Advances in Biology Laboratory Education
Publication of the Association for Biology Laboratory Education
Volume 42, Article 10, 2022

A Six Elements Method for Teaching Scientific Writing at Scale

A. Daniel Johnson

Wake Forest University, Department of Biology, 1834 Wake Forest Road, Winston-Salem NC 27109, USA
(**johnsoad@wfu.edu**)

Scientific communication helps students learn to state problems and present claims precisely, summarize evidence to support those claims, and demonstrate their reasoning. Regular scientific writing with expert feedback is particularly effective at developing critical and applied thinking process skills. The challenge is how to do this consistently in large multi-section introductory biology courses with multiple instructors. We developed a data-rich **Six Elements Method** for teaching scientific writing in our multi-section introductory biology courses. It is based in education research and proven practices from the Writing Across the Curriculum/ Writing in Disciplines (WAC/WID) literature. The general method can be repurposed to a variety of courses. In 2020, we tested our approach in online labs, and found it works equally well with only minor adjustments.

This workshop and discussion focused on: instructor pre-training, active learning exercises for students, & bins-based grading. Before the workshop, participants completed two assignments: an undergraduate text annotation exercise, and report grading using our bins-based grading methods.

Discussion focused on how to use compiled data from a class section to adjust our writing instruction strategy, and participants' questions about the process and implementation challenges.

Keywords: scientific writing, bins-based grading, instructor training, text annotation, data-driven instruction

Link To Supplemental Materials: https://doi.org/10.37590/able.v42.sup10

Introduction

Scientific communication helps students learn to state problems and present claims precisely, summarize evidence to support those claims, and demonstrate their reasoning. Regular scientific writing with expert feedback is particularly effective at developing critical and applied thinking process skills. While these skills are highly transferable to nearly all STEM careers, many students lack the level of skill needed for the STEM workforce. Why?

STEM instructors face several barriers when implementing well-structured lecture and lab writing experiences. First, large multi-section courses are nearly universal, which creates a workload challenge. Second, there is the problem of expertise. STEM faculty and graduate students have experience producing scientific writing, but few have been trained to teach scientific writing, and fewer still know what best practices are. Third, faculty may not have direct control over the instructional process; they may be supervising graduate or undergraduate TAs who do the actual student training and grading.

Our Approach

Wake Forest University (WFU) Biology graduate TAs (GTAs) grade >3100 lab reports from >850 students annually. With NSF support we developed a data-rich Six Elements Method (SEM) for teaching scientific writing in our multi-section introductory biology courses. Starting August 2016, we transitioned three longstanding introductory labs for majors (each emphasizing a different sub-discipline) to this Six Elements Method. Two new lab courses that replace the existing courses launched in August 2019; they also teach writing this way. In 2020, we tested this approach in our online labs, and

found that it works equally well with only minor adjustments.

SEM is not meant to be a fixed set of activities, but rather a data-driven approach to writing instruction and assessment overall. It combines:

1. A single standardized style and format that we use across multiple courses;
2. Structured training activities for the writing instructors and graders;
3. Active learning exercises for undergraduates;
4. Automated, form-based pre-submission review of reports prior to grading;
5. Bins-based (vs. points-based) grading; and
6. Scaffolded reflective coaching feedback.

Each element provides data that inform instructional goals. For example, Elements 1 and 4 provide students with clear end goals for their writing, and low stakes feedback to support self-correction. Data from Elements 2, 5, & 6 let us monitor variation between instructors while Elements 3, 4, and 6 show us students' current knowledge and skills, and surface gaps that need additional work.

Participants completed two activities:

- A short text annotation exercise that we give to GTAs, then again to the undergraduates. The summarized results are used to adjust writing instruction strategy.
- Scoring 3-4 sample student reports from our archives using our bins-based scoring model.

Group discussion focused on challenges to implementing the types of instructor pre-training, active learning exercises, and bins-based grading demonstrated by the pre-meeting activities.

Workshop Outline

EXERCISE 1: Reading Annotation Exercise

Pre-Workshop Instructions for Participants

This is an abridged version of a text annotation exercise that our GTAs and students complete at the start of our lab unit on Scientific Writing. Before doing this assignment, our students read and take a short quiz on the ***BioCore Resource Guide***, which is our local manual outlining writing requirements, basic biostatistics, and data visualization. Students complete a short practice section during lab, then finish the remainder of the annotation at least 24 hours before their next lab meeting. The annotation is done online, and GTAs get a summary report of their section. GTAs then use the data from the annotations to adjust which topics they emphasize during the following week's instruction.

Please complete the exercise based on what your own local writing expectations are for introductory courses; one of our discussion goals will be to surface differences in expectations between institutions.

Imagine yourself as a student with no prior experience reading scientific journal articles or writing scientific lab reports.

1. Read the Student Instructions (next section), then complete the in-class practice session. Follow the instructions for the paper version of the exercise.
2. When you understand the procedure, complete the two homework blocks. (Note: this is an abridged version; the most recent version of the full exercise (in Supplement 1) has 7 blocks and takes ~1 hour to complete.)
3. Please keep written notes to bring to the workshop about which questions or terms were unclear, and what you personally think students are MOST likely to get wrong or have misconceptions about.

Instructions for Students

This is a low-stakes diagnostic exercise. You will be graded mainly on effort, not whether you give the right answers. We have three reasons for giving you this assignment.

- It introduces terms we will use as we talk about scientific writing, and let you practice using them.

- Next week your class will talk about lab report writing requirements in detail. To make good use of time, your GTA will use a summary of your class's responses to this exercise to decide what topics to talk about in depth, and what topics everyone understands already.
- Data show that errors students make on this annotation exercise predict what errors they are likely to make on lab reports. If you learn how to avoid mistakes you make now, your lab reports will improve.

Try out the Practice Example in class. Once you see how it works, you can complete the remainder of the exercise for homework.

If you are using **our online interactive version**, open the link to the exercise provided by your GTA.

1. Read each prompt to find out what you should mark.
2. Click and drag to highlight one or more words. Then click one of the displayed annotation tags to mark the highlighted text. To remove a tag, highlight the text again and click "Remove" from the label options shown.
3. Mark longer phrases or sentences first. Then go back and mark shorter phrases or individual words. If you do not see an example of an item, simply leave that annotation tag unused.
4. Answer the follow-up questions.

