

Generating Biological Hypotheses in a “Course-based Undergraduate Research Experience” (CURE) Course

Priyanjali Ghosh

Department of Neurology, University of Massachusetts-Amherst

Abstract

Hypotheses are often incorrectly defined as educated guesses or if/then statements. Further confusion arises from the fact that not all disciplines define hypotheses similarly. The difficulty in distinguishing between statistical and biological hypotheses has led to the erroneous substitution of a hypothesis for a prediction. These misconceptions are pervasive amongst scientists from all levels of training. The inquiry-based structure of CUREs creates a conducive space for clarifying these misconceptions. However, lectures in a first attempt CURE course were not sufficient in clearing this confusion. Studies show that highly structured courses with active learning techniques successfully increase students' comprehension, this lesson complements the in-class lecture with three activities – designing hypotheses and predictions for real-world observations, ii) workshops on generating hypotheses and predictions from published article excerpts, and iii) formulating hypotheses for students' in-class experiments. Students participated in these activities in pairs and received Socratic feedback on how to improve their responses. Students' responses from both the lecture-only semester and the lecture+active learning semester were evaluated by analyzing the accuracy of hypotheses produced by students, i.e., i) Is there a correct distinction between the hypothesis and predictions? ii) Are students testing the correct phenomenon? and iii) Are students including a biological mechanism? Overall, for each of these three criteria, students from the lecture+active learning cohort out-performed students from the lecture-only cohort. In conclusion, this lesson supports the established notion that structured active learning increases students' comprehension and achievement.

Citation: Ghosh P. 2021. Generating biological hypotheses in a “Course-based Undergraduate Research Experience”(CURE) course. *CourseSource*. <https://doi.org/10.24918/cs.2021.3>

Editor: Tracie Marcella Addy, Lafayette College

Received: 7/14/2020; **Accepted:** 11/30/2020; **Published:** 2/10/2021

Copyright: © 2021 Ghosh. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

Conflict of Interest and Funding Statement: The author does not have a financial, personal, or professional conflict of interest related to this work.

Supporting Materials: Supporting Files S1. Generating biological hypotheses – Exit card activity presentation slide; S2. Generating biological hypotheses – Scientific method lecture slides; S3. Generating biological hypotheses – Workshop 1 worksheet; S4. Generating biological hypotheses – Assignment 1 instructions; S5. Generating biological hypotheses – Assignment1 feedback slide; S6. Generating biological hypotheses – Workshop 2 worksheet; S7. Generating biological hypotheses – Assignment 1 revision instructions; and S8. Generating biological hypotheses – Hypothesis accuracy scoring data.

*Correspondence to: ghosh20p@gmail.com

Learning Goal(s)

Students will:

- Learn about a research topic by reviewing prior research.
- Identify unexplored research avenues and formulate novel experimental questions.
- Understand and practice the initial steps of the Scientific Method (i.e., asking questions, formulating hypotheses, and predicting outcomes).

Learning Objective(s)

Students will be able to:

- Create foundational knowledge about their research topic by reading review articles and primary research papers.
- Use this knowledge to formulate novel research questions.
- Formulate a hypothesis statement that is testable, falsifiable, and rooted in a biological mechanism.
- Learn differences between a hypothesis statement, a research questions and experimental predictions.
- Evaluate survival data to determine virulence properties of a given microbe.

INTRODUCTION

An essential step in the cyclical scientific method (1) is the formulation of hypotheses and predictions. Frequently, hypotheses are incorrectly defined as *educated guesses* or *if/then* statements (2,3). This confusion arises from the fact that not all disciplines define hypotheses in the same way. Statistical hypotheses are formal claims about the nature of a population or an existing pattern. For example, “*Male chickens have a different average foot size than females*” – it states a fact but does not provide a “cause” for why those numbers are different

(4-7). In contrast, a biological hypothesis would provide an explanation like “*Sexual selection by females has caused male chickens to evolve bigger feet than females*” (7). Therefore, biological or explanatory hypotheses are testable and falsifiable mechanistic explanations of existing phenomena. In other words, a biological hypothesis should account for facts relating to a specific problem, and it should lead to predictions that can be tested (3,8). It is the inability to distinguish between the two (i.e., statistical vs. biological) that leads to the erroneous substitution of a hypothesis for a prediction. A recent survey showed that this confusion was systemic in the science

community, where 42% of surveyed science textbooks (n = 54) defined a hypothesis as a prediction (3). Informational interviews conducted in 2006 revealed that 81.2% (n = 1176) of high school students defined a hypothesis in the form of a predictive statement, such as: *If I do X, then Y will happen* (3). As a result, it comes as no surprise that students carry this confusion well past high school. In college, it is hard to clear up these misconceptions in the standard lecture and lab courses. For example, in traditional undergraduate cookbook labs, students verify existing data; there is no scope for critical thinking and inquiry. Thus, when students are required to write hypotheses for these canned labs, they usually write predictions for their experiments – further perpetuating the systemic misunderstandings about hypotheses and predictions (8-10).

With the shift towards inquiry-based pedagogy, instructors now have the opportunity to clarify these misconceptions about hypotheses and predictions. Specifically, CUREs (Course-based Undergraduate Research Experience) create a unique learning environment where students are encouraged to participate in every step of the scientific method. CUREs promote intellectual autonomy; undergraduates learn to make observations, problem-solve, troubleshoot, and analyze data in an iterative manner. Thus a CURE is the perfect backdrop wherein misconceptions about the hypothetico-predictive nature of science can be clarified (10-12). Nevertheless, even in the context of a CURE, merely explaining the concept of a biological hypothesis is not sufficient to improve student responses. The thought process required to construct a biological hypothesis is challenging and requires practice. Students need specific, well-designed activities to learn how to write an accurate hypothesis (8).

