Teaching Critical Thinking and Writing Skills with CPR

Robert Hodson[1], Todd Nickle[3], Linda Dion[4], and Dee Baer[2]

[1,4]Department of Biological Sciences
University of Delaware
[2]University Writing Center
University of Delaware
[3]Mount Royal College
Calgary, AB

[1] hodson@UDel.Edu
[2]TNickle@mtroyal.ca

Abstract: This workshop explores methods of teaching critical thinking and writing skills applied to laboratory reports. It is a challenge to teach these skills to large numbers of students and to teach report evaluation skills to new graduate student laboratory instructors. We conducted a carefully designed project in the Fall Semester 2006 in which Calibrated Peer Review (CPR) was tried as a way to train students by using peer and self evaluation and a rubric testing for the target skills. Honors and conventional students in an introductory biology laboratory experience were assigned in small groups to CPR or standard (untrained) peer review. After the semester, a trained team of assessors evaluated rough and final drafts of early (first) and late (last) of four reports spread over the semester. Responses on pre and post questionnaires were also examined, and participating teaching assistants were queried. We present outcomes, provide a computer-based "walk though" of a CPR cycle of activities (prepared with Wink; a poster – included in this 2007 ABLE collection – presents this software), and open the floor to discussion. Although participants might attend to hear about the quantitative data collected, there are three other significant aspects: (1) the materials prepared for the project which included a very useful rubric, a handout on preparing tables and figures (graphs and images), and a handout highlighting key scientific writing rules, (2) a revised and successful teaching plan for the sequel Spring 2007 course, and (3) the method used to include tables and figures in the CPR form designed only for html-tagged text.

Introduction

This ABLE mini workshop is about using and evaluating CPR (Calibrated Peer Review) as an aid in teaching students, and laboratory teaching assistants (TA's), critical thinking and writing skills for

laboratory reports (also called technical papers). Observations were made during a project funded by academic and service units at the University of Delaware. We begin with recognition of communication and evaluation problems encountered with students who are writing reports and TA's who are grading them. The funded project and implementation plan in the first semester introductory biology course are described, followed by a summary of initial results. Our experience led to a revised plan for the sequel semester course, and we conclude with a summary of its second-round results.

Laboratory reports were modeled on primary literature with certain components omitted and others abbreviated. The Introduction section had a brief statement on background of the topic, a hypothesis or experimental question as appropriate, and justification of the hypothesis proposed or question asked. A Methods section was omitted, but a brief description could be given in the caption to a figure or title of a table if it was necessary for data comprehension. There was a traditional Results section with text and data in the form of figures or tables. Last there was a Discussion section in which there were statements about support or rejection of the hypothesis with justification, limitations to the interpretations that could be drawn including any known errors in data collection, relevance of the findings to known biological facts or concepts (usually obtained from a textbook), and a possible future related investigation that logically follows current work.

Students generally fall short of producing high quality reports because they do not know how to think critically and communicate clearly and properly in the biology discipline. Teaching assistants, usually first year graduate students, are generally under-developed in the same thinking and writing skills which carries over to difficulty in evaluating student work. Faculty can't cover for TA shortcomings because of heavy teaching and research responsibilities.

CPR was developed by chemists at the University of California – Los Angeles (UCLA) with funding from the National Science Foundation and Howard Hughes Medical Institute. It is a "Web-based program that enables frequent writing assignments even in large classes with limited instructional resources" (http://cpr.molsci.ucla.edu/). One of us (Todd Nickle) has used the program extensively. Another author (Bob Hodson) wondered if CPR would be helpful in training students to better express themselves in a confident, competent manner, and how TA's participating in a CPR cycle might benefit with respect to their assessment abilities.

A CPR cycle as originally developed at UCLA progresses through four stages and in our project it was allotted one week.

- In the first stage text is created (in our adaptation this is the laboratory report). Students are often not comfortable with making a multi-paragraph piece of work. CPR gives them this opportunity. It's truly open-ended, as just about any topic can be fodder for text creation. Students have a limited time to get the text in; with our project it was 4 days after completion of the laboratory investigation.
- When the text-entry phase is complete, students enter the calibration and review phase, stage 2. The switch from text-entry to calibration and review occurs at the same time for all students (aside from those individuals who need this adjusted because of personal circumstances). This means that all students should have an idea about what the assignment is about. When first logging on in this new phase, students must complete three calibrations regarding essays provided by the instructor. One is good, one is bad, and one is a moderate piece of work. The bad and moderate pieces have errors intentionally introduced to resemble

errors students frequently make on essay-style assignments. There are guided questions that focus each student's attention on the text. If the students pick up most of the errors, they are considered to be "good markers" and given a high Competency Index. Missing substantial numbers of errors will force the student to re-calibrate on that particular essay and the student Competency Index is decreased. This index tells the computer how reliable the student is as a marker.

- After the Competency Index has been calculated, Stage 3 is begun when each student reviews (anonymously) three other student essays. Actual student essays are graded by three student reviewers, and the mark is a weighted average from the three reviewers: highly competent markers are given a higher influence on the average for that grade. This instills a sense of fairness in the student grades.
- Reviewers also make comments on the work. As mentioned, a weighted average for the essay quality is determined at this stage.
- Finally, in Stage 4 students get their OWN work back to review using the same guiding questions they had for calibrations and reviews. Their job is to consider how well they did the assignment in the light of calibrations and the sample writing of their peers. The grade they get here is not the mark they give their own essay, but rather how closely it resembles the weighted average from the reviewers.

Thus, CPR accomplishes

- modeling good and bad characteristics of essays,
- giving students a chance to practice writing skills,
- giving students a chance to practice reviewing skills as a critic, and
- giving feedback on essay quality which encourages self-evaluation and reflection.

Because in our project we were trying to train for specific skills we felt it important to give students and TA's guidelines and grading criteria before they wrote anything. Thus we developed three handouts for them. Interestingly it turned out that the three documents were some of the most valuable outcomes of the project. (1) The first document was instructions for presenting data in the form of figures, i.e. charts and graphics, and tables according to generally accepted conventions one would find in primary research journals (**Appendix I**). Previously we had given some guidelines for this, but usually verbally and not as thoroughly. (2) Second, a set of writing conventions was prepared (**Appendix II**). We hoped, if nothing else, students would learn to describe results in past tense and that "data" is plural!! (3) Third, we developed a scoring rubric (**Appendix III, version 1**). Students were advised to use this in creating reports and TA's used it for grading them. The first two documents mentioned, presenting data and writing conventions, aided TA's in grading reports because they had specific things to find and evaluate.

The CPR program as developed at UCLA had to be customized to be useful for laboratory reports at the University of Delaware.

1. First, we set up the CPR "course" with our own rubric and instructor-written reports for the calibration stage (Stage 2). The two sets of calibration reports were written by co-investigators. Bob Hodson wrote the high quality version and Linda Dion modified them to produce mid and low quality versions. We have since learned it would have been better to use actual student work because it is difficult to emulate student writing. Unfortunately student samples were not available at the time.

