
Start with a very general statement: cite a very general physics rule.

Move to something less specific: specifically, how the rule applies to the particular circuit

Move to something specific: how the circuit rule affects a certain bulb

This organization scheme makes clear the connections between the ideas and it helps your reader understand how one physics rule can be applied to circuits, and then how it affects bulbs. Hopefully, this will also make it easier to write.

For the writing activity today, we want you to take the content (not the sentences themselves) from the three circled sentences and write a sentence or two (at the most) that organizes it in the way we suggest. Think about the relationship between the different ideas in the content. You will need to add outside information to your sentences to tie the content together.

Sentences from Handout (the circled parts are indicated here by bold font):

“Brightness depends on power, which is represented by current multiplied by voltage. Bulb 1 will be the brightest, then bulb 2, and bulb 3 and 5 will be the most dim. Bulb 2 will be 4 times as bright as 3 or 5. Resistance from A1 to C1 is 1.5x.5 than from A1 to D1.”

This is NOT the writing activity for today. Write this as if it were a part of a larger paragraph. It should only be around 3 sentences. Remember that Dedra is here to help you with the physics. If you have questions, please ask her.

B.6 Third Explicit Writing Instruction

Lesson Plan 3: Instructions
DEDRA: Before we start on today’s assignment we have some feedback for you from the previous essay. Although the writing was overall better, and pieces of many essays were quite good, most of the essays contained at least some factual problem or inconsistency. Ideas were contradicted within a single paragraph. This indicates two things: first, you are still struggling with these concepts, and second, you didn’t go back to consider if what you wrote made sense. Today you will need to do the latter, and in order to help with the former, I am providing you with my “solutions” to the previous two writing assignments. Please read them over and make sure the content makes sense to you.

Before we start on today’s writing assignment we have a brief worksheet for you help with some of the problems we saw from last week’s essays. The sentences provided are from student essays (not in this class), and each one has some factual problem or word use issue which is problematic. Correct the sentences by replacing the incorrect word(s) with ones which will make the sentences accurate.

HAVE THEM SPEND 5 minutes on the worksheet

CAT: In the first class, we talked about the general concepts you’ll need to complete these exercises. On the second class we gave you some suggestions to improve on the body of your essays. Today, however, you are doing a slightly different assignment. You are going to have to write specific instructions on how to apply Kirchhoff’s rules to any simple battery and resistor circuit. We are going to give you some suggestions on how to write instructions, which will be slightly different than what we suggested for the last assignment.

An outline might be a good way to think about writing this assignment.
Again, like last week, you will need to start with a general physics ideas: this time in specific start by defining Kirchhoff’s rules

Then you need to write down the specific instructions on how to use the rules to solve for all currents in any simple resistor and battery circuit.

- Use instructive verbs (imperative form), “do this, do that”

- Order will be important, make sure that your instructions are in an order that make sense to another reader. Use words like, “first,” “second,” “third,” and “then,” and “next”

- If you are doing an outline, only write one, or at the most two, instructions per line.

- * Even though you are writing directions, you still need to explain to your reader WHY you are telling them to do things!

- So (before or) after each instruction or set of instructions, where it is appropriate, you need to say why you are going what you are doing. You are going to need to connect your instructions to the physics rules that govern the circuit.

- You need to do this ... because ...

- The organizational scheme for last’s week essay was like an upside triangle. This week’s is more like a pyramid. You need to start with something specific, but connect it to something general.

Still need to connect specifics and generalities like we have discussed on the first day of writing and last week! You need to consider the readability of your instructions, because you are going to give your instructions to another group and they are
going to have to use your instructions to solve for the currents in a circuit. If your instructions are not clear, or are incomplete, they will not be able to do this. So we have been talking all week about the importance of making your writing clear, easy to understand and readable because other people are going to read it, and today you are going to actually have other people read it. You’ll get to see first hand how easy/hard it is to read what your group wrote.

BEFORE we start: tell me, “what is current conservation?”, “what is Kirchhoff’s junction rule?” and “what is Kirchhoff’s loop rule?”

DEDRA: Each person should take 10 minutes (roughly) to write an outline about what information and explanations need to be provided to solve a new circuit.

Then the group should write a single set of instructions (including explanations) which will be used by another group (~10 more minutes)

Then the instructions will be passed to another group who will use them EXACTLY as they are written to try to solve for all currents in a circuit we will provide.

Any group which produces instructions that WORK will get 5 bonus points for today’s lab grade. It is therefore beneficial to look at the instructions critically as if your group finds problems with them you will prevent your classmates from receiving a better grade.

Remember that Dedra is here to help you with the physics. If you have questions, please ask her.