Studies have shown that highly structured courses that incorporate active learning methods have successfully increased students’ interest, comprehension, academic success, and retention in the sciences (13-15). This lesson has been designed to help students understand the concept of and develop the skill of writing biological hypotheses. In McPherson’s published learning exercise (9), the author lectures on the difference between hypotheses and predictions and assigns 5-10 articles (recent, primary literature articles) for students to read, and asks them to think about the following questions:

1. Did the paper propose hypotheses or predictions?
2. Were the hypotheses or predictions clearly stated?
3. Were the hypotheses or predictions testable?
4. Were the hypotheses or predictions tested?
5. On a scale of one to five, how comfortable are you with the paper’s conclusions?

Student responses are evaluated based on how well they applied the lecture material in answering these questions. Student answers serve as the basis of in-class discussions where the author spends time clarifying existing confusions and misconceptions.

What makes this lesson different is that it is taught in the context of a CURE, and it focuses only on the formulation of hypotheses. Like McPherson’s paper (9), it begins with a lecture that explains the concepts and provides the students

with a scaffold to generate hypotheses. Using this scaffold, students participate in group discussions to produce real-world examples to highlight the difference between a hypothesis and a prediction. Next, students participate in workshops [Supporting File S3. Generating biological hypotheses – Workshop 1 worksheet, Supporting File S6. Generating biological hypotheses – Workshop 2 worksheet] in which they are given excerpts from the introduction section of primary literature articles and are asked to consider the following:

1. What observations do the authors make?
2. What biological mechanisms do the authors discuss?
3. Do they explicitly mention hypotheses and predictions?

Since most published primary literature articles rarely state a hypothesis explicitly, this activity gives students the chance to use the scaffold and contents of the introduction (i.e., answers to questions 1 and 2) to generate hypotheses and predictions. Finally, students write hypotheses and predictions (relevant to the CURE course content) that they subsequently test via experiments during the semester. The lesson gives students practice in formulating hypotheses and predictions and creates a sense of ownership over their work. Students have the opportunity to test whether or not their predictions were correct and show how their results can support or refute their hypotheses; this is an entirely different approach from the McPherson learning exercise, where students evaluate published experiments. Thus, this highly structured, active learning lesson provides instructors with the opportunity to guide students into learning and understanding the concept of a biological hypothesis and prediction.

Intended Audience

This lesson was designed for Junior/Seniors in the Biology Major at a large research university. However, it can certainly be adapted for any Biology course if the instructor provides the appropriate context and background information. It would undoubtedly be best to do this in an inquiry-based lab course where the students can test their hypotheses—it gives them a sense of ownership.

Required Learning Time

This lesson was designed to be taught during a four-hour lab that meets once a week. The entire lesson spans over 5-6 weeks, including the week in which students work on the lesson on their own (i.e., homework week). If adapting for another course/context, the instructor can cover the initial lecture and workshops in three hours (could be done over two meetings), followed by another hour (during a different meeting) for feedback and the second workshop. Details about the lesson timeline is outlined in Table 1.

Prerequisite Student Knowledge

The most important prerequisite is a working knowledge of how to read a scientific paper and a general understanding of the scientific method. Since the lesson was designed for Juniors/Seniors, it is likely that the students have encountered these concepts in prior STEM courses. However, to keep the course inclusive, the following resources are provided:

- Reading Primary Literature: <https://www.sciencemag.org/careers/2016/03/how-seriously-read-scientific-paper>

- The Scientific Method: <https://www.khanacademy.org/science/high-school-biology/hs-biology-foundations/hs-biology-and-the-scientific-method/v/the-scientific-method>

Prerequisite Teacher Knowledge

Regardless of whether this lesson is being taught in the context of a CURE or only to Juniors/Seniors, instructors should have a thorough understanding of hypotheses and predictions. Since there is a systemic confusion about the topic, instructors (faculty/staff/TAs) should familiarize themselves with the concept of a biological hypothesis. In addition to the sources mentioned in the introduction section of this lesson, the instructors can also refer to this slide deck from P.K Strode <https://science.jburroughs.org/mbahe/APBio/APBioHandouts/00ResearchHypothesisNABT.pdf>

SCIENTIFIC TEACHING THEMES

Active Learning

For this CURE, students are randomly assigned as pairs for the whole course. Thus, the group activities in which the students have to come up with real-world examples of hypotheses and predictions and the workshops are all a type of think-pair-share, active learning activity.

Assessment

Students routinely participate in anonymous exit-card activities; they are asked to define hypotheses and predictions on small notecards and submit them before they leave the course, providing a low stakes formative assessment, that helps tailor the subsequent lecture. The primary assessment for this lesson is generated when students (as a group) write hypotheses and predictions pertinent to their experiment. This assignment [Supporting File S4. Generating biological hypotheses – Assignment 1 instructions, Supporting File S7. Generating biological hypotheses – Assignment 1 revision instructions] assesses the students' ability to evaluate existing scientific information critically and their ability to use that information to generate accurate and relevant hypotheses and predictions. Each hypothesis (written by the group) is evaluated against the following three criteria:

- Criterion #1: Is there a correct distinction between the hypothesis and predictions?
- Criterion #2: Are students testing the correct phenomenon?
- Criterion #3: Are students including a biological mechanism?