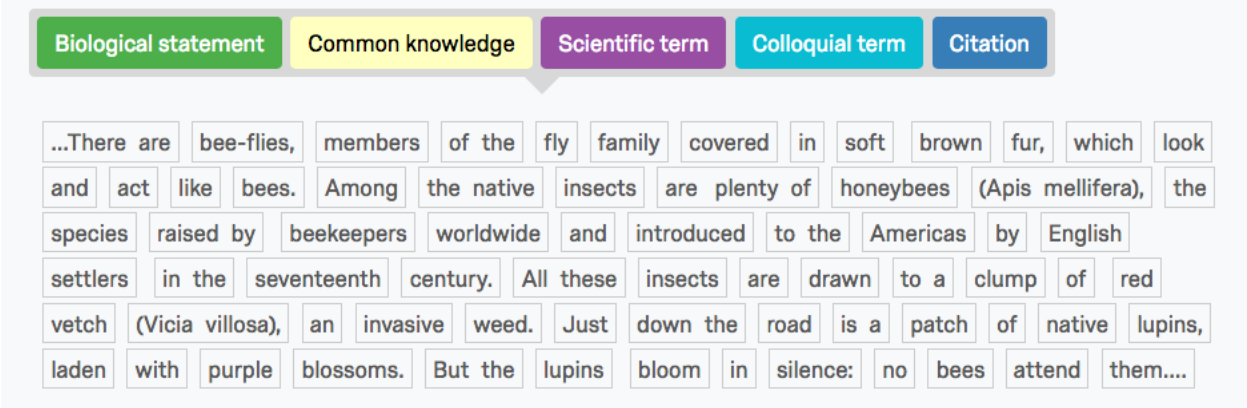
If you are using the **paper version of the exercise**, pick up a pre-printed copy from the desk in the lab and follow these instructions.

1. Read each prompt to find out what you should mark.
2. Circle or highlight your chosen text phrase or word.
3. Write a short obvious abbreviation over the circle, like "BioSt", "SciTm", or "ClqTm" (short for "biological statement," "scientific term," and "colloquial term," respectively.)
4. Answer the follow-up questions.

Practice Example

The paragraph below is a short excerpt from a published article.

1. Find and mark 3 biological statements (statements about an established biological fact or result that needs to be supported with evidence).
2. Mark any common knowledge statements.
3. Mark all scientific terms that someone without college-level biology would not know.
4. Mark all colloquial terms you think are not appropriate for a scientific article.
5. Finally, label all citations you can find.



Follow-Up Questions

Based on the wording of this example, for what audience was the text written?

- O Scientific Community
- O General (Lay) Public

Does this text read like it is part of a scientific paper?

- O Yes
- O No
- O I don't know

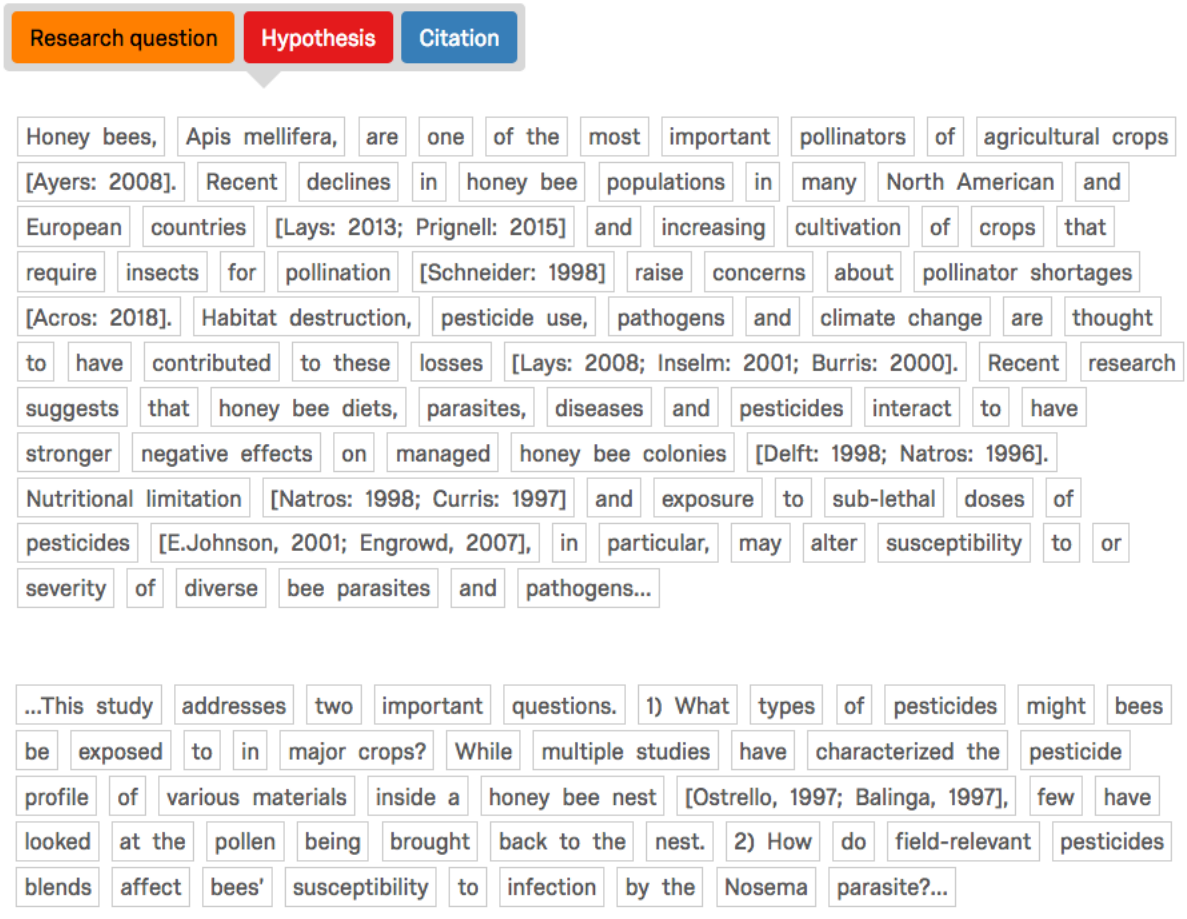
Homework

The rest of this assignment will take about 1 hour to complete.

Text Block #1

Annotate the text below by clicking on phrases or words and choosing from the given labels. Mark whole sentences first and then mark individual terms.

1. Find and label the research question or hypothesis.
2. Label 2 citations.
3. Do not use a label if that item is not present.



Block 1 Follow-Up

Given its contents and structure, where does this text belong in a scientific paper?