2. Second, we had to develop a way for students to include data (charts, tables, images) in their reports. The CPR software was developed to accommodate only text entry into an html form. Microsoft Word or PDF documents are not accepted. All tags such as paragraph separators (<p>) and line breaks (
) have to be entered explicitly by typing them in. However charts (graphs) and pictures are stored as separate image files, and tables are extremely difficult to construct using tags. To overcome these limitations we modified the report creation and submission protocols. A personal server at UD was enlisted and a password-protected dropbox created for image files. The dropbox allowed students to submit image files using ordinary ftp (file transfer protocol) while still maintaining pretty good server protection from unauthorized users, and it made it very difficult for students to download the work of others because the dropbox did not list file content.

   Here briefly is the protocol that CPR-using students followed to create a laboratory report such that peers could fetch a document complete with tables and figures for reviewing in Stage 3.

   - The report first draft was assembled using Microsoft Word. Tables and charts were created with Microsoft Excel and transferred into the Word document.
   - The Word document was saved as html. This created a folder containing image files, and it also created html-tagged text with embedded tables.
   - Students sent the entire folder with images to the satellite server.
   - The html file "page source" was opened and searched for all image tags. These tags were modified to point to specific files in the satellite server.
   - The entire modified page source was copied and pasted into the CPR form's text entry box.

   Did students find this difficult? Some didn't but too many did. We learned after the first report cycle to create an Adobe Flash tutorial on the procedure using the free Wink tool (http://www.debugmode.com/wink/) and this greatly helped students with less computer experience.

3. A third customization had to do with assigning points and allowing students to benefit from peer comments. The original CPR program is configured to award points for calibration, peer review, and self review. We eliminated this because we wanted TA's to do all of the actual grading. Instead we banked on students giving helpful written comments in the peer review stage, assuming there would be enough incentive from pride or realization that to expect helpful comments from peers one should give them. This was a mistake. One has to give points to encourage serious effort.

4. A fourth customization had to do with grading. A fifth CPR stage was added in which authors revised their report based on peer comments and submitted the final draft to their TA for a grade.

The entire cycle, from the time an investigation was completed to the time a report was returned with grade, was given two weeks.

The project plan was designed as a pilot study involving only a subset of students and TA's. We selected 5 of 30+ laboratory sections and thereby involved about 80 out of 700 students and 5 out of about 20 TA's. Two sections had honors students who were mostly biology majors. One of these was designated a CPR test group and the other served as a "control" in which peer review was performed within teams of three students without formal training. Three sections had conventional students who were a mix of majors including biology, nursing, exercise science, biochemistry, and others. One section was designated a CPR test group and two were controls. These three conventional sections were taught by graduate TA's who also taught a second section not involved in the study and lacking peer reviewing. The hope was they could provide valuable feedback on student performance by comparing their two differently treated sections.

Four variables were incorporated into the plan. These were the type of peer review (with and without CPR training), the type of student (conventional versus honors), the laboratory report draft (first versus final in which peer comments if any were incorporated), and report number (number 1 and number 4 which was the last).

Some variables could not be controlled. One was student behavior. We heard one CPR group student comment that peer review was not taken seriously because TA's assigned grades. This didn't seem logical to us but if widespread could have affected the outcome. Examining some CPR group peer comments even after calibration training indicated a wide range of correctness and usefulness, and it was therefore understandable that some students did not trust peers and incorporate their comments.

Students participating in the project (CPR and non-CPR) wrote four peer-reviewed reports while all other students wrote twice as many without peer review to balance the work load. Students in CPR groups participated in two cycles that had a calibration stage and two cycles bypassing calibration (Stage 2) and going directly to the peer review stage (Stage 3). TA's participated in two CPR cycles, both with calibration stages.

At semester's end a new rubric (**Appendix III, version 2**) was created to better assess the project goal of critical thinking improvement. Rubric questions were grouped into "critical thinking" and "writing" categories, and a fourth level, above expectation, was added to each question.

An assessment team of four experienced undergrads was assembled and trained (Dee Baer) using a "norming" approach. Each assessor read and scored a sample report and then scores were discussed. This was repeated once. The outcome was pretty good agreement amongst the assessors, and they also better understood rubric questions and suggested ways for rubric improvement.

A total of 134 reports which included all four types of variables was scored. Each report was scored by two assessors working independently. If they disagreed by more than 10% a third reading by one of the investigators (Linda Dion) was added. Means were calculated and pairs were compared using a 2-tailed t-test. Sample sizes ranged from 4 to 16, and there were 32 t-tests performed.

Score means on a percent scale for rough ("first") and final drafts of the first (#1) and last (#4) reports are reported in **Fig. 1** Data are shown for conventional students using CPR, conventional students in a control group (not using CPR), honors students using CPR, and honors students in a control group. Also shown are two means for grades assigned by TA's.

Overall the scores from assessors were low compared to TA grades. Part of this is likely due to different rubrics being used by TA's (and students) and assessors. The latter had a more demanding scale of 0-4 rather than 0-3. Almost no scores of four were awarded, which in effect means that the highest possible score was about 75%. If 25 points are added to each mean, then the range of assessor scores would be similar to TA grades.

The most important comparison is between students using CPR and students in control groups. The only significant difference was that conventional students in the CPR group scored higher than the control group on the first draft of the last report. In the other comparisons there is a suggestion that CPR was beneficial in some cases and not in others and thus there is no consistent pattern. Unfortunately we could not include a no-peer review control group to reveal if peer reviewing as such was beneficial. Some students commented that CPR was of no value to them, but a few thought it was very valuable to be able to get ideas for writing and critical thinking from other

students in this way. We received many comments that incorporating figures and tables into text for CPR was too difficult and discouraging.

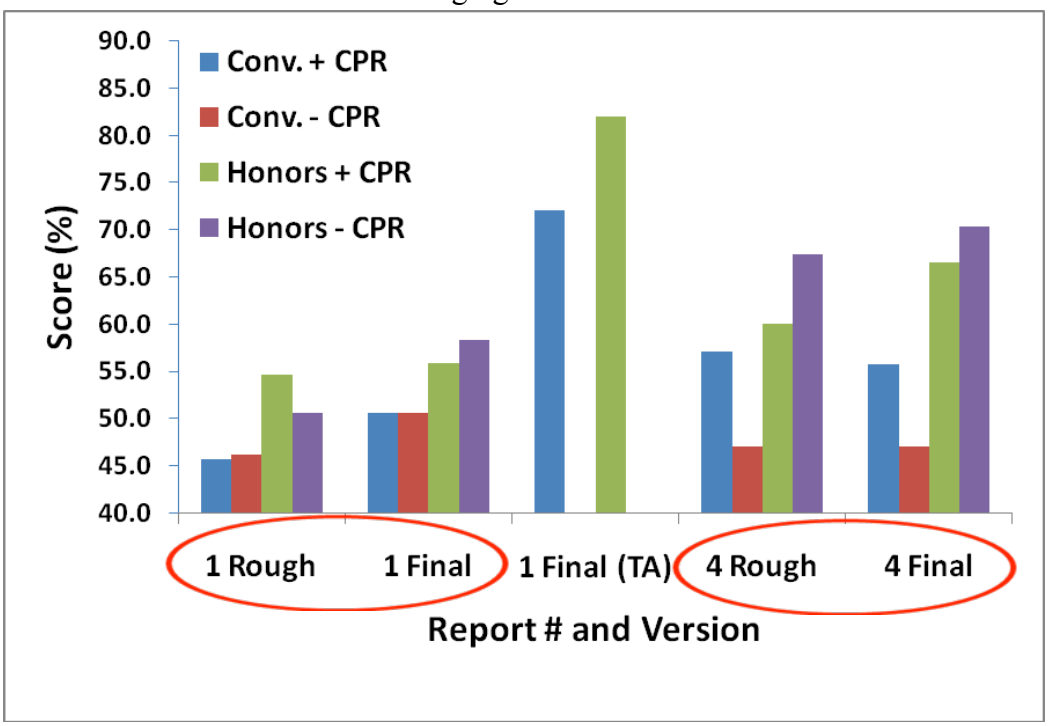


Figure 1. CPR Project Assessment Results. Peer-reviewed laboratory reports were from conventional (Conv.) and honors program students trained (+CPR) and not trained (-CPR) for peer reviewing. The final drafts of the first and fourth (last) laboratory reports would have incorporated any peer reviews from the rough draft. Except for the two sets of scores assigned by TA's, all other report scores were assigned by the independent assessors.

