Managing the Improvement of Student Written Presentation Using a Complete Measurement Technique

Gay B. Stewart and John C. Stewart

November 14, 2000

Abstract

This paper presents a complete, unbiased, and reproducible measurement technique for written technical communication in a physics class. A measurement based on this technique shows serious presentation problems in a traditional class. Policy changes that have produced measurable improvements in presentation are presented.

1 Introduction

In the last decade, improving the poor performance of physics students on conceptual understanding examinations [1] has become a major focus of many physics departments. This poor performance was a surprise to the physics community because traditional instruction relies on the assignment of non-trivial calculations to teach physics. These calculations would seem to require a mastery of conceptual skills for their successful completion. We present a measurement of the written communication of a traditional university mechanics course which shows very low levels of written English and technical drawing. Since the conceptual components of a technical calculation are often exposed in the words and drawings in the solution, this may indicate that part of the failure of traditional instruction methods lies in the level of technical writing accepted. In this paper, a complete unbiased measurement technique for written technical communication is introduced, and applied in detail to student solutions of a calculus-based physics problem. A pattern of progressive improvement of student technical writing is quantitatively demonstrated, and the course policy which led to that improvement is discussed.

2 Measurement Technique

Our program to improve student written presentation in introductory physics began with the measurement presented in Table 1, which compares the counts of the number of basic language elements found in graded student writing in an introductory calculus-based mechanics course (Class I) and its second semester electricity and magnetism course (Class II) at the University of Arkansas. The results in Table 1 estimate the number of presentation elements in total graded communication (excluding lab reports) for a semester of the class. The *symbol* is a mathematical variable and the *number* is a number and any associated units, so the units are not counted as symbols. If the conceptual content of student work is contained in their output of words and diagrams, it would be no surprise if Class I did a poor job in the conceptual education of students. This simple measurement also shows that Class I students are working largely with numbers, where Class II students are working mostly with symbols, indicating better mathematics presentation in Class II.

There is no grand principle of education at work producing the difference in the two classes. Class I makes homework optional and assigns multiple choice tests. Class II collects homework and gives substantive exams with a strong conceptual component. The results for Class I are a natural outgrowth of the trend toward mechanical evaluation and lower grading loads. The results do not immediately imply that Class I is inferior, but strongly suggest that detailed research should be performed to verify that quality has been maintained with such low levels of written output.

It should be noted that the measurement in Table I is reproducible, relatively free of researcher bias, and fully comparable between institutions. It could be a part of a basic collection of measurements used to characterize all technical classes.

2.1 A More Detailed Measurement Technique

The simple measurement above was very powerful. It showed unambiguous problems in the student conceptual presentation in those 13 words a semester. The technique can be refined to give a more detailed representation of technical writing by taking advantage of the rich set of mathematical and graphical symbols used in the solution of a physics problem. Physics uses many special mathematical symbols, such as vector markings ($\vec{F}$, $\hat{F}$), vector operations ($\vec{A}\cdot\vec{B}$, $\vec{A}\times\vec{B}$, $|\vec{A}|$), calculus operations ($\int$, $\frac{d}{dx}$), algebraic operations ($\sqrt{a}$, $+$, $-$), and others. Physics solutions also contain line drawings composed of simple two-dimensional shapes and special drawing elements such as coordinate axes, vector arrows, mathematical symbols, motion indicators, etc. A count of the special mathematical and graphical symbols in a problem is a quantitative measurement of the richness of presentation.

To extend the object counting technique to include the special objects, a complete list of the objects must be developed that divide mathematical and graphical parts of a solution into countable objects. This list of objects should be not be biased by the researcher's preconception of the solutions content or the researcher's beliefs about what is good or bad in a solution. Since counting all presentation elements is time consuming, we track student presentation by examining a single representative problem which is assigned unchanged each semester although with a different problem number so it is not readily identfiable. Observation of student writing has shown that the manner of presentation is fairly consistent for a given set of course policies and therefore measurement of one problem is a good estimate of the general quality of solutions. To develop a complete un-biased list of problem elements, a good instructor solution to a problem is produced, and a few student solutions are chosen at random. Using a technical writing program, a reproduction of the sample problems is produced. The reproduction is compared to the originals and missing elements are identified. This has proven a very effective mechanism for identifying missing presentation elements.