- O Introduction
- O Abstract
- O Materials and Methods
- O Results
- O Discussion

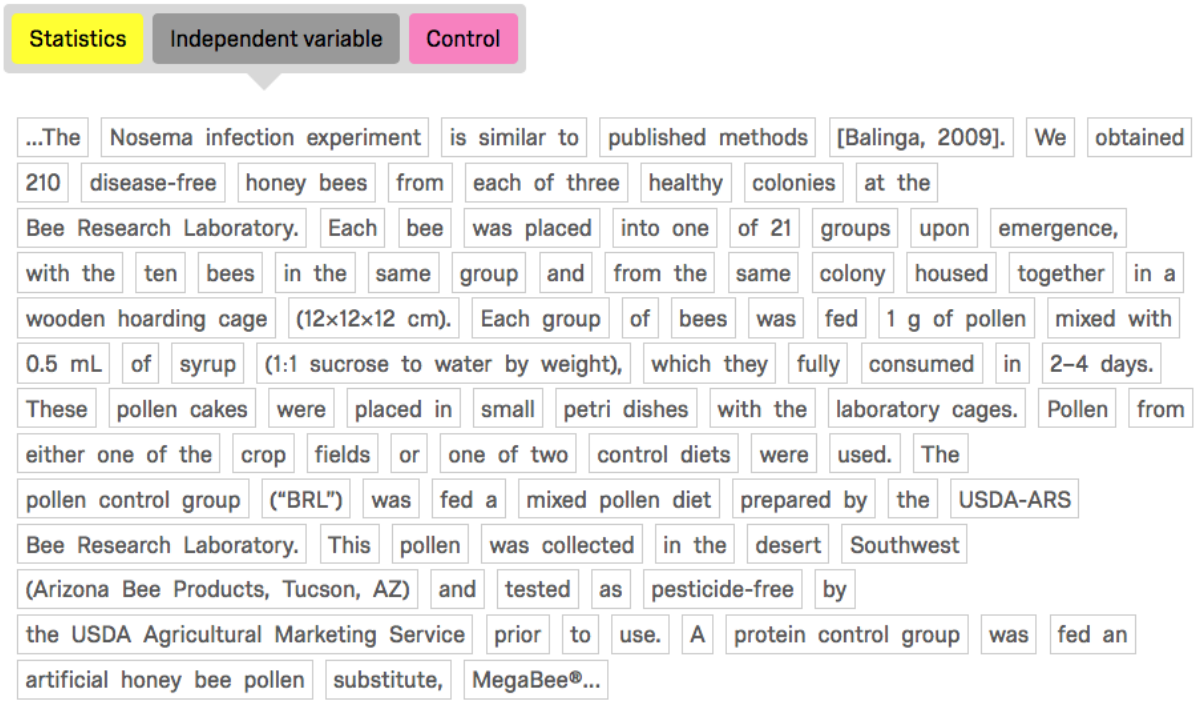
Which of the following statements are true about the Introduction of a scientific paper?

	No	I don't know	Yes
Contains biological statements.	O	O	O
Has citations.	O	O	O
Shows reproducibility of the experiment.	O	O	O
Objectively states outcomes.	O	O	O
Is a summary of the paper.	O	O	O
Is written in past tense.	O	O	O
Contains hypothesis or research goals.	O	O	O
Explains why study is relevant.	O	O	O
Subjectively interprets the findings.	O	O	O

Text Block #2

Annotate the text below by clicking on phrases or words and choosing from the given labels. Mark whole sentences first and then mark individual terms.

1. Look and label all statistical analyses or procedures.
2. Identify the control and independent variable in the experiment.
3. Do not use a label if that item is not present.



Block 2 Follow-Up

Given its contents and structure, where does this text belong in a scientific paper?

- O Introduction
- O Abstract
- O Materials and Methods
- O Results
- O Discussion

Which of the following statements are true about the Introduction of a scientific paper?

	No	I don't know	Yes
Contains biological statements.	O	O	O
Has citations.	O	O	O
Shows reproducibility of the experiment.	O	O	O
Objectively states outcomes.	O	O	O
Is a summary of the paper.	O	O	O
Is written in past tense.	O	O	O
Contains hypothesis or research goals.	O	O	O
Explains why study is relevant.	O	O	O
Subjectively interprets the findings.	O	O	O

EXERCISE 2: Using Summary Data From Annotation Exercise

We use this exercise to teach GTAs how to interpret the data collected from the online interactive exercise. Their goal is to use the summary data to adjust what topics they emphasize as they talk with students about writing. The accompanying set of slides for this exercise are reprinted in Appendix A; the Powerpoint file is in Supplement 2.

Pre-Workshop Instructions for Participants

Put yourself in the place of the GTA leading a lab section. Now look at the reprinted in Appendix A.

1. Slides 1-12 provide background about how the annotation activity fits into our instructional goals, and how GTAs might structure their class debriefs. Read Slides 1-12 so you understand the goals for the debrief sessions within the larger context of teaching scientific writing.

2. Slides 13-24 are an example of the summarized data from ONE section of 16 undergraduates for the questions posed in Exercise 1 (live data reports are different for each instructor and section.) Look at the data and take written notes about which misconceptions about scientific writing are more important to correct, and which errors you could give lower priority. Also write down ideas for how you might adjust what you tell students about scientific writing based on these data.

3. Slides 25-29 show ways that GTAs can manage the Week 2 debrief with their students. Compare our suggestions with your ideas. Please keep written notes to bring to the workshop for discussion.

EXERCISE 3: Bins-Based Report Grading

We no longer grade lab reports using a points-based rubric. Instead we use a bins-based scoring model that is very different from how most GTAs were evaluated. We train new GTAs by having them score a set of "training reports" of different quality, then debriefing their scores with experienced GTAs and faculty.

Pre-Workshop Instructions for Participants

Imagine yourself as a novice GTA grading lab reports. Read the training guide and scoring guidelines, then:

- Using the bins-based method, assign a score of 1-4 to each of the 4 lab reports provided.
- As you grade, please take written notes about how long it took to grade each report, items in our criteria that made it hard to assign a score, and questions you have about implementation. Bring these to the workshop for discussion.