Writing Activity:

Instructions: Correct the following sentences. Some have serious problems; others are minor issues of word choice. The circuit is provided on the right to remind you of the content being discussed. (See Fig. B.2).
1. The capacitor has been fully charged and no longer accepts a voltage drop from the source.

2. The voltage drop across the second bulb is less because a portion of the voltage has to drop across the capacitor.

3. The voltage goes through A then splits and goes through the capacitor.

4. At 0 s, all of the voltage will flow through bulb A.

5. At the junction the capacitor will take all of the voltage.

6. Initially the battery gives off some voltage and is carried through the circuit through a current.

7. Bulb A would have half the current and bulb B would have half the current.

8. A retains all the voltage from the battery and thus retains all the current.
C.1 First Lab Quiz

1. (3 points) Rank each bulb in the following circuit (Fig. C.1) according to their brightness.

2. (2 points) For the circuit given above, describe the current flow between junctions J1 and J2. In specific, compare the current through each branch between those two junctions.

3. (2 points) Is a battery a constant current or a constant voltage source? Explain clearly how you know this (providing an explicit example may make your response more clear).

4. When a bulb is placed in series with another bulb:
   a. (1 point) How will the brightness of the two bulbs compare to each other?
   b. (1 point) Explain your response including reasoning based on current flow.
   c. (1 point) How will the brightness of the first bulb after the 2nd bulb is added compare to before it was added?
d. (1 point) Explain your response including reasoning based on resistance and current.

5. When a bulb is placed in series with another bulb:
   a. (1 point) How will the brightness of the two bulbs compare to each other?
   b. (1 point) Explain your response including reasoning based on current flow.
   c. (1 point) How will the brightness of the first bulb after the 2nd bulb is added compare to before it was added?
   d. (1 point) Explain your response including reasoning based on resistance and current.

C.2 Second Lab Quiz

132 Lab Group Quiz - Circuits II
1. (3 points) Determine how the brightness of each bulb changes from an initial
time just after the circuit is connected to a final time a long time later (~10 seconds).
(The capacitors have equal capacitance and are initially uncharged.) See Fig. C.2.

2. (3 points) For the circuit above, explain the charge accumulation on both plates
of each capacitor from immediately after the circuit is connected to a long time later
(~10 seconds).

3. a) (1 point) For the circuit shown below, compare the direction and magnitude
of the current through bulb B with the current through bulbs A and C. See Fig. C.3.

   b) (2 points) Explain your answer to 3a by considering the charge accumulation
on each capacitor plate.
4. (3 points) What will be the voltage across the capacitor in the following circuit a long time after the switch is closed? R1 and R2 are both 5 ohms, V is 10 V and C is 10 mF. If you can do this without doing a calculation, please explain your reasoning. See Fig. C.4.

5. (3 points) Explain what will happen in the above circuit if after we have fully charged the capacitor, we open the switch? If instead of plain resistors we had bulbs in this circuit, which would light up, R1, R2, both or neither? Would it/they stay light?

C.3 Third Lab Quiz

1. (3 points) Explain Kirchhoff’s junction rule: how do you use it and what general principle does it come from?

2. (3 points) Explain Kirchhoff’s loop rule: how do you use it and what general principle does it come from?

3. (3 points) Write each junction equation for the following circuit (see Fig. C.5).

4. (3 points) Write 3 loop equations for the circuit.
5. (3 points) Take $V_1 = 10 \text{ V}$, $V_2 = 5 \text{ V}$, $R_1 = 200 \text{ ohms}$, $R_2 = 100 \text{ ohms}$, $R_3 = 100 \text{ ohms}$, and $R_4 = 200 \text{ ohms}$. Solve for each current in the circuit.

**C.4 Course Exam Questions**

1. Barrett’s Final Exam: First Multiple Choice Question

   A $+2 \ \mu\text{C}$ and a $-1 \ \mu\text{C}$ charge are placed on the corners of a right triangle as shown. The direction of the force on a $-3 \ \mu\text{C}$ negative charge placed at the lower left-hand corner is best represented by which arrows?

2. Reay’s Final Exam: Capacitor and Battery Circuit

   Find the charge on the $3 \ \mu\text{F}$ capacitor. (The circuit contains two capacitors in parallel with a third; all in parallel with a battery.)

