JOURNAL OF MICROBIOLOGY & BIOLOGY EDUCATION
DOI: https://doi.org/10.1128/jmbe.v18i3.1367

Perspectives

# Using Backward Design in Education Research: A Research Methods Essay †

Jamie L. Jensen[1], Elizabeth G. Bailey[2], Tyler A. Kummer[1], and K. Scott Weber[3]
[1]*Department of Biology, Brigham Young University, Provo, UT 84602;*
[2]*Department of Biology, Georgetown University, Washington, DC 20057;*
[3]*Department of Microbiology and Molecular Biology, Brigham Young University, Provo, UT 84602*

**Education research within the STEM disciplines applies a scholarly approach to teaching and learning, with the intent of better understanding how people learn and of improving pedagogy at the undergraduate level. Most of the professionals practicing in this field have 'crossed over' from other disciplinary fields and thus have faced challenges in becoming experts in a new discipline. In this article, we offer a novel framework for approaching education research design called Backward Design in Education Research. It is patterned on backward curricular design and provides a three-step, systematic approach to designing education projects: 1) Define a research question that leads to a testable causal hypothesis based on a theoretical rationale; 2) Choose or design the assessment instruments to test the research hypothesis; and 3) Develop an experimental protocol that will be effective in testing the research hypothesis. This approach provides a systematic method to develop and carry out evidence-based research design.**

## INTRODUCTION

With increasing frequency, accomplished biologists meet with discipline-based education researchers (DBER scholars) to ask for assistance in designing a robust education study testing the effectiveness of a particular pedagogical technique in their classrooms. This is likely a common experience among DBER scholars, as we are integrated in our disciplinary fields and share our expertise with our colleagues. Colleagues like these biologists are experts in their particular disciplinary field and in experimental design but have not been formally trained in performing education research. Thus, we often try to relate their education experiment to the research that they conduct. We do so by asking, *What is your hypothesis? How will you test it? What results would confirm or refute your hypothesis?*

While this is an appropriate way to approach education research, many researchers find it frustrating to apply the objective, quantifiable, scientific, experimental design they use in their field to education research experiments involving human subjects, classroom environments, and less controlled conditions (at least not the type of control with which they are familiar). We are certainly not saying that education research cannot be run in an empirical, quantitative, and controlled manner; it can. Determining how to do this, however, poses a problem that traditional researchers often find frustrating, especially when they are unacquainted with learning theory.

In this article, we introduce a framework for educational research design called Backward Design in Education Research (BDER), patterning it after backward curricular design (1). This method builds upon the Scholarship of Teaching and Learning (SoTL) and teaching-as-research (www.umass.edu/ctfd/teaching/cirtl.shtml) approaches and scaffolds research design to assess pedagogical effectiveness that goes a step beyond "what works" questions to "why" questions in order to be more generally applicable to the education community. With BDER, the goal of the research project (i.e., what hypothesis do we want to test) is first identified and then used by the researcher to help design the appropriate curricula, assessments, and learning activities for the study. Just as backward curriculum design is used to improve the student learning experience (2, 3), we believe that BDER can lead to improvements in our ultimate research goal: providing evidence-based pedagogy and a deeper understanding of causal mechanisms for the broader education community.

*Corresponding author. Mailing address: 4102 LSB, Department of Biology, Brigham Young University, Provo, UT 84602. Phone: 801-422-6896. Fax: 801-422-0090. E-mail: Jamie.Jensen@byu.edu.
Received: 8 June 2017, Accepted: 29 July 2017, Published: 5 October 2017.
†Supplemental materials available at http://asmscience.org/jmbe

## WHAT ARE DBER AND SOTL?

The National Research Council (4) defined discipline-based education research (DBER) by the common goals that many disciplines within the sciences share: to understand how students learn within the discipline, to study the teaching strategies that help students learn, to inform classroom practices, and to study the path to inclusivity of all students in these disciplines. DBER sets itself apart from traditional educational psychology and cognitive science research by taking an empirical approach to teaching and learning through the lens of disciplinary expertise. The example we use is SoTL-based with implications for broadening findings to the wider education research community. While SoTL is considered a relatively young field (5), it has its roots in the larger development of DBER. DBER encompasses a much broader view of research than just classroom interventions and tests of pedagogy (see www.unl.edu/dber/action-research-sotl-dber for a comparison between DBER and SoTL). However, SoTL is an excellent place for researchers to start as they build more generalizable understanding of teaching and learning.