Instructions for GTAs

Overview

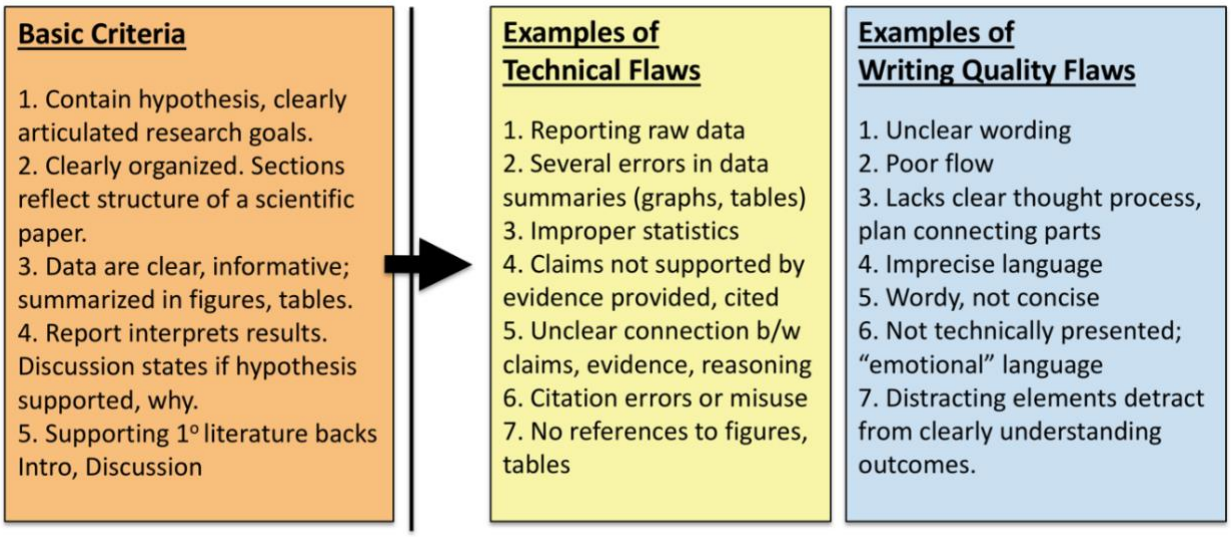
When grading and commenting on lab reports, you should be your students' writing coach, not a copy editor or proofreader. We also want report scores to be consistent between GTAs for similar quality work. Finally, we do not want you to spend excessive time grading each report. Ideally an experienced GTA will spend 10-15 minutes on an initial submission, and under 5 minutes on a revised report.

To make all this possible we use a bins-based scoring method. You won't use a points-based rubric. Instead, you will give each report one of these "bins scores" that reflect the quality overall:

4 = Acceptable (translates in the gradebook to a score of 95, or A)
3 = Needs Minor Revisions (translates to a score of 85, or B)
2 = Needs Major Revisions (translates to a score of 75, or C)
1 = Submitted but Unacceptable (translates to a score of 65, or D)
0 = Not Submitted (translates to a score of 0)

What to Do As You Grade	Why Do It This Way
Focus on the largest flaws and errors first. Work down to smaller errors later.	Making small corrections does not compensate for big uncorrected errors, and students often cannot distinguish between them. You want to focus students' attention on the issues that have the greatest impact on overall score first.
In comments, ask questions that encourage students to reflect on their writing.	The goal of "reflective coaching" is to model the kinds of questions students should ask themselves. With practice, students start to adopt these questions and self-correct.
Limit the number of comments you give to 3-5 per text page, plus some overall comments about the whole report.	Students only process and respond to a limited number of feedback items. Also, students tend to weight all comments equally. When they get too many comments, they tend to make simpler corrections first and leave larger, more important issues uncorrected.
Refer students to appropriate resources for guidance.	We provide extensive guidelines. Rewriting basic information that students can and should find on their own is not using your grading time efficiently. Students need to develop the habit of seeking out their own answers instead of looking to us for them.
Do not copy-edit unless absolutely necessary.	It is appropriate to point out where writing is vague or unclear, but again students need to develop the ability to self-correct rather than expecting they will be shown what to "fix" every time.

Scoring Criteria



Criteria for scoring lab reports are divided into **Basic Criteria**, **Technical Flaws**, and **Writing Quality Flaws**.

- The Five Basic Criteria are **minimum** requirements; either a student's report meets them, or it does not.
- The items listed under Technical Flaws or Writing Quality Flaws are the most frequently encountered errors; there could be other flaws.

Workflow

Grading a stack of 32 lab reports should not swallow up your entire week, or even a weekend. To avoid that you need to:

- Allocate your time appropriately.
    - Plan for 10-15 minutes per report for initial submissions, 5-10 minutes for revisions.
    - Keep a timer on hand; your phone usually works well. If you fall behind, decide whether you are tired and should take a break, or are spending too long on each report.
- Avoid letting one report take over your time.
    - Occasionally a report needs so much work that a face-to-face meeting with the student to discuss the problems would take less time than writing out comments. When this is the case, stop and schedule a meeting.
- Grade in a way that helps students but ALSO is time efficient.

This is how we recommend grading so you are more efficient. Use this the first few times while you gain more experience. With time you may find another way that works better for you. That is fine; applying our expectations consistently is more important than how you get there.

First Pass Sorting

1. Open each report in MS Word.
2. SKIM it (**1 minute or less**), looking for the 8 features in the table below. When you see one, attach an empty comment box (you will refer back to these in the next step.)
3. As you skim, divide the reports into 3 **provisional** groups.
    - Clearly unacceptable. One or more basic criteria are obviously missing.
    - See some flaws.
    - No obvious flaws.

First Pass Features	Group
1. Are all required **sections** there?	"No" on ANY item means report goes into "**Unacceptable**" group
2. Do you see **citations** in Introduction AND Discussion? Look for [Name: Year]	
3. **Quickly** read last 1-3 lines of Introduction. Is there a **hypothesis** near end of Introduction?	
4. Is there **a table or figure** summarizing data?	
5. Quickly skim first **1-3 lines of Discussion**. Does author reference their hypothesis?	"No" should go into "**Some flaws**" group – could be moved back to Unacceptable later
6. Does the **flow and wording** sound reasonable for a technical audience?	"No" should go into "**Some Flaws**" group
7. Do **figures or data tables** at end look right? 8. Do **citations** at the end look generally right?	Present, but you see errors means report goes in "**Some flaws**" group.
Nothing stands out in first brief skim through	Put in "**No obvious flaws**" group

Second Pass: Double-Check and Read Deeper

This time read the full text of each report. Try not to spend more than 4 minutes per report (1 minute/per page). Don't grade one entire group at once. Take a report from each group in turn. Why, you ask?

- It helps you avoid getting frustrated when grading.
- You are more likely to subconsciously change grading standards if you keep grading reports of similar quality.
- Your first pass was an initial sort only. As you re-read you may see that you sorted a report incorrectly and should move it into a different group.

You have three goals this time.