A hypothesis that meets all three criteria is considered to be 100% accurate.

Inclusive Teaching

Listed here are the action items for creating an inclusive classroom and lesson:

1. Students create their own name cards, in which they can include their pronouns. Since there is a lot of discussion in the lesson, it is necessary to know how to address students. All of this information is not always populated in the class roster, so it is important to ask!
2. Although the lesson is designed for Juniors/Seniors, it

is advisable to review the fundamental concepts of The Scientific Method and How to read primary literature to ensure that all students are on the same page.

3. All lesson-related materials are posted on the course website, and handouts are provided in class. The classroom is also equipped with computers and printers. This provides students easy access to learning material.

LESSON PLAN

Class Period #1: Exit-Card Activity

At the end of the class session preceding the lecture/discussion on *Formulating hypotheses and predictions*, I do the exit-card activity [Supporting File S1. Generating biological hypotheses – Exit card activity presentation slide]. I ask each student to anonymously write down the definition of a hypothesis and a prediction on note cards and turn it in before they leave. I encourage every student to participate by being transparent; I let them know that it is ok not to know the answer and that I will use their answers to tailor the class lecture and discussion.

Class Period #2

Item #1: Lecture on The Scientific Method and Formulating hypotheses and predictions.

The day of the lecture, I begin by giving students a brief overview of *The Scientific Method* [Supporting File S2. Generating biological hypotheses – Scientific method lecture slides]. I highlight the importance and purpose behind each step and show the students how they will have the opportunity to participate in every step of *The Scientific Method* during the course. After this brief overview, I show students the tabulated answers from the exit-card activity [Supporting File S2. Generating biological hypotheses – Scientific method lecture slides], creating a starting point for lecture and discussion on hypotheses and predictions [Supporting File S2. Generating biological hypotheses – Scientific method lecture slides]. In this lecture, I explain the difference between statistical and biological hypotheses, provide them with the definition of biological hypotheses and predictions, provide them with real-world examples, and end with a scaffold that will help students write hypotheses and predictions.

Item #2: Class discussion on formulating hypotheses and predictions for real-world observations.

After the lecture, each pair is tasked with writing a real-world i) observation, ii) question, iii) hypothesis, iv) predictions [Supporting File S2. Generating biological hypotheses – Scientific method lecture slides]. The lab partners work on this for 5-10 minutes, and then each group reports back to class and presents their real-world examples. During the discussion, I check to see if each hypothesis has the key components, i.e., the hypothesis is testable, falsifiable, and must include a mechanism, and that the predictions are related to the stated hypothesis.

Item#3: Workshop #1.

In this active learning exercise (Workshop #1 [Supporting File S3. Generating biological hypotheses – Workshop 1 worksheet]), the groups read short excerpts from the introduction section of published primary literature. The groups

are instructed to use the information in the excerpts to identify the observations and questions that the authors are making/asking in the introduction. With this, the groups are asked to formulate a hypothesis and prediction using the scaffold that was given to them during the lecture. For this workshop, it is helpful to use a variety of examples; some excerpts mention a hypothesis, whereas the others do not explicitly state it – they simply allude to it. This workshop pushes the students to use the limited content to find i) what the authors are testing, ii) what is a potential mechanism behind their tests, and iii) whether the predictions support/falsify their hypothesis. This workshop also highlights that professional scientists may not explicitly state a hypothesis, but they nonetheless provide all the key components of a hypothesis in their introduction.

Workshop #1 has four excerpts. I give the groups 15 minutes to work on the first excerpt. During this time, I circulate in the classroom and help guide discussions or clarify any issues that may arise. I also encourage students to research scientific terms that they may not understand in the excerpt. After the 15 minutes, we regroup as a class and discuss the hypotheses and predictions. Once again, the class discussion focuses on the key components of a hypothesis and whether it aligns with the predictions. After the class discussion, the groups move on to the next excerpt, and we go through the same process again.

Homework (group activity) – Assignment #1: Observations, Hypotheses, and Prediction.

After Workshop #1, I assign students a recent review article relevant to the course. These review articles can be extensive, so I pick relevant excerpts from the article and assign them to the students. For example, for this lab, I picked the article titled *Thyroid hormone regulation of neural stem cell fate: From development to aging* (16). Students are given two weeks to read the review article and write an assignment titled: Assignment#1: Observations, Hypotheses, and Prediction [Supporting File S4. Generating biological hypotheses – Assignment 1 instructions]. As the title of the assignment suggests, students write observations based on their reading and use that information to write a hypothesis and predictions that are aligned with their experiment.

Class period #3

Item #1: Feedback on homework and lecture review.

Once the students submit the assignment, I go over their answers, and provide them with some general feedback in the following class period. I always show my students data so that there is buy-in [Supporting File S5. Generating biological hypotheses – Assignment1 feedback slide]. Therefore, during the feedback, instead of providing possible correct answers, I show them a graph (or a table) depicting the successes and errors in their hypothesis/prediction formulations [Supporting File S5. Generating biological hypotheses – Assignment1 feedback slide]. Following the feedback, I quickly recap the contents from the hypothesis/prediction lecture [Supporting File S2. Generating biological hypotheses – Scientific method lecture slides] and then move on to Workshop #2 [Supporting File S6. Generating biological hypotheses – Workshop 2 worksheet].