3. Reay’s Final Exam: Charged Plates Question

   Three square aluminum plates, each $1 \ \text{m} \times 1 \ \text{m} \times 1 \ \text{cm}$ thick are spaced $10 \ \text{cm}$ apart, as shown below. The plates are connected to an electrical circuit, and the right-most plate is connected to a metal rod stuck in the Earth. The
plates will have the same potential as the point in the circuit to which they are connected, and the right-most plate will be at 0 volts.

a) Draw equipotential surfaces at 20 volt intervals on the above figure.

b) What is the distance in meters between the equipotential surfaces?

c) Calculate the magnitude and direction of the electric fields between each of the two plates.

d) Calculate the magnitude and direction of the force on an electron located between the central plate and the rightmost plate.

4. Reay’s Final Exam: Current Carrying Wires

Four infinitely long wires each have currents $i = 100 \text{ A}$ going into or coming out of the page as shown. Also, $d = 0.01 \text{ m}$.

a) What is the magnetic field at the origin?

b) What would be the magnetic force on a wire segment 1 meter in length placed at the origin with a 50 A current coming out of the page?

c) Now, remove the wire segment of part b. What would be the direction and magnitude of the magnetic force on a moving particle with 1 $\mu\text{C}$ of charge and velocity 5 m/s moving in the positive $x$-direction at the origin?
The essay assignments given in the Winter quarter Physics by Inquiry course are:

1. Other than simply using “trial and error”, explain how to balance two objects of unequal mass in a way which would be clear for an intelligent but uninformed person. Also explain how this applies to real world examples such as using scales or children on see-saws. Refer explicitly to in class observations and/ or activities.

2. Your friend is given a long length of string with equally spaced markings on it. They are then given the assignment of determining the circumference of various types of tires. Explain to your friend (who is intelligent, but not in your physics class) what needs to be done in order to use the string to determine if any two tires definitely have the same circumference, may have the same circumference, or definitely do not have the same circumference. Briefly discuss what is similar and/or different between making these measurements and comparing the lengths of straight objects with rulers. Where possible, relate the ideas to activities you did in class (but don’t just refer to the section numbers!)
3. Sixty milliliters of a homogeneous mystery substance has a mass of 90 grams. Write an essay explaining in detail how you can use this information to predict the mass of 90 milliliters of the same substance. Include in your essay a description of how this process is related to the idea of the ratio of mass to volume of a material. Include a meaningful introductory and concluding sentences, as well as good transitions between sentences.

4. Write an essay to explain how to determine if an object will sink or float in some given liquid. Be sure to include how to make the prediction and how to test it as well as how it relates to the concepts of mass, volume and density. Include your own in-class observations to support your reasoning.

5. Considering the discussion on p. 73 in Exercise 13.8 and 13.9, write a short essay to explain how a mass-volume graph can be used to determine how a piece of wood and a piece of clay can be combined so that the inhomogeneous object will sink or float in a given liquid of arbitrary density. Indicate the role played in your graph by the density of the wood, the density of the clay, and the average density of the inhomogeneous object.

The essay assignments given in the Spring quarter Physics by Inquiry course are:

1. Explain clearly to an intelligent but uninformed friend what is meant by open, closed, and short circuits. Explain these in the context of turning on a light with a switch: when is the circuit open, closed, and how might it short? Refer explicitly to in-class observations and/or activities if necessary. Pay close attention to writing a meaningful introductory and concluding sentence, as well as the transitions between sentences.
2. When given networks of bulbs which are connected to a battery, explain clearly to an intelligent but uninformed friend how to determine if the networks are independent or dependent. Also give one example for each type of network. Refer explicitly to in-class observations and/or activities if necessary. Pay close attention to writing a meaningful introductory and concluding sentence, as well as the transitions between sentences.

3. Explain clearly to an intelligent but uninformed friend why a circuit diagram of the wires in your house may be more useful than a diagram showing the layout of the wires. Include in your essay a description of how the process of redrawing the circuit can clarify the relationship between or among networks and might be related to the distinction between series and parallel circuits. Refer explicitly to in-class observations and/or activities if necessary. Pay close attention to writing a meaningful introductory and concluding sentence, as well as the transitions between sentences.

4. Explain clearly to an intelligent but uninformed friend how to connect an ammeter and a voltmeter in a circuit in order for them to function correctly. Be sure to discuss the differences in their operation, and to discuss the resistances of each device and how that relates to the way they are connected. Refer explicitly to in-class observations and/or activities if necessary. Pay close attention to writing a meaningful introductory and concluding sentence, as well as the transitions between sentences.

5. Explain clearly to an intelligent but uninformed friend how to find out whether a circuit can be decomposed into series or parallel components. Discuss how
this is useful for understanding complicated circuit diagrams like those for your home wiring. Refer explicitly to in-class observations and/or activities if necessary. Pay close attention to writing a meaningful introductory and concluding sentence, as well as the transitions between sentences.
BIBLIOGRAPHY


286