1. You already marked several items with comment boxes. This time you should confirm that they are actually present/ flawed/ absent.
2. Find the 2-3 **highest impact corrections** that each student needs to make. These are what you will point out in your reflective coaching comments that go on the first page of the report, with the student's bins score. Remember, these comments should directly reference the criteria.
3. Identify and provide short comments on other errors. Limit these to 3-5 per page. Address these errors by:
   - Asking reflective coaching questions, or
   - Referring students to the Resource Guide or other reference sources.
   - Do not copy-edit.

More specific strategies for reading and marking up a report from each group the second time through.

- **Unacceptable Group:**
  - If one of 5 Basic Criteria is indeed missing, keep the report in this group.
  - Correcting Basic Criteria errors should be the student's top priority. On the front page, *list which required items are missing, and the score.*
  - If you make comments about technical or writing flaws, put them in the short comments on individual pages.
  - Do not spend an excessive amount of time. *A report that does not meet basic criteria should take LESS time to grade, not more.*
- **Some Flaws Group:**
  - As you read these, decide which of these 3 sub-groups is most appropriate:
    - Flaws in *writing* only. Report gets a Bins Score = 3 (Minor Revisions Needed)
    - Flaws *in technical execution* of stats, figures, tables, etc., only. Report gets a Bins Score = 3 (Minor Revisions Needed)
    - Flaws in *BOTH* writing *AND* technical execution. Report gets a Bins Score = 2 (Major Revisions Needed)

- As you divide the reports, look for the larger/global errors the student should address *first*. What are 2-3 corrections that the student could make that would make the report fundamentally better?
- In the front page comment, summarize the most important corrections needed, and the score.
- Add no more than 3-5 short comments per page. **Use these comments to point out smaller corrections, not the global issues.** Page-level comments should be questions or refer to other sources whenever possible.

- **No Obvious Flaws Group:**
  - Double check that you did not overlook any writing or technical flaws.
  - Identify 2-3 points where you think the report could be improved.
  - In the front page comment, summarize the most important areas the student could improve, and the score. Report gets a Bins Score = 4 (Acceptable)
  - Add no more than 3-5 short comments per page. **Use short comments to point out smaller corrections, not the global issues.** Page-level comments should be questions or refer to other sources whenever possible.
  - Remember that even if a report is in the highest possible group for grading, it can always be better.

Sample reports for you to practice scoring are provided in Supplement 3.

Materials

The procedure described does not require specific equipment or software. Previously we used the institutionally licensed Qualtrics survey software to collect lab reports, and to run the annotation exercise entirely online. We now use a paper version of the annotation exercise and collect reports via our Canvas LMS. Both ways have worked equally well.

Supplement 2 contains the PowerPoint slides that we use to train GTAs and is the source of the figures in Appendix A. Examples of reports written by our students are provided in Supplement 3. The bins score that we gave each report are included as the first digit in the R_id# at the top of the report; for example, R_4zFUf… means the report earned a score of "4/Acceptable."

Notes for the Instructor

The Six Elements Method (SEM) for teaching and grading writing described here is not meant to be a fixed set of activities, but a data-driven approach to writing instruction and assessment overall.

Impacts on Student Effort, Grades, Attitudes, and GTAs

When we first developed this approach to writing instruction, we selected 5 "basic criteria" as the minimum effort required for students to earn a score of C or higher on lab reports. Looking at a historical sample of reports, 36% of the initial submissions (work that students thought was "grading-ready") and 10.4% of **revised** reports were missing at least 1 of these 5 minimum basic criteria. We do not think this is intentional, but rather suggests that many introductory level students do not fully understand our expectations, so are **unable** to recognize that their submitted work does not meet requirements. Students' answers to annotation follow-up questions about writing criteria showed a similar pattern of misunderstanding expectations.

We have been extremely pleased by the results from switching to SEM. One year after implementing it, only 5.5% of initial submissions and 2.0% of revised reports failed to meet all 5 minimum basic criteria.

SEM also reduced variation in scores between GTAs. Prior to implementing it, mean scores assigned to lab reports varied 6-11% between GTAs for the same assignments in the same course, and up to 12% between GTAs assigned to different courses. Post-implementation, both between-GTA and between-course variation declined to 3-4%. Inter-GTA grade variation remained at 6-12% in a control course that did not adopt the new instruction and grading model.

Student attitudes towards writing also improved. On final course evaluations, the average number of negative open response comments about writing instruction or grading dropped 28.8%, from 46.7% of all student comments pre-implementation to 17.9% post-implementation.

As more of our faculty learn how bins-based grading works, more are adopting it. Most recently, lecture instructors in our 100-level majors sequence have adopted bins-based grading for scoring open-response questions on the lecture exams.

Feedback From VIABLE 2021 Workshop

On post-workshop evaluations, many participants said they were looking for a more effective and time-efficient way to teach and grade student writing. As one participant pointed out, "grading efficiently and fairly remains a point of anxiety for even seasoned faculty."

Another participant made this very insightful observation: "The bins-based grading was a new concept to me & I enjoyed learning about how it worked. Although, I probably subconsciously already put my student's papers in bins!" That is one of the key ideas behind this method: it formalizes what we do already. Through experience we develop a professional sense of which students' writing meets expectations, needs minor or major revision, or is fundamentally flawed. Points-based grading is thought to provide unbiased justification for student grades, but in fact can be very subjective, especially when working with GTAs, who may have very different ideas of what is or is not acceptable. Bins-based scoring helps GTAs develop this professional skill sooner, while simultaneously teaching students norms of professional writing.

Participants also shared many useful insights and good follow-up questions during the sessions.

Design Process

Several participants asked how we chose our ~20 criteria used in the bins scoring model. We began by asking "what are the most common errors we have seen in the past in student writing, that keep students from getting the highest grade possible? We found 40-50 common errors that were our starting point for establishing the bins scoring framework.

Next we asked, "which of these are the 5-6 biggest, most fundamental errors that our students make that they need to fix first? What are we spending the most time telling students to fix, that students should be able to fix themselves?" These became our basic criteria.

Then we asked, "do the other common errors have a pattern? Could some of these be combined into more general classes?" For us, common errors could be divided into technical flaws (improper statistics, not graphing properly, citation errors, etc.) and writing flaws (poor flow, improper language, not telling a complete, logical story, etc.)

One suggestion was to add a fifth scoring bin. This was not possible during the NSF study but is something we are considering for the future. One idea is to have 3 (vs. 2) general criteria above the basic level: technical flaws, writing flaws, and logic and evidence (or argument) flaws.