Many practicing DBER scholars have "crossed over" to education research from the discipline in which they received their graduate training. Transitioning into this discipline is not without its challenges. For one thing, many researchers are transitioning from studying populations with more tightly controlled conditions (e.g., mice, flies, bacteria) to human subjects, who bring a plethora of extraneous variables into the equation as well as serious ethical considerations and constraints. Educational research has been compared with clinical research, as it faces many of the same challenges (6). Another challenge for discipline-specific researchers is learning the relevant learning theories to support your research. Thankfully, there are many resources available that give a great overview of several relevant theories (e.g., *How People Learn* (7), *Scientific Teaching* (8), *Educational Psychology* (9)). Other resources are also available to introduce researchers to some of the unique methodologies involved in this discipline (e.g., *Discipline-Based Education Research: A Scientist's Guide* (10), *Entering Research: A Facilitator's Manual* (11)). While all of these resources can facilitate components of research projects, this article presents a simplified and direct framework for complete education research design.

## HOW IS BACKWARD DESIGN APPLIED?

In order to discuss ways in which to properly design SoTL research, we first need to introduce backward design for curriculum development (1). Backward design is a useful method of designing learning activities with the end-goal in mind. This process consists of three steps: 1) Identifying the desired result, i.e., defining your learning outcome, 2) Determining the acceptable evidence, i.e., designing your assessments, and 3) Planning the learning experiences and instructional materials you will use. The purpose of backward design is ultimately to improve student performance by following a purposeful designing process that allows instructors to align their teaching practices with the outcomes they are trying to achieve. Backward design has played an influential role in the design of courses throughout higher education, some being described in the scientific literature (e.g., 12–18). Backward design is now implemented and taught in the National Academies Summer Institute (19) for training current faculty and the National Science Foundation–funded Faculty Institutes for Reforming Science Teaching (20) programs for post-doctoral training.

Applying a form of backward design to research is not a new idea. One example from the field of market research is called "Backward Market Research," (21) which consists of eight steps that ultimately resemble Wiggins and McTighe's Backward Design for curriculum development. The key to backward market research lies in identifying the desired outcome (i.e., what data would answer the question you are asking) before embarking on the project to avoid mindless fishing expeditions. As Pearson (22), a statistician, points out, "Mindless fishing expeditions are unlikely ever to catch a fish worth eating" (p. 16). The same is true in education research. In an effort to collect data and try to take a quantitative, data-driven approach to teaching, novice researchers often make the mistake of plunging into the data collection process without considering the underlying pedagogical problem they are trying to solve. To avoid aimless data collection, we introduce a structured approach, a Backward Design in Education Research (BDER) approach, taking components of both the Andreasen (21) and the Wiggins and McTighe models.

### Three steps in backward design in education research

Table 1 compares Wiggins and McTighe's Backward Design principles for curriculum design and BDER principles. Filling in both columns will help researchers design experiments that will lead to the most useful results (also provided in the Appendix with examples, Tables S2–S5).

**Step 1: Defining your research question.** As an example, a group of biologists wanted to integrate sequencing and analysis of students' microbiomes into their advanced micro/molecular biology courses. For the purpose of this article, we will refer to them as the Microbiome researchers. On initially meeting with the Microbiome researchers, we asked them, "What is your research question?" These researchers had a clear hypothesis in mind: they hypothesized that by having access to their own personal microbiome data, students would be more motivated, which would lead to greater conceptual understanding. This hypothesis answered the causal question, "Why might personal data collection facilitate microbiology learning?" Defining your research question is defining clearly the reason for your research and articulating a distinct and testable hypothesis.