Item #2: Workshop #2 and group check-ins.

I conduct Workshop #2 in the same way as Workshop

#1; only this time, it goes quicker since only one excerpt is examined. After completing Workshop #2, I hand back the assignments and ask them to review and identify their mistakes. I make it a point to visit each group to check-in with them, clarifying anything they need without giving away a potential answer.

Homework/Final Assessment: Assignment #1 Revision

Finally, the groups are asked to redo assignment #1 [Supporting File S7. Generating biological hypotheses – Assignment 1 revision instructions]; they must write an introduction section for their intended experiment followed by a revised version of their hypotheses and predictions. Students will submit their revisions the following week, and I correct and assign grades to this version of the assignment. After the grading is complete, I return the assignments to the students and provide them with the potential correct answers (the content’s scope allows for multiple correct hypotheses).

TEACHING DISCUSSION

Observations and Conclusions

I have been teaching this CURE course for the past three semesters, and in my first semester (Spring 2019), I noticed that students did not have a clear idea about hypotheses and predictions, and they were often interchanging the two. In fact, throughout the past three semesters, I noticed that students from each semester often considered a hypothesis as an *If, then* (which is prediction), or an *educated guess*. This observation was evident from the responses in the exit-card activity [Supporting File S1. Generating biological hypotheses – Exit card activity presentation slide], and is in direct alignment with previously published observations (2,3). Some example responses from students include:

- *A hypothesis is like “If I do this, then that will happen.”*
- *A hypothesis is a thought out “if” and “then” statement*
- *A hypothesis is an idea of if something is true, then something will happen*
- *A hypothesis is a statement/proposal that is formulated based on a prediction*
- *An educated guess on a known topic*
- *An educated guess as to the outcome of an experiment. Doesn’t need to be right*

To help clarify this, I spent some lecture-time in the Fall of 2019 to help them understand the concept [Supporting File S2. Generating biological hypotheses – Scientific method lecture slides]. However, a simple lecture was not enough to clear up misconceptions, and that observation led to the genesis of this lesson. As a result of this course modification, I can now compare the responses to assignment #1 [Supporting File S4. Generating biological hypotheses – Assignment 1 instructions] between the lecture-only cohort from Fall 2019 and the lecture+active learning cohort in the Spring 2020 (*henceforth, will only be referred to as Fall and Spring semester*). For additional information about how the two semesters differ from one another, please refer to Figure 1. Students from both semesters used the same review article (16) to generate hypotheses and predictions for assignment #1 [Supporting File S4. Generating biological hypotheses – Assignment 1 instructions] (Figure 1).

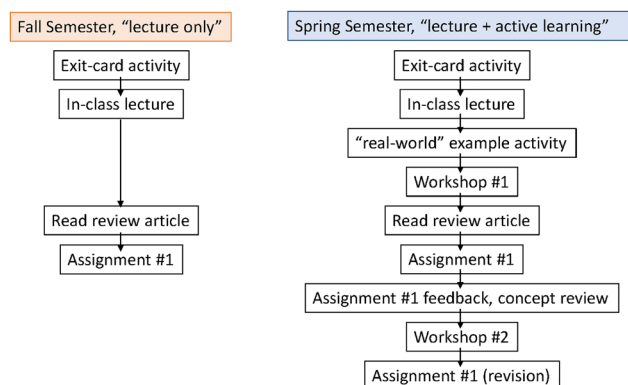


Figure 1. Differences in teaching approach. This schematic represents the difference in teaching between the Fall (lecture only) and the Spring (lecture+active learning) semesters. In the Fall, students received a lecture and proceeded to read the review article and then assignment#1. In the Spring, the students participated in two active learning activities before reading the review article and writing the assignment. The received additional practice and feedback after the first submission and were given a chance to revise the assignment.

The hypotheses were evaluated for three criteria:

- Criterion #1: Is there a correct distinction between the hypothesis and predictions?
- Criterion #2: Are students testing the correct phenomenon?
- Criterion #3: Are students including a biological mechanism?

If a hypothesis satisfied all three criteria, it was considered to be 100% accurate and was given an accuracy score of 3. For every missing criterion, 1 point was deducted. Interestingly, in the Spring semester, some hypotheses alluded to a mechanism, but the responses were not specific (Table 2); these responses received partial credit (score of 0.5) for that criterion. However, it must be noted that the partially correct response was

considered as a 0 when calculating the accuracy per criterion (Figures 2B, 3C, Table 3, and [Supporting File S8. Generating biological hypotheses – Hypothesis accuracy scoring data]). Example student responses and their corresponding accuracy scores are shown in Table 2. Additionally, it is worth noting that while students wrote both hypotheses and predictions, only the accuracy of the hypotheses was measured. The reason for this is that if a hypothesis was generated accurately, the students were automatically phrasing their predictions in the *If... then* format. Having students write out both statements helped me determine the extent of their confusion.