Buy-In and Resistance

As one participant pointed out, students may want detailed rubrics. Like many participants, our students are obsessed with knowing the exact number score and points they received. We know from experience that GTAs can be just as insistent on having something with points to add up. How do we overcome resistance to bins-based scoring?

We found introducing it at the very start of the lab sequence reduced resistance, because it is the only way our students have gotten feedback on their writing. This took time; we continued to have pushback as long as there were students and GTAs who had worked under a points-based grading model. Once those cohorts passed through, bins-based grading became the norm.

Some lingering resistance is inevitable, but we have found two effective arguments against going back to points. Students buy into the first argument quickly: lab grades are higher. We looked closely at grades and writing quality before and after implementing bins-based scoring, and found that writing grades improved significantly, due mostly to a large drop in the number of basic mechanical errors. This in turn shifted student and GTA attention towards higher-order elements of writing sooner. Eliminating points forces students and GTAs to engage in a larger conversation about how ideas are communicated in science rather than point-harvesting arguments.

Our other argument for this approach is that most professional writing (i.e., grant proposals, journal manuscripts) is evaluated holistically, not point-by-point.

Bins-based grading is a form of specifications grading; Nilson (2014) outlines the rationale for specifications grading in detail for anyone needing still more supporting arguments.

One of the perennial arguments from faculty is that it is impossible for students to earn a perfect score of 100%. We have three counter-arguments. First, the bins scores convert to letter grades; we use the median percent grade for that letter grade. Second, we have data to say confidently that this grading strategy does not depress class averages, and in fact raises them. Third, earning a 100% score means the report is perfect. It does not have a single error in the logic and rationale, methods, data analysis and visualization, citation of sources, or writing mechanics (including spelling and grammar.) There is nothing that can be changed or improved. Awarding a perfect 100% sets up false expectations for students who may assume that the same level of writing effort will still be worthy of perfect scores in more advanced courses. This is not hypothetical; we have had several students try to use 100% scores on reports in 100-level courses as evidence that they were graded unduly harshly by GTAs in higher level classes.

Goals & Design of Homework Assignments

Several participants had questions about how we designed the homework assignments. Some of our writing-related assignments are designed to let students practice specific skills and get feedback. Other assignments like the annotation activity provide actionable data about what students know so the instructor can address misunderstandings before they impact grades.

The annotation activity is meant to evolve over time, adding and removing blocks depending on what flaws we see in students' writing. To create our initial version of the annotation exercises, we took a critical look (with GTAs) at our writing requirements, asking: "what parts of the instructions do students seem to misunderstand?" Then we designed and tuned the annotation homework assignment to test their knowledge of those items.

For example, our students historically have struggled to identify where citations are needed, and to insert them correctly. To explore this, one of the blocks in our first version of the exercise focused specifically on identifying the specialized terms and concepts that warrant citation, vs. common knowledge. When we looked at the summary data after the first semester, we found ~30% of students mis-identified concepts and terms that needed citations, so we kept that block of questions for the subsequent semesters.

Our first annotation exercise was much longer than the current version, taking students more than 2 hours to complete. Looking at the summary data for that first semester, we found that >90% of students correctly answered over half of the questions we asked. These low-error questions were eliminated, and only items where more than 10% of students answered incorrectly stayed in the

annotation exercise for the next semester. After two more semesters we had an annotation activity that aligned very well with the most common errors that GTAs saw on reports.

Implementation Challenges

Many participants had concerns about how to implement SEM without computer support. We built a complex lab report collection platform to meet research needs, not because SEM requires specialized technical support. For example, a simple web form is sufficient to address one of the most pervasive basic writing problems: incomplete reports. Supplement 4 shows a simple Google Form for collecting lab reports. Students cannot submit an incomplete lab report because all questions are set as "required."

Most institutional LMSs (Canvas, Moodle, Blackboard, D2L, etc.) provide even richer survey and assignment-building tools. Another alternative is to use an institutionally licensed survey system like Qualtrics or SurveyMonkey. These work extremely well if they do not have to be closely integrated with other computer systems.

New tools are constantly becoming available. For example, we are retiring both the Qualtrics online version (shown in Exercise 1) and paper version in favor of a new annotation activity built on the free web collaboration tool Hypothes.is. Our advice is to look at what technologies are available locally, then design an annotation that leverages those tools to provide insights into the current state of students' knowledge.

Specific Requests

Several participants liked having example reports and wanted more. We will be publishing a more in-depth analysis of student writing that includes a supplemental data store of ~4400 lab reports that were graded using the bins scoring method. Contact the author to request early access. Others wanted examples of the fully commented reports; these have been added to Supplement 3.

Several participants pointed out that SEM was hard to understand without seeing the entire plan in overview or knowing what is in our Writing Resource Guide. This is a challenge for our first-semester GTAs too. We provide them with a longer outline showing how our system works, and more detailed instructions on making comments on reports (including examples of reusable comments). This document is provided as Supplement 5.

In response to user requests we created an open-source version of the BioCore Writing Resource Guide that users can modify to fit local needs. This will be presented in an ABLE workshop in 2022. An open-source version of the web form for automated pre-submission checking of reports prior to grading will be released in June 2022; contact the author for details.

Finally, participants pointed out that it was unclear when/how reports are separated from the 3 provisional groups to the 4 final scoring bins. The text has been revised to make that clearer.

Cited References

Nilson, LB. 2014. Specifications grading: restoring rigor, motivating students, and saving faculty time. Stylus Press, Sterling VA. 184pp,

Acknowledgments

Thank you to Dr. Sabrina Setaro and Ms. Shannon Mallison who helped design, develop, and implement the Six Elements Model for Writing. Thanks as well to all of the GTAs and undergraduates who provided insights and suggestions for improving this program. The program and data presented were developed with support from NSF grant #1712423 to Dr. Johnson.

About the Author

A. Daniel (Dan) Johnson is a Teaching Professor and the former Core Curriculum Coordinator in Biology at Wake Forest University in Winston-Salem, NC. Since 1998 he has been developing inquiry lab activities and lab support resources across the range of biology. He also is part of the BioTAP Network for improving GTA professional development.
.