One can take these student results in two ways. (1) Training with CPR did not result in significantly superior student work and is therefore not worth the effort it takes to set it up. (2) Training with CPR did not result in significantly inferior student work and it could be quite useful in courses with large numbers of students and few or no TA's.

As for the five TA's, two were unable to participate in CPR cycles consistently and their experiences had little impact. The other three TA's commented that the custom aids and rubric were beneficial but that students did not seem to appreciate or benefit from CPR. It is unclear if CPR is an effective way to teach laboratory report grading.

Modification of plan for sequel (spring) semester

As the sequel course in the Spring Semester neared, we thought about the norming approach used to train the assessors and how successful it appeared to be in getting across expected writing and

thinking outcomes. Perhaps we could better train TA's using it, and in turn they could train their students the same way.

A new plan was developed for a TA team of about 15 mostly second and greater year graduate students with a few experienced undergraduates.

- TA's were trained with two norming sessions. Sample student reports were used with student permission. [Example reports are shown in **Appendix IV**; these reports are from summer school students but they are similar in quality to those used in the spring semester norming.] Dee Baer led the first one and biology faculty the second. In the first session conducted before semester start the scores on the first sample report (mediocre quality) ranged from about 20 to low 30's out of 42! After discussion the range narrowed significantly, low to mid 20's. In the second session two weeks later with a different student report (same topic though) the range remained acceptably narrow. We concluded that norming is a good way to train TA's.
- TA's in turn used the first sample student report for one norming session with their students. We also collected report means for each of the four reports from each TA and shared these data with all. TA's and faculty commented how helpful it was to be able to compare their grades to those of others.
- Again four reports with the modified journal article format were assigned on an every other week schedule allowing TA feedback to benefit the next effort. Peer reviewing was omitted. For the in-between weeks collected data were analyzed and reported in a Results section as figures and tables but no text. This at least forced students to grapple with data analysis and communication.
- A third rubric (**Appendix III, version 3**) was developed in which the questions were essentially the same as in the previous version (#2) but some had the range of possible scores shortened from 0-4 to 0-3. For the skills demonstrated by these questions we felt a score of 4 – above expectation – was not attainable. If a student achieved a score of 3 on all questions they would have a grade of 85% which is a B. That seemed about right.

Means in units of percent for the four reports from both conventional and honors students are shown in **Fig. 2**. The four means from conventional students ranged from 70% to 85% with an overall mean of 79, and the range with honors students was 84 to 90% with an overall mean of 88. These values are very satisfactory. They show good progress over the semester and are not inflated as they have been in previous years. **We conclude that the norming method of training for both students and TA's is better for our program than peer review with CPR.**

In the fall semester of 2007 we plan to again use the norming approach. We will have mostly first year graduate student TA's, and we will separately track means of critical thinking skills and writing skills parts of report scores in addition to overall scores.

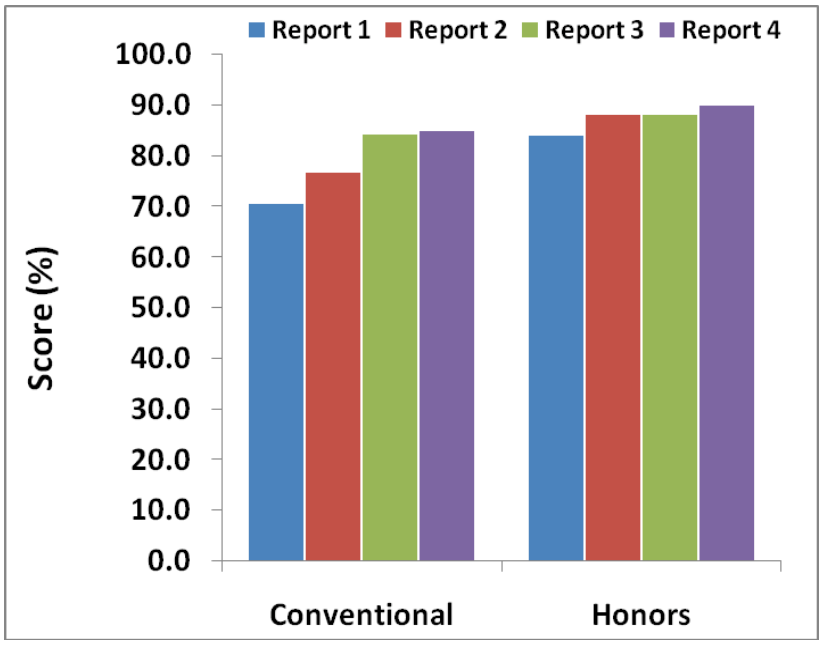


Figure 2. Average scores on laboratory reports from students trained by TA's using the norming method. All report scores were assigned by TA's.

ACKNOWLEDGEMENTS

Assistance was gratefully received from the following individuals at the University of Delaware: Prof. Stephen Bernhardt (Chairperson, English Department); Mary Boggs, Kimberly Milligan, and Kalyani Telu (graduate students), and Amanda Bayley, Georgia Giebel, and Samantha Smith (undergraduate students). The project was funded by the following entities at the University of Delaware: Center for Teaching Effectiveness, General Education Initiative, and Information Technologies.

Appendix I – Chart and Table Construction

CHARTS

Many different types of charts can be found in science journals. In science writing charts are called "figures" and numbered sequentially, e.g. Figure 1, Figure 2. What follows pertains to two of the most common types – XY and bar or column charts – and presents their most basic conventions.

In General

- An XY chart is used for data in which the independent variable is continuous, like temperature, time, and wavelengths of light.
- A bar or column chart is used for data in which the independent variable is discontinuous as with categories such as types of ecosystems, species, and colors (e.g. red, green, blue). It is usually 2D but may be made 3D to improve comprehension.

As you read the following do's and don'ts, note that they are intended to improve clarity which translates into speed of data comprehension for your readers. Readers are very busy people and they want to understand with a minimum of effort!

XY Charts

All quantitative data[1] should be presented graphically in a chart unless they are more suited for a table. **Do not present the same data in both forms; data in a table must be differentiated from data in a chart.** For example, you could show raw data in a table to reveal variation between samples and then show averages in a chart.