Our observation uses the LaTex technical typesetting program [2, 3]. It has superior mathematics reproduction and an appropriate set of simple line drawing functions for the reproduction of student diagrams. It is familiar to many scientists, a valuable skill for future scientists to learn, and freely available over the internet. The LaTex document which reproduces the sample problems is converted into a basis set of presentation objects by examining the code for special characters and LaTex commands. These characters and commands represent the presentation objects in the solution. They range from simple special characters $+$, $-$, to characters representing operations $|$, to simple math functions $\sqrt{a}$, to complex drawing functions. Different mathematics reproduction programs will produce different sets of primatives. LaTex has the very useful feature for this research that the description of the document is easily readable. Note, the development of the complete object list only has to be done once, it is not repeated each semester.

With the set of basic objects, an observer examines a population of student solutions or other writing. The observer counts the number of these presentation elements which would be required to reproduce the solution. This does not require actually reproducing each solution, but is merely a count of the number of objects which would be required to do so. This is fairly straight forward if the collection of basic objects is relatively complete. New objects are added as they are discovered in student solutions. The measurement accuracy is improved if a full mechanical reconstruction is done, but this increases measurement time 100-fold.

Any mechanism for developing a list of basic presentation objects to count will provide a measure of the written communication. The process above minimizes the researcher bias of the measurement, since the selection of the basic objects to count is controled by the requirement that the sample problems be reproduced. If LaTex or the emerging internet math standard, MathML, could be adopted as a measurement standard, this method would allow very detailed measurements of the written elements of science engineering classes, which were reproducible, un-biased, and comparable between institutions. The method is intrinsically unbiased by our opinion of what is in the problem and our belief of what is good about the problem. The measurement can actually be carried out by a student with little or no expertise in education. Like any good scientific measurement, the measurement shows us what is in the problem, not what we expect to be there, or what we believe is good or educationally relevant.

2.2 Measurement Results

The refined measurement technique was applied to a representative problem. The problem was assigned unchanged in the Spring 1998 and Spring 1999 semesters. The selected problem was neither one of the hardest or easiest problems in the class. For the presented measurement, the following problem was used:

> A wire of length 16cm is suspended by flexible leads above a long straight wire. Equal but opposite currents are established in the wires such that the 16cm wire floats 1.5mm above the long wire with no tension in the suspension leads. If the mass of the 16cm wire is 14g, what is the current?

The results of applying the refined measurement technique to the instructor's solution, 30 student solutions from the Spring 1998 semester, and 30 solutions from the Spring 1999 semester are shown in Tables 2, 3, and 4. Table 2 reports the result of counting the basic objects of the solution as before. The solution was divided into four different regions; (1) A re-statement of the problem (Problem Statement), (2) Statements defining or giving the value of some of the symbolic variables in the problem (Variable Definitions), (3) The diagram, and (4) Everything else which encompasses the actual reasoning of the solution, called the "Core Solution". For entries which were sometimes missing from the problem, the number in parenthesis represents the number out of the 30-problem sample where the feature was present. The averages are taken over the number of solutions possessing the feature.

For the problem selected, 92 special characters and commands were required to reproduce the sample problems. Table 3 presents the totals for the objects used in our class engineering for objects found in the diagrams and Table 4 presents the totals for the "Core Solution". For different engineering purposes, different sets of symbols will be interesting. Subscripts are symbol subscripts. Vector markings are either the arrow over a vector symbol ($\vec{F}$) or the hat over a unit vector symbol ($\hat{F}$). "Drawing Objects" are the line, circle, vector, rectangle, and curves in the diagram and "Total Lines and Curves" are the number of lines or curves it takes to make a drawing object. For example, a rectangle takes four lines and so would count 4 toward "Total Lines or Curves", but one toward "Drawing Objects". The number of drawing objects and total lines and curves is a good indicator of the complexity of the drawing.