Appendix A

Slides for Exercise 2

Part 1: Background and Goals

Overall Goals of This Activity

- Help students implement what they read in Resource Guide
- Start conversation b/w GTAs, students about scientific writing
- Introduce language of science communication
- Provide GTAs with data to guide debrief
  - Give names to the errors we see most often
  - Surface group vs. individual knowledge gaps, misconceptions

GTA Goals When Introducing Activity in Week 1

- Reinforce this is a "diagnostic activity".
- Do in-class portion THIS week
  - Be sure ALL understand instructions before leaving
- Remind them:
  - They CAN refer back to Resource Guide (not closed book activity)
  - Need to download PDF of their answers, upload to LMS assignment, AND bring a copy next week

GTA Goals for Debrief at End of Week 1

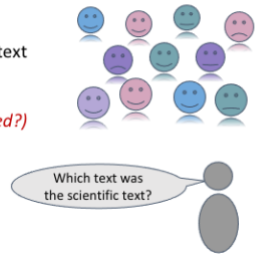
Spend 20 minutes or less

Help for YOU as instructor:
- Reinforce goals of exercise
- Be sure all students know:
  - How to interpret questions
  - How to generalize questions

Help STUDENTS as writers:
- Exercise vocabulary of scientific writing
- Start to evaluate others' writing systematically

Debriefing at End of Week 1

Keep it short:
- Which text was a scientific text and **why**?
  (Do students recognize language, structure expected?)
- What is the purpose of the citations?
  (Do students understand role of sources?)



GTA Goals for Debrief in Week 2

Plan for 20-30 minutes

Help for YOU as instructor:
- Provide diagnostic / predictive data
- ID students' biases, misconceptions
- Uncover most common errors as a class
- Start evidence-driven conversation w/students

Help STUDENTS as writers:
- Get low-stakes "active reading" experience
- Start to evaluate **own** writing systematically

Using Data From GTA Annotation Reports

1. Review Annotation Report (or submissions) for your sections
2. ID what you think are main problems students have
3. Write **reflective questions** to help students surface, solve these problems
   - Prioritize BASIC criteria, then technical/writing
   - Address biggest, most common errors FIRST
4. Document, share your students' main problems, your questions with other GTAs, Coordinators

Sample Questions – Introduction Issues

1. **Hypothesis**:
   - Why is it a research question and not a hypothesis in this case?
   - How might you re-phrase the research question to turn it into a hypothesis?
2. **Citations**: What is the difference b/w a citation, quote?
3. **Structure**: What are minimum requirements for Introduction?

Sample Questions – Methods Issues

1. **Control**: What is purpose of control in experiments?
2. **Variables**: What is difference between dependent, independent variable?
3. **Statistics**: Why use statistics in a scientific experiment?
4. **Structure**: In what tense is Methods section written? Why?

Sample Questions – Results Issues

1. **Statistics**: What statistical information do you put in Methods vs Results?
2. **Structure**:
   - What is the most important goal of the Results section?
   - What should NOT be in Results?
   - What internal references should be in the Results?

Sample Questions – Discussion Issues

1. **Structure**: What is the boundary line between the Results and the Discussion?
2. **Interpretation**: When you interpret something about your results, what should you also do shortly after?
   - (Show examples from literature of interpretation followed by citations as demonstration)

Sample Questions – Issues in Abstract

1. **Citation**: Why should there be no citations in the Abstract?
2. **Structure**:
   - How is Abstract different from other parts of report?
   - Why should you write the Abstract LAST?

Sample Questions - Titles

- Pick two strong and two weak titles from OTHER section's data
- Ask students to ID strong vs. weak and **why**.

Part 2: Student Response Data

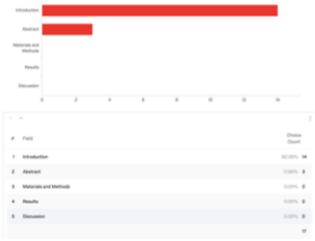
Practice For Week 2: What Can You Learn From The GTA Summary Report?

- The next 11 slides show summary data collected from ONE lab section.
- Write down:
  - Primary problems, misconceptions
  - 3-4 potential questions you might ask

What Can You Pull Out of This GTA Report?

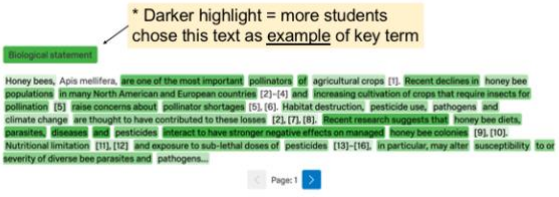
Original question: What are the main features of the ***Introduction***?

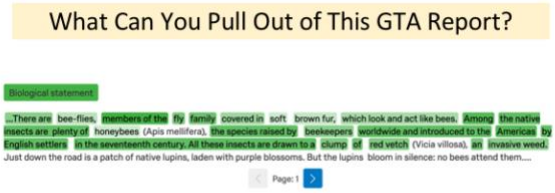
What Can You Pull Out of This GTA Report?

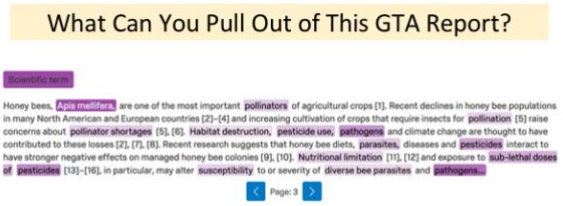


Original question: in what part of a report does this section belong? (*Correct answer is Introduction*)

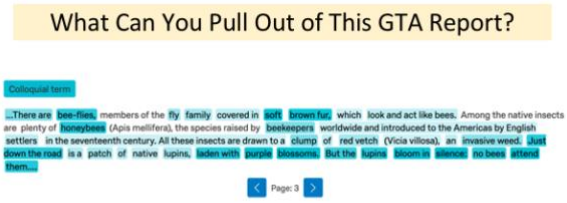
What Can You Pull Out of GTA Summary Report?




What Can You Pull Out of This GTA Report?
Biological statement
...There are bee-flies, members of the fly family covered in soft brown fur, which look and act like bees. Among the native insects are plenty of honeybees (Apis mellifera), the species raised by beekeepers worldwide and introduced to the Americas by English settlers in the seventeenth century. All these insects are drawn to a clump of red vetch (Vicia villosa), an invasive weed. Just down the road is a patch of native lupins, laden with purple blossoms. But the lupins bloom in silence: no bees attend them....
Page: 1


What Can You Pull Out of This GTA Report?
Scientific term
Honey bees, Apis mellifera, are one of the most important pollinators of agricultural crops [1]. Recent declines in honey bee populations in many North American and European countries [2]–[4] and increasing cultivation of crops that require insects for pollination [5] raise concerns about pollinator shortages [5], [6]. Habitat destruction, pesticide use, pathogens and climate change are thought to have contributed to these losses [2], [7], [8]. Recent research suggests that honey bee diets, parasites, diseases and pesticides interact to have stronger negative effects on managed honey bee colonies [9], [10]. Nutritional limitation [11], [12] and exposure to sub-lethal doses of pesticides [13]–[16], in particular, may alter susceptibility to or severity of diverse bee parasites and pathogens...
Page: 3