- The independent variable is assigned to the X axis. Values on both the X and Y axes increase as they proceed away from the origin. Values do not have to begin at the origin with zero however.
- Data points are shown with geometric symbols such as circles, triangles and squares. These may be filled or empty. Do not show values next to data points.
- Lines should always be used in conjunction with data points. Lines represent expected values between the data points shown. They are usually continuous but may be dashed, dotted, or otherwise broken for clarity, and they may be smoothed or simply connect one data point to the next.
- Black color is preferred for both symbols and lines because this has greater clarity when printed with a non-color printer.
- Each line (series of connected data points) should be identified with a legend that identifies the symbols. The legend may be placed right beneath the chart area or in an empty space within the chart area.
- Each axis is labeled with both a descriptive parameter (e.g. Time on X axis) and a unit (e.g. min). The parameter is capitalized and the unit that follows is enclosed in parentheses, e.g. Time (min).
- Each axis has short crossing lines, called "ticks," indicating where values are positioned along the axis. Only a minimum of ticks should be used, e.g. one every minute of time rather than every second of every minute. Also only a few of the ticks (considered the major ones in Excel) should be given values, e.g. 0, 5, 10 minutes rather than 0, 1, 2, 3, 4, 5, and so on. Doing so improves clarity without compromising comprehension.

[1] Remember, *data* is the plural of *datum*. Use a plural pronoun and the correct verb when using the word *data*.

- The values used on the axis ticks should be rounded numbers in uniform progression. Data values do not have to be plotted such that they line up with a tick. Thus, use 0, 1, 2, etc. ticks for time but not 0, 0.8, 1.6, 2.4. This makes it easier for the reader to locate an intermediate value along an axis.
- The chart area should be kept free of vertical or horizontal lines (called "gridlines" in *Excel*) for improved clarity. The reader can line up data points with axis values using a straight edge.
- **The figure caption should be placed <u>below</u> the figure.** Its text should be a meaningful statement about what the data represents. The goal is to convey enough information that the reader understands what is being presented without consulting the main text, i.e. the figure "stands alone". Complete sentences are not required. After this descriptive statement, one or more sentences may be included that add to reader understanding.
- Figures should be numbered and placed in the same order they are cited in the text. Do not make the reader hunt for the figure.
- A "trend line" (e.g. least squares regression) should be used <u>only</u> if you expect the data to be a linear progression. An example would be the regression of absorbance (Y axis) on concentration (X axis) for a solution of protein.
- Statistical "whiskers" should be included on the data points if either the standard deviation or the standard error of the mean has been calculated. When using *Excel* special steps must be taken to obtain these descriptive statistics.

Column (vertical) and Bar (horizontal) Charts

- Again black color is preferred. If two or more columns or bars are grouped together you can use different types of hatching to differentiate them.
- For figure caption/placement and axis labels instructions given above should be followed. The X axis is reserved for the independent variable just as for XY charts.
- Statistical "whiskers" should be included at the extreme of each column or bar if descriptive statistics of sample variation have been calculated.

TABLES

Tables are used very sparingly in science journals. They are good at showing, in an organized way, quantitative information that has a variety of elements such as a description of a plant in terms of height, number of leaves, age, and so forth. They also are used to present non-quantitative information composed of text such as a list of bacterial mutants used in an investigation.

Although tables can be presented in a variety of ways, certain features are standard.

- **The title is positioned <u>above</u> the table.** We use the term "title" for a table and "caption" for a figure so you will not confuse the two. The title should have enough information that the reader understands what it represents without reference to the body of the report; it does not have to be a complete sentence. Additional information in sentences or phrases may be included. An example would be some explanation of methods specific to the data in the table (i.e. not general methods presented in a separate section) or identification of symbols used for results of statistical tests such as * for significant (95% confidence) and ** for highly significant (99% confidence).
- Tables are numbered consecutively in the order they are referenced in the text and independently of figures. Thus a report may have a Table 1 and a Figure 1.
- Each table is composed of one or more columns and rows. Each column must have a heading composed of a parameter and, if appropriate, units. Examples would be "Species" and "Temperature (^{o}C)." Column headings are placed in the same row and separated from data rows below by a solid line. Vertical lines between columns are generally not used. An enclosing line around the four sides of a table is a good way to distinguish the table from neighboring text.
- Many kinds of tables use the first (left hand) column to define entries in rows. An example would be treatment number or a treatment descriptor.

Appendix II – Eight Writing Rules to Remember

#1 Point of view
#2 Active or passive voice
#3 Tense – past or present
#4 Pronouns
#5 Word Choice – objectivity
#6 Word Choice – tricky word
#7 Concision
#8 Complete sentences and run-on sentences

1. Point of View – 1st, 2nd or 3rd person?

- Avoid 1st or 2nd person in reports.
- Use 3rd person or a neutral subject in lab reports.

Examples:
1st I, we, us, our **2nd** you, yours **3rd** he, she, they, his, etc

NO: *I conducted the experiment,* or *Our results show that...* or *We next filled the tube with solution*.
YES: *The experiment was completed*. (Notice passive voice; see # 2 for more on passive voice.)

2. Sentence Structure – Active or Passive Voice?

In general, lab reports are written in passive voice:

- **Active voice** has an active subject for the verb: *The researchers performed the experiment in three stages*. (Here, you know who did the performing—the researchers. The sentence emphasizes the researchers.)
- **Passive voice** has no active subject. There's a verb, but you don't know who or what is performing the action of the verb: *The experiment was performed in three stages*. (Who is performing? We don't know, but it is not important; thus the three stages are emphasized, not who performed them.)

3. Tense – Past or Present?

- **Introduction Section Only:** consistent present tense allowed
- **Results Section**: past tense used when reporting results
  NO: *Anoles and frogs given CO2-free air show a decline in respiration as the temperature drops.*
  YES: *Anoles and frogs given CO2-free air showed a decline in respiration as the temperature dropped.*
- **Do not mix tenses** within a sentence as in the following:
  Percent error for the obtained unknown one was 59.3% and for unknown two is 68.8%.

4. Pronouns – Do They Match up?

Pronouns (they, it, she, their, etc.) should match the word they refer to (called an antecedent). You wouldn't say, "The child gave their toys away." Here's another example that is trickier because of the prepositional phrase – of soybean root nodules – that comes between the pronoun and its antecedent.

NO: *The clump of soybean root nodules was inserted into their tube*. (Their is incorrect; the antecedent isn't nodules, but clump.)
YES*: The clump of soybean root nodules was inserted into its tube.*

Be careful with pronouns that go with "data." *The data and their implications can be written tomorrow.* [See # 6 for more details.]

5. Word Choice – Are Words Clear and Objective?

5 A. Lab reports should be written as objectively as possible, with **no unwarranted interpretation or subjectivity**. Avoid words like clearly, definite, drastic, obvious, always, never, interesting (such as "the lizard gave an *interesting* response to being probed. . . ." Interesting to whom?) Below is a sentence from a Discussion. *Poorly chosen words are italicized*; better choices are underlined.

NO "Results from this *interesting* experiment *clearly prove* that lizards and frogs decrease their respiration rate as temperature drops, whereas chicks *obviously* increase their respiration in response to this temperature change."
YES: "Results from this experiment indicate that lizards and frogs may decrease their respiration rate as temperature drops, whereas chicks may increase their respiration. . . ."
An alternative to the above YES example could be "The results of this experiment indicate that lizards and frogs undergo a decrease in respiration rate when temperature drops, whereas. . . ."