Overall, the trend indicates that the students from the Spring semester performed better than those from the Fall. In the Spring, 6 groups (out of 16) received an accuracy score of 3 in comparison to 3 groups (out of 16) in the Fall. Moreover, 4 groups in the Spring received an accuracy score of 0 in comparison to the 8 groups in the Fall (Figure 2A, [Supporting File S8. Generating biological hypotheses – Hypothesis accuracy scoring data.]). If we look at the accuracy in response to each criterion, we will see that once again, the students in the Spring semester performed better than those in the Fall. The most striking improvement was seen for criterion #1; 68.8% (11/16 groups) of the class successfully structured a hypothesis and did not confuse it as an *If... then* predictive statement (Figure 2B, [Supporting File S9. Generating biological hypotheses – Hypothesis accuracy scoring data.]). This a huge improvement from the Fall semester, when only 25% (4/16 groups) of the class were able to construct a hypothesis correctly (Figure 2B, [Supporting File S8. Generating biological hypotheses – Hypothesis accuracy scoring data.]). Noticeably, the criterion that the students struggled with the most was including a biological mechanism in their hypothesis (Figure 2B, [Supporting File S8. Generating biological hypotheses – Hypothesis accuracy scoring data.]). This can be attributed to i) they identified an inaccurate or unrelated mechanism, ii) they had difficulty understanding the dense content in the review article, and iii) they had a pre-existing misconception about the definition of a hypothesis.

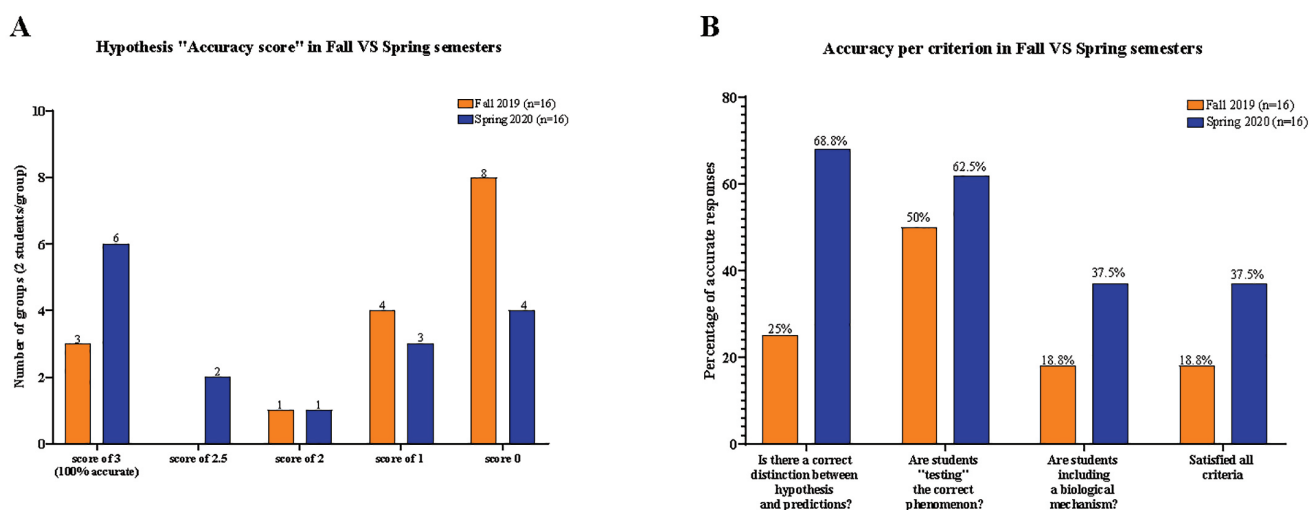


Figure 2. Hypothesis accuracy comparison between the Fall and Spring semester. (A) Depicts the comparison in accuracy scores between the Fall (lecture only; orange bars) and Spring (lecture+active learning; navy bars) semesters. There was a 50% increase in the number of groups with a score of 3, and a 50% decrease in the number of groups with a score of 0, in the Spring. (B) Depicts the accuracy percentage for each hypothesis assessment criteria. Groups in the Spring semester (navy bars) performed better in every criterion, with the most striking improvement visible for the first criterion, where 68.8% of the responses showed a distinction between hypotheses and predictions.