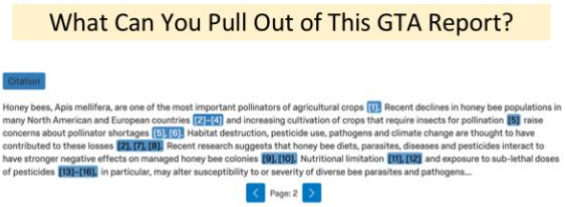

What Can You Pull Out of This GTA Report?
Scientific term
...There are bee-flies, members of the fly family covered in soft brown fur, which look and act like bees. Among the native insects are plenty of honeybees (Apis mellifera), the species raised by beekeepers worldwide and introduced to the Americas by English settlers in the seventeenth century. All these insects are drawn to a clump of red vetch (Vicia villosa), an invasive weed. Just down the road is a patch of native lupins, laden with purple blossoms. But the lupins bloom in silence: no bees attend them....
Page: 2


What Can You Pull Out of This GTA Report?
Colloquial term
...There are bee-flies, members of the fly family covered in soft brown fur, which look and act like bees. Among the native insects are plenty of honeybees (Apis mellifera), the species raised by beekeepers worldwide and introduced to the Americas by English settlers in the seventeenth century. All these insects are drawn to a clump of red vetch (Vicia villosa), an invasive weed. Just down the road is a patch of native lupins, laden with purple blossoms. But the lupins bloom in silence: no bees attend them....
Page: 3


What Can You Pull Out of This GTA Report?
Research question
...This study addresses two important questions. 1) What types of pesticides might bees be exposed to in major crops? While multiple studies have characterized the pesticide profile of various materials inside a honey bee nest [27], [28], few have looked at the pollen being brought back to the nest. 2) How do field-relevant pesticides blends affect bees' susceptibility to infection by the Nosema parasite?...
Page: 3


What Can You Pull Out of This GTA Report?
Hypothesis
...This study addresses two important questions. 1) What types of pesticides might bees be exposed to in major crops? While multiple studies have characterized the pesticide profile of various materials inside a honey bee nest [27], [28], few have looked at the pollen being brought back to the nest. 2) How do field-relevant pesticides blends affect bees' susceptibility to infection by the Nosema parasite?...
Page: 1


What Can You Pull Out of This GTA Report?
Citation
...This study addresses two important questions. 1) What types of pesticides might bees be exposed to in major crops? While multiple studies have characterized the pesticide profile of various materials inside a honey bee nest [27], [28], few have looked at the pollen being brought back to the nest. 2) How do field-relevant pesticides blends affect bees' susceptibility to infection by the Nosema parasite?...
Page: 2


What Can You Pull Out of This GTA Report?
Citation
Honey bees, Apis mellifera, are one of the most important pollinators of agricultural crops [1]. Recent declines in honey bee populations in many North American and European countries [2]–[4] and increasing cultivation of crops that require insects for pollination [5] raise concerns about pollinator shortages [5], [6]. Habitat destruction, pesticide use, pathogens and climate change are thought to have contributed to these losses [2], [7], [8]. Recent research suggests that honey bee diets, parasites, diseases and pesticides interact to have stronger negative effects on managed honey bee colonies [9], [10]. Nutritional limitation [11], [12] and exposure to sub-lethal doses of pesticides [13]–[16], in particular, may alter susceptibility to or severity of diverse bee parasites and pathogens...
Page: 2

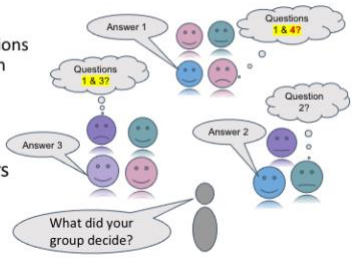
Part 3: Debrief Strategies

WEEK 2: How to Debrief Students

- These are time-efficient ways to explore annotation homework with students
- Which debrief method are you likely to use? Why?

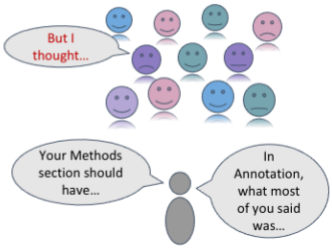
Wk. 2 Debriefing Option 1: Think-Pair-Share

- Split into groups:
  - Give each 2-3 questions
  - Each group works on different questions
  - Answer in group
  - Report back to class – 2 min./qu.
- GTA corrects, answers lingering questions

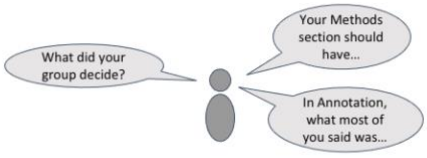


Debriefing Option 2: Intermix General w/Details

- General o'view of report format
- Add findings from exercise as appropriate
- GTA answers lingering questions

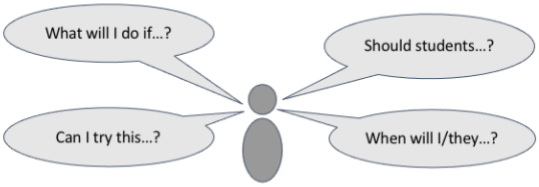


Group Discussion by GTAs



- Which debrief method are you likely to use? Why?

Your Questions?



Mission, Review Process & Disclaimer

The Association for Biology Laboratory Education (ABLE) was founded in 1979 to promote information exchange among university and college educators actively concerned with teaching biology in a laboratory setting. The focus of ABLE is to improve the undergraduate biology laboratory experience by promoting the development and dissemination of interesting, innovative, and reliable laboratory exercises. For more information about ABLE, please visit http://www.ableweb.org/.

Advances in Biology Laboratory Education is the peer-reviewed publication of the conference of the Association for Biology Laboratory Education. Published articles and extended abstracts are evaluated and selected by a committee prior to presentation at the conference, peer-reviewed by participants at the conference, and edited by members of the ABLE Editorial Board. Published abstracts are evaluated and selected by a committee prior to presentation at the conference.

Citing This Article

A. Daniel Johnson 2022. A Six Elements Method for Teaching Scientific Writing at Scale. Article 10 In: Boone E and Thuecks S, eds. *Advances in biology laboratory education*. Volume 42. Publication of the 42nd Conference of the Association for Biology Laboratory Education (ABLE). https://doi.org/10.37590/able.v42.art10