5 B. Avoid sweeping generalities (especially when the sample size was one and only one experiment was performed).

NO: "*As shown* by this experiment, *animals respire more slowly* when temperatures decline." [Is this conclusion always valid?]
YES: "Results indicate that animals may respire more slowly when temperatures decline*."* [Less sweeping in its conclusion]

6. Word Choice – What's Correct?

a. **Significant** should be avoided unless one means statistical significance. Used in a report, "significant" implies statistical significance.

b. Prove is not allowed. Rather, results "support" (not prove) the hypothesis. This is particularly true of our labs, in which sample size is always small and experiments are rarely repeated sufficiently to create confidence in the results. Avoid implication that results are conclusive (conclusions are always tentative in science, with the expectation that they can be changed with ongoing experiments). See 5 B above.
NO: *We proved our hypothesis to be correct*.
YES: *The results indicate that the hypothesis was correct.*

c. **Data** is a plural noun. **Datum** is a singular noun.
NO: *The data shows interesting changes*.
YES: *The data show incremental change.* (Note change to the word interesting. See 5A for details on that word.)

d. **Bacteria** is a plural noun**. Bacterium** is a singular noun.

e. **Effect/affect:**
Use *effect* when you want the noun or the result**.**
YES: *The effect of the liquid on the powder was immediate.*

Use *affect* when you want the verb.
YES: *The liquid did not affect the overall results.*

7. Information – Is it Concise?
Scientific writing should be concise; do not include information that is obvious to the reader from reading the figures:

NO: *The enzyme concentration and average rate were graphed on an X-Y coordinate in Figure 1 to exhibit the relationship between the enzyme catalase and the amount of time it took for the disk to float back up to the top*.
Or *The data in Fig. 1 were connected in order to determine whether the relationship between average rate and enzyme concentration was linear.*

YES: *Figure 1 shows that there was a linear relationship between catalase concentration and the reaction rate.*
Or *There was a linear relationship between catalase concentration and the rate of enzyme catalysis (Fig. 1).*

(Note that there is no need to point out how a graph was made, so long as a proper title and caption are included and axes are correctly labeled. Nor is it necessary to describe what the X and Y axes represent.)

8. Are the Sentences Complete?
Every sentence must have a subject and a predicate (also called the verb).

NO: *Only by comprising the following constants: (a) the soybean plants are grown under the same condition, (b) no inhibitors are added to prevent or slow the reactions of the enzyme nitrogenase, (c) there are no significant flaws in the experimental setup.*

At first glance this sentence may seem complete. However, try reading it aloud, and you can hear that something's missing. What occurs after "comprising the following constants"? The writer forgot the verb that tells the reader what happens then. Further, who or what is doing the comprising? There's no subject for that. Reading your drafts aloud, especially to someone else, can help you catch these types of errors.

YES: *Only by comprising the following constants: (a) the soybean plants are grown under the same condition, (b) no inhibitors are added to prevent or slow the reactions of the enzyme nitrogenase, (c) there are no significant flaws in the experimental setup, can the experiment yield tentative results.* Now there is a subject (experiment) and a predicate (can yield). The group of words forms a complete sentence now.

Run-on sentences occur when you combine without proper punctuation two or more groups of words (also called clauses) that could be sentences by themselves**. Two could-be sentences combined with only a comma or with no punctuation are said to be "run-on."**

NO: *As the enzyme concentration increased, the time for the paper disk to float to the top decreased, the* ***highest*** *concentration tested in the experiment was 100 units catalase/ml was the fastest to float to the top.* [Note the incorrect word, "highest; it implies height, not amount.]

YES: *As the enzyme concentration increased, the time for the paper disk to float to the top decreased. The disc with the* ***greatest*** *concentration of catalase tested in the experiment, 100 units /ml, was the fastest to float to the top*. [Note change from "highest" to the preferred "greatest."]
Or *As the enzyme concentration increased, the time for the paper disk to float to the top decreased. Overall, the greatest concentration of catalase used in the experiment, 100 units/ml, produced the shortest time to disk flotation, i.e. fastest reaction rate.*

Appendix III – Scoring Rubrics

VERSION 1 (used the first semester by students and TA's)

Use this as a guide when writing your reports and when critiquing the reports of your peers. Your TA will also use this rubric for grading purposes. Your writing should be aimed at a reader who knows basic science, but is **not** familiar with the experiment or its background. You could think of this person as yourself a year or two from now when the details of the investigation have faded. If you were writing a complete formal lab report, you would organize it into four sections in a particular order – Introduction, Methods, Results, and Discussion – and you might include an Abstract which is a condensation of all four sections into one paragraph. In your section of BISC207 you will omit Methods and the Introduction will present only the hypothesis and prediction/s of what you expect to observe.

Use the following scoring key:
A = meets requirements = 3 pts
B = approaches requirements = 2 pts
C = well below requirements = 1 pt
A value of 0 pts may be used if something is missing.
Maximum score = 36

Report Topic:

Author Name:

Assessor Initials:

Style Questions	Score	Pts
1. The Results section is correctly organized. It begins with text, and figures (and tables if present) are cited, and they are placed within the text close to where they are cited.	**A B C**	
2. Figures are correctly constructed based on the handout provided.	**A B C**	
3. Data (figures, tables) are cited in the text, in numerical sequence, and in the appropriate places.	**A B C**	
4. Grammatical and spelling errors are absent. The Results section uses only past tense. Personal pronouns are avoided. Simple declarative sentence structure is used as much as possible.	**A B C**	
Content Questions		
5. A testable hypothesis is present and leads into the report. [More than one hypothesis may be used.]	**A B C**	
6. The Results section text tells the reader what the experimenter observed rather than leaving it up to the interpretation of the reader. There is meaningful text, not statements such as "The results are shown in Fig. 1" and nothing else. The data are described in terms of trends and not each individual datum. All data produced in the investigation and relevant to the hypothesis have been analyzed and reported.	**A B C**	
7. The observations are analyzed correctly. Averages are calculated, and standard deviation is calculated where possible and presented. The equations for these calculations are not included.	**A B C**	
8. The best instrument for data presentation (table versus figure, type of figure) was chosen.	**A B C**	

9. There is a statement about whether the null hypothesis was supported or rejected and therefore the experimental hypothesis supported. The statement "the hypothesis was proved" or any of its variations is absent.	**A B C**	
10. Limitations on data interpretation are presented. Real errors in data collection, not potential ones, are described and explained.	**A B C**	
11. The significance of the observations relative to known data or concepts is presented.	**A B C**	
12. There are suggestions for realistic and meaningful new investigations.	**A B C**	

VERSION 2 (used between semesters by the project assessment team)

Use the following scoring key:

No points if item is missing
1 point = well below requirements
2 points = approaches requirements
3 points = meets requirements
4 points = meets requirements extremely well
Maximum score = 44

Code:
Report# 1 4 (circle)
Assessor Initials:
Critical Thinking Score:
Writing Score:
Overall Score:

Critical Thinking Skills Questions	Scores (circle)
(1) Results Section. The Results section text tells the reader what the experimenter observed rather than leaving it up to the interpretation of the reader. There is meaningful text, not statements such as "The results are shown in Fig. 1" and nothing else. Sentences referring to a figure or table without including a summary statement on content are avoided. The data are described in terms of trends and not each individual datum. The best report includes all data relevant to the hypothesis or scientific question.	**1 2 3 4**
(2) Data Quality and Presentation. The data are good quality, i.e. useful. The most appropriate method of data presentation (table versus figure, type of figure) was chosen. The most appropriate method allows the reader to quickly and correctly interpret the results.	**1 2 3 4**
(3) Discussion. There is a statement in the Discussion about support of the hypothesis. In the best report justification for this statement is present. The statement "the hypothesis was proved" or any of its variations is absent.	**1 2 3 4**
(4) Discussion. Any limitations on data interpretation are presented. Real errors in data collection (if present), not potential ones, are described and explained. Errors are not fabricated if absent. A statement about absence of errors is present if justified.	**1 2 3 4**
(5) Relevance. In the Discussion comparison of the observations to known data or concepts and the general applicability of the observations are presented. Reference may be made to the course textbook for this.	**1 2 3 4**
(6) Extension. There are suggestions for realistic and meaningful new investigations connected in some logical way to the current investigation.	**1 2 3 4**
Writing Skills Questions	**Scores (circle)**
(7) Overall Report. The report is organized into three sections (Introduction, Results, and Discussion) with a heading for each. The appropriate content is in each section: **Introduction** has a scientific question and/or hypothesis, **Results** has observations made and no discussion, and the **Discussion** revisits the hypothesis and interprets the observations. Methods should not be presented anywhere in the report except sparingly and briefly in a table or figure legend as necessary for understanding.	**1 2 3 4**
(8) Figures. Figures are correctly constructed based on the handout provided.	**1 2 3 4**
(9) Results. The Results section is correctly organized with text preceding figures (and tables if present). Data (figures, tables) are placed close to where they are being presented in the text, they are cited in the text in numerical sequence, and the	**1 2 3 4**

citations are in the appropriate places.	
(10) Mechanics. Grammatical and spelling errors are absent. The Results section uses only past tense. Personal pronouns are avoided.	**1 2 3 4**
(11) Style. Writing style and content are concise and avoids inflated prose (padding).	**1 2 3 4**

VERSION 3 (used the second semester by students and TA's)

Use the following scoring key:
0 point = item is missing
1 point = well below requirements
2 points = approaches requirements
3 points = meets requirements
4 points = exceeds requirements
Maximum score = 42

Report Topic:
Author Name:
Assessor Initials:

Critical Thinking Score (26):
Writing Score (16):
Overall Score (42):

Critical Thinking Skills Questions	Value	Score
(1) Introduction Section. Some background information on the topic and a clearly written, testable hypothesis with justification are presented. When an investigation is not "hypothesis driven" a scientific question with justification is substituted.	**0 - 4**	
(2) Results Section Overall. The Results section text tells the reader what the experimenter observed rather than leaving it up to the interpretation of the reader. There is meaningful text, not statements such as "The results are shown in Fig. 1" and nothing else. Sentences referring to a figure or table without including a summary statement on content are avoided. The data are described in terms of trends and not each individual datum. All available data relevant to the hypothesis or scientific question are represented.	**0 - 3**	
(3) Results: Data Quality and Analysis. The data are of good quality and analyzed correctly including use of statistics where appropriate. They reflect effective lab technique and can be interpreted to yield meaningful conclusions. There is evidence the author clearly understood the observations.	**0 - 4**	
(4) Results: Data Presentation. The most appropriate method of data presentation (table versus figure, type of figure) was chosen. The most appropriate method allows the reader to quickly and correctly interpret the results.	**0 - 3**	
(5) Discussion: Hypothesis and Relevance. There is a statement about support or rejection of the hypothesis with justification. If there is more than one hypothesis all are addressed. Any statement or implication that the hypothesis was "proved" is absent. Also in the Discussion results are compared to published data or concepts and discrepancies explained. The course textbook and scholarly online resources may be used for this.	**0 - 4**	
(6) Discussion: Limitations. Any limitations on data interpretation are presented with explanation. If errors in data collection are known to have been present their impact on data interpretation is explained.	**0 - 4**	
(7) Discussion: Extension. There are suggestions for realistic and meaningful new investigations connected in some logical way to the current investigation. "Realistic" means the materials would be available and measurements could be made. The best ideas go well beyond merely suggesting more samples or better techniques.	**0 - 4**	

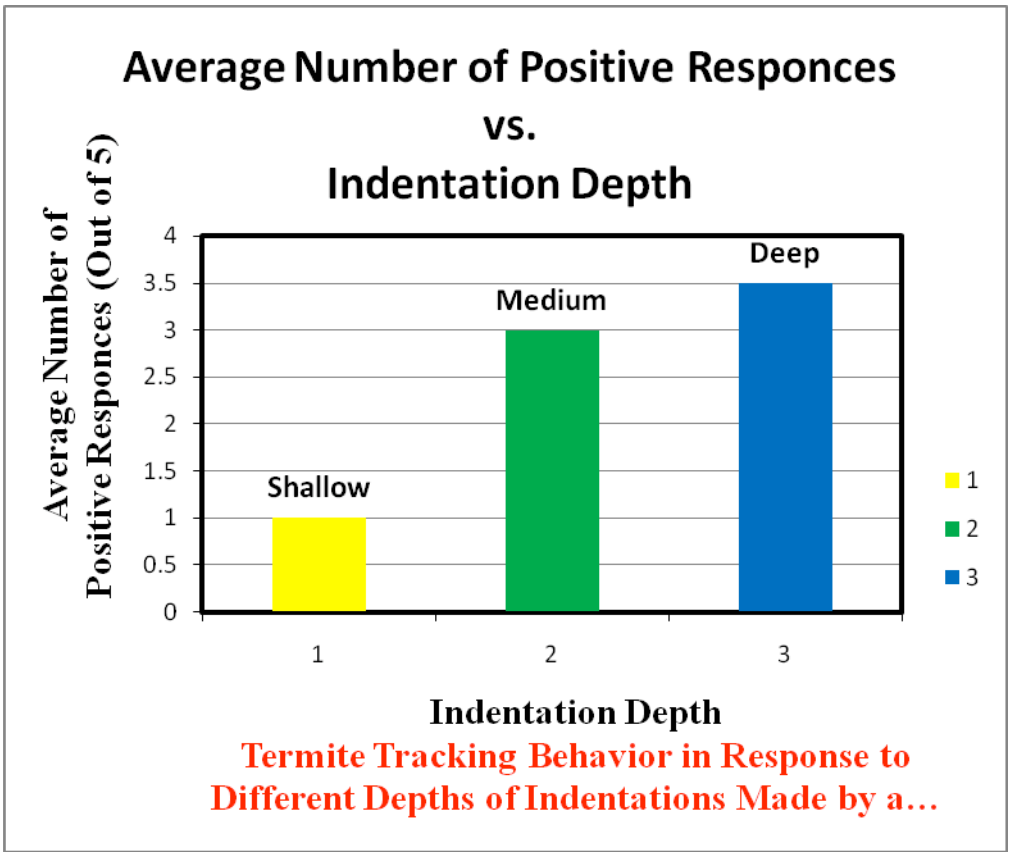
Writing Skills Questions	Value	Score
(8) Report Organization. The report is organized into three sections (Introduction, Results, and Discussion) with a heading for each. The appropriate content is in each section: **Introduction** has some background information and a scientific question and/or hypothesis, **Results** has observations made but no discussion, and the **Discussion** revisits the hypothesis and interprets the observations. Methods should not be presented anywhere in the report except sparingly and briefly in a table or figure legend as necessary for understanding.	**0 - 3**	
(9) Results: Organization. The Results section is correctly organized with text preceding figures (and tables if present). Data (figures, tables) are placed close to where they are being presented in the text, they are cited in the text in numerical sequence, and the citations are in the appropriate places.	**0 - 3**	
(10) Results: Figures. Figures are correctly constructed based on the handout provided and easily understood.	**0 - 3**	
(11) Mechanics. Grammatical and spelling errors are essentially absent. The Results section uses only past tense. Personal pronouns are avoided.	**0 - 3**	
(12) Style. Writing style and content are concise and coherent; inflated prose (padding) is avoided.	**0 - 4**	
Total	**42**	

Appendix IV – Samples of Student Reports Used in Norming

[Evaluations are given after each report.]