3 Applying Measurement Results to Improve Student Writing

Examination of Spring 98 student solutions, with the quantitative measurement and the instructor solution in mind, showed the following disturbing pattern. Not only were the conceptual steps of the solution not stated, the appropriate diagramatic and mathematical structure to perform the reasoning was also not there. So it seems very likely the conceptual steps were not being done. The students seemed to be fixing up signs, directions, and writing balance equations because of the constraints of symbolically solving the problem, not because of their knowledge of physics. Since the conceptual steps are a crucial part of solving a real world problem, it seems unlikely these students would fair well on a real world problem. In fact, research shows that students have trouble transferring problem-solving techniques to related but different problems, a difficulty reduced by greater conceptual understanding [4, 5, 6, 7]. Further, we were struck by how much conceptual structure was present in our written calculation, that was missing from the student's solution, suggesting quantitative calculation, if properly managed, could be use to supplement an explicit conceptual component of a science class.

This section presents the steps taken to improve the solution writing habits of the students in the Class II. In Spring 98, the student solution is relatively tiny, about three inches long. The diagram is very small and present in only 70% of the problems. Only half the problems contain variable definitions or a problem statement. These objects contain the majority of the written English in the problems. The students are writing, but very little of the writing is going where it is most needed, in the core solution. The students are using more symbols than numbers reversing the Class I results, but are still using far fewer symbols than in the instructor solution, indicating that mathematics steps are also being skipped. Examination of the Spring 98 data in Tables 3 and 4 show a low level of the use of vectors in an intrinsically vector problem, both in the core solution and in the diagram. The missing cross product symbol from the Spring 98 diagrams indicate the students are skipping the crucial conceptual solution step of representing the direction of the magnetic field of the fixed wire at the floating wire.

The low solution size was addressed by requiring the student to start each problem on the top of a new page. Examination of the Spring 98 solutions showed the student would make a problem fit whatever space was left on a page, and often determine the size allowed the problem before they began the problem. The diagram size was improved by the simple expedient of requiring a 7cm square diagram. Both problem size and diagram size were substantially improved in the Spring 99 problems by these requirements. The students used the additional area to improve expression, which can be seen in the increase in drawing objects and vector markings in Table 3 and in the increase in all important special symbols in Table 4. In Spring 99, the existence of the problem statement and variable definitions section, as well as a part of the core solution called the "Strategy" was graded. The strategy is a one or two sentence statement of the student's general plan for working the problem. This method of grading the existence of identifiable sub-objects was very effective as can be seen from the 80% – 90% existence rate for problem statements, variable definitions, and diagrams in the Spring 99 data in Table 2.

The problem statement, variable definition section, diagram, and strategy give the student a solution with superior structure. By requiring these sections and grading them separately, a framework is produced within which the student can focus on writing good physics in the core solution. There is no way to break the writing of the actual reasoning of the solution into cookbook pieces. A grading key is produced where points are given to both reasoning and calculation steps. Students are required to reference the names of the physical laws and principles used. Requiring the student to produce some of the problem based on a recipe leaves only one difficult part to grade, thus focusing the grading effort where it will do the most good. The required structure of the student solution attempts to balance the expert suggestions [8] on student solution style with the time cost of producing well-presented solutions and grading them.

Examination of Table 2, Table 3 , and Table 4 shows the Spring 99 students have made substantial progress toward the goal set by the instructor solution in all areas over the Spring 98 students. Particularly impressive gains have been made in problem and diagram size, words in the total problem, and the number of students using words in the core solution. The increase of subscript and vector markings in the figure and solution shows improved vector presentation as one would hope for a problem intrinsically involving a force balance.

The key to making any of this work, especially the justification of the steps in the solution, has been careful oversight and management of TAs. TAs habitually do not grade conceptual steps and students are very quick to identify a grader who will accept solutions which are just math. TAs, after all, often have the same poor technical presentation skills as their students.

The Class I entry in Table 1, Spring 98 data Table 2, and Spring 99 data in Table 2 show three stages in an improvement process for student technical writing. The Class I measurement shows what one can expect from the student's writing in a course where homework in not graded and all exams are multiple choice, the current trend in physics instruction. The Spring 98 result shows what can be accomplished with passive encouragement. The students were told the benefits of writing good solutions, and given appropriate supporting materials. The existence of a figure was graded, and the students were told that justification of problem steps would be graded. The Spring 99 measurement shows what can be accomplished when the writing process is broken down and explicitly taught and graded. The measurement presented allows an educator implementing this process to quantitative see an un-biased measure of the level of presentation in their student writing, and compare it with the level of presentation in their own goal solution for the problem.