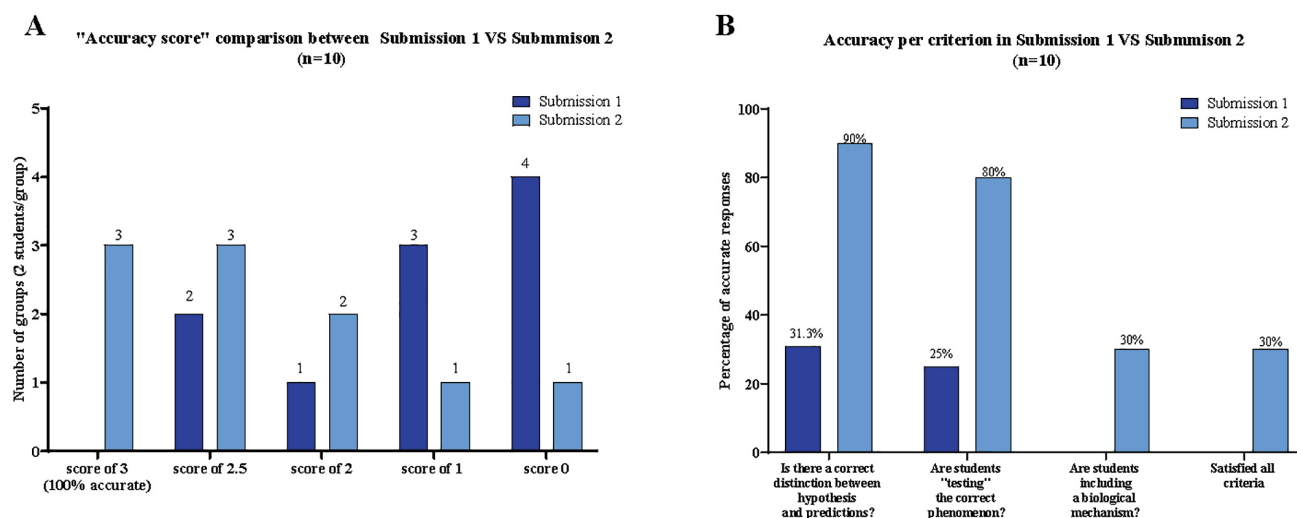


Figure 3. Hypothesis accuracy comparison between the submission 1 and submission 2. (A) Depicts the comparison in the accuracy scores between the first submission (navy bars) and the second submission (light blue) for assignment #1 in the Spring (lecture+active learning) semester. (B) Depicts the accuracy percentage for each hypothesis assessment criteria between the first submission (navy) and the second submission (light blue) for assignment #1 in the Spring (lecture+active learning) semester. Both panels show that additional practice and feedback improved student performance, and once again, the most improvement was seen for the first criterion (first light blue bar in panel B), showing that active learning helped students understand the difference between a hypothesis and a prediction.

Unlike the Fall Semester, students in the Spring semester had the opportunity to revise their hypotheses. This revision happened after they participated in Workshop #2, the goal of which was to remind students about the key concepts of a hypothesis and to give them additional practice. This time, the analysis was performed on the 10 groups that did not get an accuracy score of 3 in their first submission. For the second submission, 3 out of 10 groups (Figure 2B; groups 8, 11, 14 in Table 3) were able to write an accurate hypothesis. Out of the seven remaining groups, 4 of them showed moderate improvement for at least one criterion (Table 3). For example, group 9 was able to include an incorrect mechanism in their hypothesis (partial credit), thereby improving their score from a 2 to a 2.5. Similarly, group 12 clarified the confusion between a hypothesis and a prediction although they included an incorrect mechanism (instead of no mechanism, therefore partial credit), thereby improving their score from 1 to 2.5 (Table 3). When looking at the accuracy per criteria, the noticeable improvement was for criterion #1 (Figure 3B), showing that an additional round of practice/discussion reduced the confusion between a hypothesis and prediction statement. While only 3 out 10 groups (i.e., 30%, Figure 3B) wrote a correct mechanism, it must be noted that 4 out of the remaining 7 groups included a mechanism, albeit incorrect (Table 3). The net result is that 87.5% (14 out of 16 groups) of the Spring semester cohort, understood the fact that a mechanism must be included in a biological hypothesis. In conclusion, this lesson supports the established notion that structured active learning increases students' comprehension and achievement (13,17,18).

Suggested Modifications

While the sample size was small in this study, it was clear from the observations that students performed better and had a clearer understanding of hypotheses and predictions after the pedagogical intervention. This lesson can be adapted for any class level and any class size. However, it is most effective if the students write hypotheses and predictions that they will test in a laboratory setting.

Here are some suggestions for modifying the lesson:

1. TA support in a large course: It will be helpful for the instructor to have TA's who can circulate in the classroom and help students during the real-world discussions and the two workshops
2. TA selection and education: Regardless of the class size, if the instructor has the help of TAs, it is critical that the TAs and the instructor/s is/are in agreement about the concept. Since there is systemic confusion about writing hypotheses and predictions (refer to discussion in the introduction), it will be beneficial for the entire teaching team to set standards and expectations before the course. It is advisable to hold a pre-lesson workshop in which the teaching team can go through the discussion and workshop content.
3. Selecting the reading material: For this CURE course, students were investigating how thyroid hormones played a role in regulating neural stem cell proliferation. Thus, for this lesson, students read the most recent review article to make observations from the existing knowledge in the field and construct their hypotheses accordingly. The instructor should identify the most relevant review article that will help students understand the existing knowledge in the field. Depending on the topic of the review article, the instructor should consider teaching some key concepts that will help students read the review article. For example, prior to this particular lesson, students in this CURE course were taught the concepts of the HPT axis, neural stem proliferation, and differentiation.

SUPPORTING MATERIALS

- Supporting File S1. Generating biological hypotheses – Exit card activity presentation slide
- Supporting File S2. Generating biological hypotheses – Scientific method lecture slides
- Supporting File S3. Generating biological hypotheses – Workshop 1 worksheet
- Supporting File S4. Generating biological hypotheses –

- Assignment 1 instructions
- Supporting File S5. Generating biological hypotheses – Assignment1 feedback slide
- Supporting File S6. Generating biological hypotheses – Workshop 2 worksheet
- Supporting File S7. Generating biological hypotheses – Assignment 1 revision instructions
- Supporting File S8. Generating biological hypotheses – Hypothesis accuracy scoring data

ACKNOWLEDGMENTS

1. I would like to thank Dr. David Gross for his support and guidance in this project, and for reviewing and providing feedback on this manuscript.
2. American Society for Microbiology for allowing me to participate in the Biology Education Research: Getting Started webinar series. This course helped me learn about designing and interpreting science education research.
3. Thanks to the organizers and participants in the CourseSource Writing Studio (June 2020) for creating a very productive, virtual writing workshop. The experience was pivotal to the creation of this manuscript.
4. UMass-Amherst Office of Research Compliance/ Human Research Protection for approving IRB protocol #1733.
5. UMass-Amherst Biology Department for giving me the opportunity to teach the CURE course.