TABLE 1.
Worksheet to enable the use of Backward Design in Education Research.

| Backward Design in Pedagogy (Wiggins and McTighe) | Backward Design in Education Research | Components of a Research Project |
|---|---|---|
| What are your learning outcomes? What do you expect students to be able to do when they finish? | What is your research question? What hypothesis do you want to test? What is your theoretical rationale? | Defining your research hypothesis and theoretical rationale |
| How will you assess these learning outcomes? | How will you assess the accuracy of your hypothesis? | Choosing/designing your assessment instruments |
| What learning activities will you use to accomplish these outcomes? | What experimental protocol will you use to test your research hypothesis? | Developing your experimental protocol |

Notice that the question these researchers are trying to answer is a *why* question. A good ground rule is *Don't just ask What, ask Why*. Many SoTL research projects focus on answering a question about what works and what does not. This is certainly a great first step. But, to be truly meaningful and transferrable to other classrooms and/or disciplines, a more important focus is on *why* it does or does not work or under what circumstances it might or might not work, a hallmark of DBER 2.0 (23). A *what* approach to this problem is simply to test whether adding authentic data collection to their classroom improves learning over the course without authentic data, testing the hypothesis that *using authentic data helps students learn*. The answer to this is indeed important. However, the minute you see that it does work, you will be wondering *why* it worked and you will need to begin your experiment anew. However, if, *prior* to running a test, the researchers take the time to consider a causal mechanism and base the hypothesis in a learning theory, then the results can help to solve a pedagogical problem by offering a solution with an underlying causal mechanism. Going beyond the *what* to the *why* necessarily changes the focus of your study and the data that need to be collected.

Basing your hypothesis in a theoretical rationale can be difficult (especially for those unacquainted with learning theories). Interestingly, in a search of all education articles published in the *Journal of Research in Science Teaching* from 1965 to 2005, only 17.8% of the articles used "theory" in conjunction with "hypothesis" and "prediction" (24). Improvement in our use of theory in SoTL research is needed. So, where do we start? A thorough literature search and/or an in-depth discussion with a DBER scholar may help you to identify learning theories that justify your hypothesis and let you know that there is at least some solid evidence that would support its plausibility. While a literature search does not make you an expert, it is a good place to start as it will direct you to the writings of individuals who *are* experts in the field. Additionally, many resources are available outlining relevant learning theories (e.g., 25, 26). At this point, our Microbiome researchers had defined their research question. They had come up with a clear *why* hypothesis and they had based it on motivational theory (27, 28), i.e., the more motivated a student is, the more likely they are to learn.

**Step 2: Choosing or designing the assessment instruments.** Choosing the appropriate assessment instruments may be the most important step in this process. Let's return to our Microbiome researchers. They needed to collect two types of data: 1) data that would answer the *what*, i.e., what students learned; and 2) data that would answer the *why*, i.e., why did they or didn't they learn more in one treatment over another. Their plan was to allow each student to submit their own samples and obtain their own microbiome DNA sequence data (those in the non-treatment section would receive simulated data). They planned to have students work on assignments with the data and learn about microbiomes in the classroom. They predicted that those students who had their own data would learn more than those who did not (because of increased motivation). To assess learning (the *what*), we needed a test for students to take. But, before we could design the test, we needed to identify the learning outcomes for the activity. What did we expect students to learn? What did we expect students to be able to do when they finished the activity? Furthermore, would we expect to see a change in student learning given our treatment? In the microbiome study, we determined their original learning outcomes did not solicit the deep conceptual understanding the researchers were hoping to achieve. Through an iterative process, we redesigned learning outcomes that were truly representative of what these researchers hoped to accomplish in the class. These new outcomes resulted in assessment items that better captured deep conceptual understanding that we would predict to change given our treatment and the theoretical rationale that motivation leads to deeper conceptual understanding. Table 2 shows the progression of design of some sample learning outcomes along with the difference in assessment items we could create based on improvements to the outcomes.

So, at this point, we had one outcome measure (the *what*): students' deep conceptual learning. This is an indirect measure of our actual causal hypothesis (the *why*): By having access to their own personal microbiome data, students will be more motivated, which will lead to greater conceptual understanding. If we believe that our intervention is increasing student motivation (our causal mechanism), then there ought to be a way to address that. This takes us back to

TABLE 2.
Sample progression through expected learning outcomes for the microbiome unit of the microbiology course.