Sample Report #1
Termite Behavior Investigation

Hypothesis: If the tracking of termites is related to the depth of the indentation made by a red Paper Mate pen then we predict that the termites will most accurately follow the indentation with the deepest groove.
Results: In the experiment, the tracking behaviors of 5 termites in response to 3 different indentation depths were observed. Two trials were conducted for each of the three types of indentations (shallow, medium, and deep). As shown by the graph, on average 1 termite followed the red line in its entirety when the indentation made by the pen was shallow, while three followed the red line in its entirety when the indentation was of medium depth. Finally, 3.5 termites followed the red line when the indentation was deep.
Conclusion: One can conclude that the tracking behavior of the termites is related to indentation depth since only 1 termite followed the shallowest groove, while 3 followed the groove of medium depth, and 3.5 followed the deepest groove. Therefore, the results support the hypothesis since the most positive responses occurred when the groove was the deepest.



Score on Sample Report #1

Use the following scoring key:
0 point = item is missing
1 point = well below requirements
2 points = approaches requirements
3 points = meets requirements
4 points = exceeds requirements
Maximum score = 42

Report Topic: *Termite Behavior*
Author Name: *Undisclosed*
Assessor Initials: *Undisclosed*

Critical Thinking Score (26): *9.5*
Writing Score (16): *13*
Overall Score (42): *22.5/42 =* ***54%***

Critical Thinking Skills Questions	Value	Score
(1) Introduction Section. Some background information on the topic and a clearly written, testable hypothesis with justification are presented. When an investigation is not "hypothesis driven" a scientific question with justification is substituted.	**0 - 4**	***1***
(2) Results Section Overall. The Results section text tells the reader what the experimenter observed rather than leaving it up to the interpretation of the reader. There is meaningful text, not statements such as "The results are shown in Fig. 1" and nothing else. Sentences referring to a figure or table without including a summary statement on content are avoided. The data are described in terms of trends and not each individual datum. All available data relevant to the hypothesis or scientific question are represented.	**0 - 3**	***2***
(3) Results: Data Quality and Analysis. The data are of good quality and analyzed correctly including use of statistics where appropriate. They reflect effective lab technique and can be interpreted to yield meaningful conclusions. There is evidence the author clearly understood the observations.	**0 - 4**	***2***
(4) Results: Data Presentation. The most appropriate method of data presentation (table versus figure, type of figure) was chosen. The most appropriate method allows the reader to quickly and correctly interpret the results.	**0 - 3**	***3***
(5) Discussion: Hypothesis and Relevance. There is a statement about support or rejection of the hypothesis with justification. If there is more than one hypothesis all are addressed. Any statement or implication that the hypothesis was "proved" is absent. Also in the Discussion results are compared to published data or concepts and discrepancies explained. The course textbook and scholarly online resources may be used for this.	**0 - 4**	***1.5***
(6) Discussion: Limitations. Any limitations on data interpretation are presented with explanation. If errors in data collection are known to have been present their impact on data interpretation is explained.	**0 - 4**	***0***
(7) Discussion: Extension. There are suggestions for realistic and meaningful new investigations connected in some logical way to the current investigation. "Realistic" means the materials would be available and measurements could be made. The best ideas go well beyond merely suggesting more samples or better techniques.	**0 - 4**	***0***

Writing Skills Questions	Value	Score
(8) Report Organization. The report is organized into three sections (Introduction, Results, and Discussion) with a heading for each. The appropriate content is in each section: **Introduction** has some background information and a scientific question and/or hypothesis, **Results** has observations made but no discussion, and the **Discussion** revisits the hypothesis and interprets the observations. Methods should not be presented anywhere in the report except sparingly and briefly in a table or figure legend as necessary for understanding.	**0 - 3**	***3***
(9) Results: Organization. The Results section is correctly organized with text preceding figures (and tables if present). Data (figures, tables) are placed close to where they are being presented in the text, they are cited in the text in numerical sequence, and the citations are in the appropriate places.	**0 - 3**	***2***
(10) Results: Figures. Figures are correctly constructed based on the handout provided and easily understood.	**0 - 3**	***1.5***
(11) Mechanics. Grammatical and spelling errors are essentially absent. The Results section uses only past tense. Personal pronouns are avoided.	**0 - 3**	***3***
(12) Style. Writing style and content are concise and coherent; inflated prose (padding) is avoided.	**0 - 4**	***3.5***
Total	**42**	***22.5***

Sample Report #2
Termite Behavior Investigation

Introduction
It has been observed that termites will follow a line created on soft filter paper with a red Paper Mate ballpoint pen. There could be any of several tracking cues involved. It could be an indentation, the line color, a chemical in the ink, some other variable, or a combination of these. Here it is proposed that termites follow a groove, and it is predicted that they will most faithfully follow the deepest groove.

Results
There was a clear trend in termite response with most tracking occurring with the deepest groove (Fig. 1). To obtain these data the animals had to be placed on the groove. If they were placed elsewhere on the paper surface and allowed to explore they did not appear to identify a groove as something to follow (data not shown).