REFERENCES

1. Wilson KJ, Rigakos B. 2016. Scientific process flowchart assessment (SPFA): A method for evaluating changes in understanding and visualization of the scientific process in a multidisciplinary student population. *CBE Life Sci Educ* 15:1-14. <https://doi.org/10.1187/cbe.15-10-0212>
2. McComas WF. 1996. Ten Myths of Science: Reexamining What We Think We Know About the Nature of Science. *Sch Sci Math* 96:10-16. <https://doi.org/10.1111/j.1949-8594.1996.tb10205.x>
3. Strode PK. 2015. Hypothesis Generation in Biology. *Am Biol Teach* 77:500-506. <https://doi.org/10.1525/abt.2015.77.7.4>
4. Sterner RT. 1998. The Scientific Method: An Instructor's Flow Chart. *Am Biol Teach* 60:374-378. <https://doi.org/10.2307/4450499>
5. Ross SM. 2017. Testing Statistical Hypotheses, p. 381-432. In *Introductory Statistics*. Elsevier. <https://doi.org/10.1016/b978-0-12-804317-2.00009-6>
6. Kitchin J. 1994. 6. Basic statistical inference, p. 155-186. In *Methods in Experimental Physics*. Academic Press. [https://doi.org/10.1016/S0076-695X\(08\)60256-2](https://doi.org/10.1016/S0076-695X(08)60256-2)
7. McDonald JH. 2014. *Handbook of Biological Statistics*, 3rd ed. Sparky House Publishing, Baltimore, Maryland. <http://www.biostathandbook.com/HandbookBioStatThird.pdf>
8. Yip DY. 2007. Biology Students' Understanding of the Concept of Hypothesis. *Teach Sci* 53:23. https://www.researchgate.net/publication/273257151_Biology_students_understanding_of_the_concept_of_hypothesis
9. McPherson GR. 2001. Teaching & Learning the Scientific Method. *Am Biol Teach* 63:242-245. <https://doi.org/10.2307/4451093>
10. Weaver GC, Russell CB, Wink DJ. 2008. Inquiry-based and research-based laboratory pedagogies in undergraduate science. *Nat Chem Biol*. Nature Publishing Group. <https://doi.org/10.1038/nchembio1008-577>
11. Lawson AE. 2004. The nature and development of scientific reasoning: A synthetic view. *Int J Sci Math Educ* 2:307-338. <https://doi.org/10.1007/s10763-004-3224-2>
12. Dolan EL. 2016. Course-based Undergraduate Research Experiences: Current knowledge and future directions. *Natl Res Counc Comm Pap* 1-34. <https://sites.nationalacademies.org/cs/groups/dbassesite/documents/webpa...>
13. Freeman S, Haak D, Wenderoth MP. 2011. Increased course structure improves performance in introductory biology. *CBE Life Sci Educ* 10:175-186. <https://doi.org/10.1187/cbe.10-08-0105>
14. Eddy SL, Hogan KA. 2014. Getting under the hood: How and for whom does increasing course structure work? *CBE Life Sci Educ* 13:453-468. <https://doi.org/10.1187/cbe.14-03-0050>
15. Wilton M, Gonzalez-Niño E, McPartlan P, Turner Z, Christoffersen RE, Rothman JH. 2019. Improving academic performance, belonging, and retention through increasing structure of an introductory biology course. *CBE Life Sci Educ* 18. <https://doi.org/10.1187/cbe.18-08-0155>
16. Gothié JD, Vancamp P, Demeneix B, Remaud S. 2020. Thyroid hormone regulation of neural stem cell fate: From development to ageing. *Acta Physiol*. Blackwell Publishing Ltd. <https://doi.org/10.1111/apha.13316>
17. Freeman S, O'Connor E, Parks JW, Cunningham M, Hurley D, Haak D, Dirks C, Wenderoth MP. 2007. Prescribed active learning increases performance in introductory biology. *CBE Life Sci Educ* 6:132-139. <https://doi.org/10.1187/cbe.06-09-0194>
18. Maskiewicz AC, Griscom HP, Welch NT. 2012. Using Targeted Active-Learning Exercises and Diagnostic Question Clusters to Improve Students' Understanding of Carbon Cycling in Ecosystems. *CBE Life Sci Educ* 11:58-67. <https://doi.org/10.1187/cbe.11-02-0011>

Table 1. Lesson plan teaching timeline.