| Original Outcome | Questions to Consider | Revised Outcome |
| --- | --- | --- |
| Identify the most common bacterial phyla found in your gut<br><br>*Example Assessment Item: What is the most common bacterial phyla found in your gut?* | Do you want students to just look at the website and write this down? Or do you want them to process this information in some way? | Compare and contrast the most abundant phyla of bacteria found in your gut versus the average gut and their proposed functions in health and disease.<br><br>*Example Assessment Item: Draw a diagram showing the similarities and differences between your gut microbiome and the average person's. Explain how this might affect your health.* |
| Be able to outline the importance of factors such as antibiotics and diet (types of foods/probiotics/etc.) on microbiome diversity and composition.<br><br>*Example Assessment Item: Describe how antibiotics and diet affect microbiome diversity and composition.* | This implies that you want students to write a list of each factor and what it does to microbiome diversity. Do you want them to be able to apply this information? Do you want them to build a conceptual understanding that goes beyond rote memorization? | Predict how factors such as antibiotics and diet (types of food, probiotics, etc.) might affect microbiome diversity and composition.<br><br>*Example Assessment Item: Given a particular bacterial profile, draw a graph depicting what it might look like after two rounds of antibiotics.* |
| Be able to outline how 16S sequencing data can provide all of this microbial information.<br><br>*Example Assessment Item: Explain how 16S sequencing data provides the information we got in our sample data.* | Again, do you just want students to recall what you told them in class or do you want them to understand the implications of the uncertainty of 16S sequencing data? | Evaluate claims made about the microbiome using 16S sequencing data.<br><br>*Example Assessment Item: Sally makes the claim that because her 16S sequences suggest higher levels of bacteroidetes, she likely processes lipids more effectively. Evaluate the validity of her claim and provide reasons for your decision.* |

what we iterated before—the importance of asking *why*. If our hypothesis had just been that having personalized data leads to greater learning, our only dependent variable would be our student learning assessment. Now, let's say that we find astounding results—those with their own data significantly outperformed those without. All this really tells us is that in this particular class, with this particular curriculum, giving them a kit to submit their own data somehow benefitted them. It is not easily transferrable to anyone else's classroom and is therefore of limited use to the education research community. However, if we can show that by giving students the ability to analyze their own data, they exhibited behaviors that suggest they were more motivated to learn, now we have something potentially broadly applicable. So, what are those behaviors? There are many ways to assess motivation. We just have to find the one that is most practical and conducive to our particular classroom environment. In our example, most of the work was done online, so researchers chose to periodically interrupt the online assignments to ask questions about motivation that are aligned with motivational theory (e.g., How interested are you in the topic? How many websites have you visited outside those assigned? How much time have you spent researching this?). As you look for appropriate assessment methods, a thorough literature search or conversation with a DBER scholar can help you find robust methodologies.

**Step 3: Developing your experimental protocol.** This is the fun part and, probably as a consequence, is often where we naïvely start. The same thing occurs in teaching, and we, the authors, are no strangers to this. However, using backward design and starting with learning outcomes can greatly transform your teaching by making your actions in the classroom more meaningful, more effective, and more in line with your assessments. We believe that using BDER will similarly transform the actual research protocol you use into something more meaningful and purpose-driven, more effective in giving you the data you need (i.e., the data that would test your hypothesis), and more in line with your research assessments, making the results you gather more broadly applicable to other settings.

To explain this, let's return to our Microbiome researchers once again. They have clearly defined their research hypothesis and established how they will measure these outcomes. Now they must decide how to implement the project: How long will the intervention be? What will the intervention look like? What platform will they use? How will they collect motivational and learning data to ensure high participation rates and the least amount of sampling bias? What will the reference group do instead of the intervention? What variables will need to be controlled? What is the time frame? What approval do they need from their Institutional Review Board (IRB)?

How will they ensure group equivalence or control for non-equivalence? This last question is an important one, especially when dealing with quasi-experimental designs such as are common in DBER. In a quasi-experimental design, assignment of subjects to treatment groups is not entirely random. In the case of our Microbiome researchers, students were assigned to a treatment group based on the section in which they enrolled. It is important to ask whether this division of research subjects might introduce a variable that may account for differences seen in the outcomes of the research, e.g., if one class is a day-time class and the other is an evening class with a high percentage of non-traditional students. Thus, it is necessary to consider potential differences between groups and collect additional data that can be used in the statistical analyses to account for these potential differences. In the Microbiome project, researchers collected a measure of scientific reasoning ability from all students, as this has been shown to be an important predictor of performance in science classes (29–31). This score was used as a covariate in their analyses to control for any potential differences in academic ability between groups. Gathering standardized test scores or previous grade point averages is also common practice. Applying these values to your statistical models can be done in many ways. One common way to do this is to use multiple linear regression (see Theobald and Freeman (32), for an excellent discussion). Considering other variables, such as student demographics, may also be important in interpreting results.