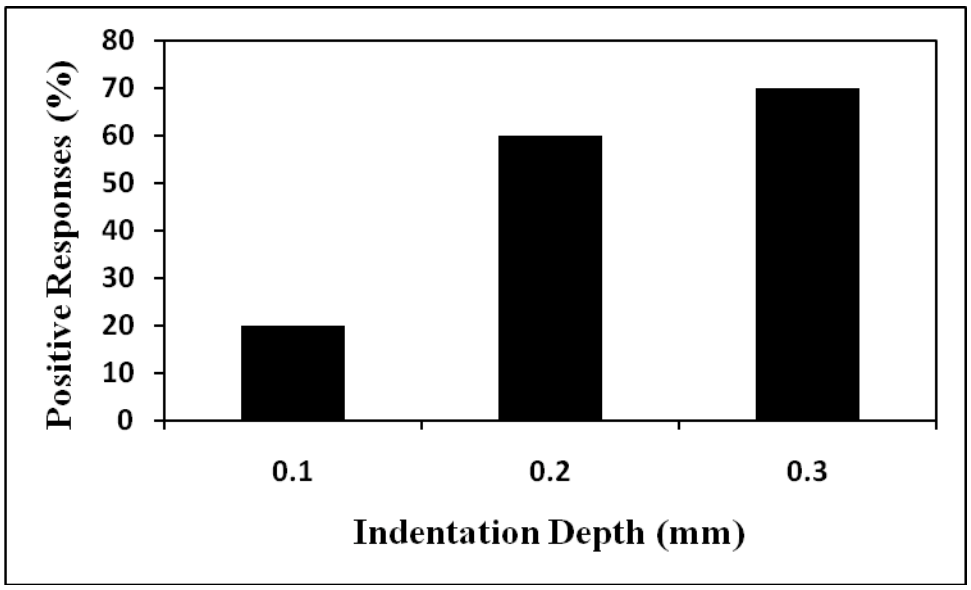
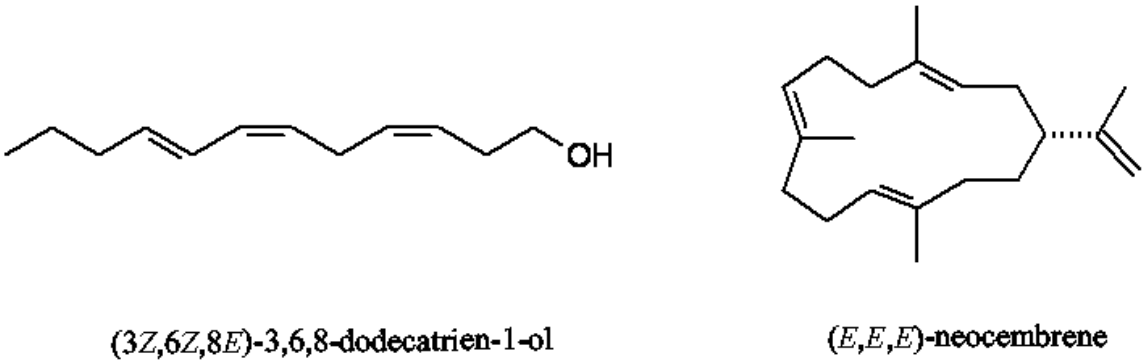


Figure 1. Average termite tracking behavior in response to different depths of indentations made by a red Paper Mate pen. Five termites were tested in two replicates.
A positive response was an individual proceeding along the indentation for its entire length.

Discussion
This study confirms published observations that termites follow some sort of trail. The hypothesis that indentation depth is a trail marker and the deepest groove provides the strongest signal was supported. Each individual gave a clear response with some following a groove faithfully and others wandering away as soon as they were placed. Although the data were not tested statistically, the increase in response between the 0.1 and 0.2 mm grooves appears to be significant, but the increase between 0.2 and 0.3 mm is probably not significant. A sample size larger than five animals might give statistically significant differences between all three treatments.

There is an alternative explanation for the results, however. The indentations were made with a pen which left a mark other than merely an indentation. The deepest groove would be expected to have the greatest deposit of ink and possibly retard more of any volatile chemical from diffusing away. This explanation is supported by the well known fact that worker termites are blind, have a scent-producing sternal gland in their posterior end, and produce at least two volatile trail-marking substances (Fig. 2). It is therefore more likely that termites

followed a substance in the ink that mimicked natural pheromones than a groove. Thus, although the data given in this report support the hypothesis, an alternative explanation was not eliminated because a negative control – groove without ink – was not included.



(3*Z*,6*Z*,8*E*)-3,6,8-dodecatrien-1-ol

(*E*,*E*,*E*)-neocembrene

Figure 2. Known termite trail-marking pheromones. From: http://www.sbs.utexas.edu/jcabbott/courses/bio208web/lectures/scimethod/trail_following_of_termites.htm

Ants are known to follow a trail directed to a food source. Although it is thought to be followed only by the same individual producing the trail substance and therefore very specific, it would be interesting to test some ant species to determine if any pen-produced chemical would serve as a trail cue.

Score on Sample Report #2

Use the following scoring key:
0 point = item is missing
1 point = well below requirements
2 points = approaches requirements
3 points = meets requirements
4 points = exceeds requirements
Maximum score = 42

Report Topic: Termite Behavior
Author Name: Undisclosed
Assessor Initials: *Undisclosed*

Critical Thinking Score (26): *20.5*
Writing Score (16): *15.5*
Overall Score (42): *36/42 =* ***86%***

Critical Thinking Skills Questions	Value	Score
(1) Introduction Section. Some background information on the topic and a clearly written, testable hypothesis with justification are presented. When an investigation is not "hypothesis driven" a scientific question with justification is substituted.	**0 - 4**	***3***
(2) Results Section Overall. The Results section text tells the reader what the experimenter observed rather than leaving it up to the interpretation of the reader. There is meaningful text, not statements such as "The results are shown in Fig. 1" and nothing else. Sentences referring to a figure or table without including a summary statement on content are avoided. The data are described in terms of trends and not each individual datum. All available data relevant to the hypothesis or scientific question are represented.	**0 - 3**	***3***
(3) Results: Data Quality and Analysis. The data are of good quality and analyzed correctly including use of statistics where appropriate. They reflect effective lab technique and can be interpreted to yield meaningful conclusions. There is evidence the author clearly understood the observations.	**0 - 4**	***2***
(4) Results: Data Presentation. The most appropriate method of data presentation (table versus figure, type of figure) was chosen. The most appropriate method allows the reader to quickly and correctly interpret the results.	**0 - 3**	***2***
(5) Discussion: Hypothesis and Relevance. There is a statement about support or rejection of the hypothesis with justification. If there is more than one hypothesis all are addressed. Any statement or implication that the hypothesis was "proved" is absent. Also in the Discussion results are compared to published data or concepts and discrepancies explained. The course textbook and scholarly online resources may be used for this.	**0 - 4**	***3.5***
(6) Discussion: Limitations. Any limitations on data interpretation are presented with explanation. If errors in data collection are known to have been present their impact on data interpretation is explained.	**0 - 4**	***4***
(7) Discussion: Extension. There are suggestions for realistic and meaningful new investigations connected in some logical way to the current investigation. "Realistic" means the materials would be available and measurements could be made. The best ideas go well beyond merely suggesting more samples or better techniques.	**0 - 4**	***3***

Writing Skills Questions	Value	Score
(8) Report Organization. The report is organized into three sections (Introduction, Results, and Discussion) with a heading for each. The appropriate content is in each section: **Introduction** has some background information and a scientific question and/or hypothesis, **Results** has observations made but no discussion, and the **Discussion** revisits the hypothesis and interprets the observations. Methods should not be presented anywhere in the report except sparingly and briefly in a table or figure legend as necessary for understanding.	**0 - 3**	***3***
(9) Results: Organization. The Results section is correctly organized with text preceding figures (and tables if present). Data (figures, tables) are placed close to where they are being presented in the text, they are cited in the text in numerical sequence, and the citations are in the appropriate places.	**0 - 3**	***3***
(10) Results: Figures. Figures are correctly constructed based on the handout provided and easily understood.	**0 - 3**	***3***
(11) Mechanics. Grammatical and spelling errors are essentially absent. The Results section uses only past tense. Personal pronouns are avoided.	**0 - 3**	***2.5***
(12) Style. Writing style and content are concise and coherent; inflated prose (padding) is avoided.	**0 - 4**	***4***
Total	**42**	***36***