Activity	Description	Estimated Time	Notes
Preparation for Class			
Prepare reading, workshop and lecture materials, and write assignments	<ol style="list-style-type: none">1. Select review article related to course research topic2. Select at least 4 excerpts for Workshop#1 (does not have to be related to course research topic)3. Select additional (at least one) excerpt for Workshop#24. Prepare lecture slides5. Write assignments6. Purchase note cards for “exit card” activity	2-3 days	It will take time to find articles that will fit the scope of the course and align with the course content and student level. Please refer to Supporting Files for inspiration and guidance.
Class Session 1			
Exit-card activity (Formative Assessment)	Students will answer two questions on a note card	~5 minutes	Questions are in Supporting File S1. Generating biological hypotheses – Exit card activity presentation slide.
Class Session 2			
Item #1: Lecture	Explain the scientific method, go over exit card answers, explain hypothesis and predictions	~20 minutes	Lecture material is available in Supporting File S2. Generating biological hypotheses – Scientific method lecture slides.
Item #2: Class discussion (Think-pair-share; active learning)	Groups will formulate hypotheses and predictions for real-world observations and present their answers to class for discussion	~20 minutes	Lecture material is available in Supporting File S2. Generating biological hypotheses – Scientific method lecture slides.
Item #3: Workshop #1 (Think-pair-share; active learning)	Groups will read excerpts from published introduction sections and use the content to write hypotheses and predictions.	~2 hours	Refer to Supporting File S3. Generating biological hypotheses – Workshop 1 worksheet.
Homework			
Reading	Students will read the selected review article	2 weeks	Refer to Supporting File S4. Generating biological hypotheses – Assignment 1 instructions.
Write Assignment #1	Students will write observations based on their reading and use that information to write a hypothesis and predictions that are in alignment to their experiment.		Refer to Supporting File S5. Generating biological hypotheses – Assignment1 feedback slide
Class Session 3			
Item #1: Feedback on homework and lecture review	Provide general feedback to the students, show them data on successes and errors, review lecture material from class session 1	~20 minutes	Feedback material is available in Supporting File S6. Generating biological hypotheses – Workshop 2 worksheet. Lecture material is available in Supporting File S2. Generating biological hypotheses – Scientific method lecture slides.
Item #2: Workshop #2 and group check-ins (Think-pair-share; active learning)	Students will read an excerpt from an introduction section of a paper and generate a hypothesis and predictions. Check-in with each group about the workshop and their assignments	~30 minutes	Refer to Supporting File S7. Generating biological hypotheses – Assignment 1 revision instructions for Workshop #2.
Homework/Final Assessment			
Revision of Assignment #1 (Summative Assessment)	Students will revise their hypothesis and predictions	1 week	Refer to Supporting File S8. Generating biological hypotheses – Hypothesis accuracy scoring data.

Table 2. Example hypotheses and scores.

Example hypotheses written by students	Criteria	Justification	Accuracy
If zebrafish are treated with levothyroxine or PTU to mimic hyperthyroidism or hypothyroidism, respectively, we will observe a change in the proliferation of NSCs in the hypothalamus because of dysregulation of Wnt/Notch signaling of <i>her2</i>	Criterion #1	Hypothesis fits the "____because____" scaffold and is not a "If____, then____" prediction	Score of 1+1+1 = 3
	Criterion #2	Students are testing NSC proliferation which is in alignment with their experiment	
	Criterion #3	Students mention a specific biological mechanism that will explain the changes in number of neural stem cells	
Thyroid hormone levels influence neural stem cell proliferation because thyroid hormones help control intracellular pathways	Criterion #1	Hypothesis fits the "____because____" scaffold and is not a "If____, then____" prediction	Score of 1+1+0.5 = 2.5
	Criterion #2	Students are testing NSC proliferation which is in alignment with their experiment	
	Criterion #3	No specific mechanism mentioned, i.e., "which intracellular pathways?"	
Using PTU to induce hypothyroidism in zebrafish will give rise to changes in neural stem cell proliferation through neurogenesis	Criterion #1	Hypothesis fits the "____because____" scaffold; student use the word "through" instead	Score of 1+1+0 = 2
	Criterion #2	Students are testing NSC proliferation which is in alignment with their experiment	
	Criterion #3	Incorrect mechanism.	
Neural stem cells are more likely to commit to an oligodendroglial progenitor fate in a low thyroid environment because TH is responsible for neural stem cells to commit to a neural progenitor fate	Criterion #1	Hypothesis fits the "____because____" scaffold and is not a "If____, then____" prediction	Score of 1+0+0 = 1
	Criterion #2	Testing cell fate instead of cell proliferation = not aligned to their experiment	
	Criterion #3	Incorrect mechanism.	
If PTU is added to the environment of a zebrafish larvae, then thyroid hormone will decrease and inhibit zebrafish development	Criterion #1	Hypothesis fits "If____, then____" scaffold which is how a prediction is written	Score of 0+0+0 = 0
	Criterion #2	Their experiment was not measuring zebrafish "development"	
	Criterion #3	No specific mechanism mentioned.	

Table 3. Improvement in hypothesis accuracy score.

Spring Semester 2020: "lecture+active learning" cohort									
Submission 1					Submission 2				
Group #	Is there a correct distinction between hypothesis and predictions?	Are students "testing" the correct phenomenon?	Are students including a biological mechanism?	Hypothesis accuracy score	Is there a correct distinction between hypothesis and predictions?	Are students "testing" the correct phenomenon?	Are students including a biological mechanism?	Net improvement	Hypothesis accuracy score
group 7	yes	yes	no (not specific)	2.5	yes	yes	no (not specific)	no change	2.5
group 8	yes	yes	no (not specific)	2.5	yes	yes	yes	correct mechanism	3
group 9	yes	yes	no	2	yes	yes	incorrect mechanism	incorrect mechanism	2.5
group 10	yes	no	no	1	yes	no	no	no change	1
group 11	yes	no	no	1	yes	yes	yes	correct test and mechanism	3
group 12	no	yes	no	1	yes	yes	incorrect mechanism	incorrect mechanism	2.5
group 13	no	no	no	0	no	no	no	no change	0
group 14	no	no	no	0	yes	yes	yes	complete correction	3
group 15	no	no	no	0	yes	yes	incorrect mechanism	incorrect mechanism	2
group 16	no	no	no	0	yes	yes	incorrect mechanism	incorrect mechanism	2
Yes	11	10	6		9	8	3		
No	5	6	10		1	2	7		
yes %	68.8	62.5	37.5		90.0	80.0	30.0		