4 Future

Counting the frequency of words used, sentences length, and paragraph length has long been a useful mechanism for evaluating the level of written material. Our use of the distribution of words, mathematical, and graphical primitives in scientific writing does not do justice to the amount of insight which can be gained from analysis of this information. The complexity of expression, the average complexity of equations, the density of equations, and their distribution throughout a material would be a valuable information for characterizing the material, and a valuable tool to the author. This information is too time expensive to be gathered globally as long as it rests on human observation. Once standards for storing text containing mathematics are created and broadly accepted, tools which give authors access to this type of information can become as common as the tools which extract the level of the English used.

Acknowledgements

We thank Dr. T.C. Burt for insisting that TEX was the correct observation language for this experiment and T.C. Burt, M.B. Martinez, and R. Coffee for their patience in observing physics problems. This work was supported in part by a grant from the National Science Foundation.

References

[1] David Hestenes, Malcolm Wells, and Gregg Swackhammer. Force concept inventory. *The Physics Teacher*, 30:147, 1992.

[2] Leslie Lamport. *Latex A Document Preparation System, User's Guide and Reference Manual*. Addison-Wesley, Reading, Mass., 1994.

[3] Donald E. Knuth. *Tex: The Program*. Addison-Wesley, Reading, Mass., 1993.

[4] William C. Robertson. Detection of cognitve structure with protocol data: Predicting performance on physics transfer problems. *Cognitive Science*, 14:253, 1990.

[5] D.E. Egan and J.G. Greeno. Acquiring cognitive structure by discovery and rule learning. *Journal of Educational Psychology*, 64:85, 1973.

[6] R.M. Gagne and L.T. Brown. Some factors in the programming of conceptual learning. *Journal of Experimental Psychology*, 62:313, 1961.

[7] D.N. Perkins and G. Salomon. Are cognitive skills context bound? *Educational Researcher*, 18:16, 1989.

[8] Donald Scarl. *How to Solve Problems, Fourth Edition*. Dosoris Press, Glen Cove, New York, 1994.

List of Tables

Table 1: **Estimated Totals for Written Communication for One Semester of a Science Class** Class I is an introductory calculus-based mechanics course. Class II is the introductory electricity and magnetism which is taken after Class I.

Object	Class I Object Count	Class II Object Count
Words	13	2378
Numbers	438	1104
Symbols	245	2429
Figures	8	56

Table 2: **Average General Properties of a Student Problem** When a number appears in brackets next to one of the entries, it is the number of non-zero counts out of 30 that contributed to the average. The average is taken over the non-zero items. For example, if only 7 solutions had words within them, then the average reported would be over those seven, not over all thirty problems in the sample.

Object	Instructor Solution	Student Spring 98 Solution	Student Spring 99 Solution
Problem Length(cm)	44	7.8	21
Diagram Area(cm^2)	58	16(21)	47(29)
Problem Statements	1	exists(18)	exists(26)
Variable Definitions	1	exists(15)	exists(25)
Words	172	20(17)	64
Numbers	20	15	19
Symbols	58	26	37

Table 3: **Measurement of the Figure Presentation Objects in a Single Student Written Problem** The student solution entries are averaged over 30 problems.

Object	Instructor Diagram Count	Spring 98 Diagram Average	Spring 99 Diagram Average
Numbers	0	2.1	1.6
Symbols	8	2.9	4.8
Subscripts	2	0.8	2.7
Vector Markings	3	0.4	1.2
Drawing Objects	15	9.8	14.6
Total Lines and Curves	25	17.1	25.9
Cross Product	1	0	2.7

Table 4: **Measurement of Core Solution Presentation Objects in a Single Student Written Problem** The student solution entries are averaged over 30 problems.

Markup Object	Instructor Core Solution	Spring 98 Core Solution	Spring 99 Core Solution
Words	105	13(7)	20(19)
Numbers	12	11	11
Symbols	52	24	32
Subscripts	9	5.0	10.3
Vector Markings	14	0.6	4.6
Modulus Lines	4	0.4	0.6
Cross Product	1	0.2	0.7