By asking these questions, these researchers can go beyond testing whether this particular pedagogical technique is better than the traditional curriculum, to testing *why* it was more successful. Designing the experimental protocol should actually be straightforward if you have completed steps 1 and 2. You should have already established your hypothesis and know exactly what instruments you will use to test it. Now, you just have to figure out the logistics of carrying it out. In the case of the Microbiome researchers, they created an online assignment that required students to use their data (personal or simulated) in a program that allowed researchers to track student access and periodically insert motivation-assessing questions. It was a robust SoTL design to test their hypothesis, with the addition of addressing the causal mechanisms behind what they were seeing in order to apply these results more broadly.

## CONCLUSION

In this paper, we have presented a novel framework for designing effective education research projects and called it Backward Design in Education Research (BDER). We have patterned it after the curriculum design strategy, Backward Design (1), and spring-boarded from a SoTL approach. We have also included some helpful worksheets to help researchers in using this method. It is our hope that this framework can scaffold researchers transitioning into education research, provide a starting point for undergraduate and graduate students beginning their career in DBER, and offer a straightforward and easy-to-follow approach to classroom research in our efforts to improve and strengthen our field of inquiry. We hope that the use of this method will result in a more purposeful and designed approach to our research, providing data that truly test our hypotheses, by focusing on causal mechanisms based in theoretical rationales. As a result, we will better understand student learning, and produce findings that are transferable between courses, institutions, and even disciplines. BDER can serve as a guiding framework as we endeavor to discover the best ways to help students learn.

## SUPPLEMENTAL MATERIALS

Appendix 1: Backward Design in Education Research worksheet and examples

## ACKNOWLEDGMENTS

The authors declare that there are no conflicts of interest.

## REFERENCES

1. Wiggins G, McTighe J. 2005. Understanding by design (expanded 2nd ed.). Association for Supervision and Curriculum Development, Alexandria, Virginia.
2. Fink LD. 2013. Creating significant learning experiences: an integrated approach to designing college courses. Wiley & Sons, Inc., San Francisco, CA.
3. Wiggins G, McTighe J. 2011. The understanding by design guide to creating high-quality units. Association for Supervision and Curriculum Development, Alexandria, VA.
4. National Research Council. 2012. Discipline-based education research: understanding and improving learning in undergraduate science and engineering. The National Academies Press, Washington, DC.
5. Boyer EL. 1990. Scholarship reconsidered: priorities of the professoriate. Carnegie Foundation for the Advancement of Teaching, Princeton, NJ.
6. Robson RL, Huckfeldt VE. 2012. Ethical and practical similarities between pedagogical and clinical research. J Microbiol Biol Educ 13(1):28–31.
7. Bransford JD, Brown AL, Cocking RR (ed). 2000. How people learn: brain, mind, experience, and school. The National Academies Press, Washington, DC.
8. Handelsman J, Miller S, Pfund C. 2007. Scientific teaching. W. H. Freeman, New York.
9. Ormrod JE. 2000. Educational psychology: developing learners. Pearson, Upper Saddler River, NJ.
10. Slater SJ, Slater TF, Bailey JM. 2010. Discipline-based education research: a scientist's guide. W. H. Freeman, New York.
11. Branchaw JL, Pfund C, Rediske R. 2011. Entering research: a facilitator's manual. W. H. Freeman and Co., New York.

12. Armbruster P, Maya M, Johnson E, Weiss M. 2009. Active learning and student-centered pedagogy improve student attitudes and performance in introductory biology. CBE Life Sci Educ 8:203–213.
13. Bergstrom G. 2011. Content vs. learning: an old dichotomy in science courses. J Asynch Learn Netw 15:33–44.
14. Carlson DL, Marshall PA. 2009. Learning the science of research, learning the art of teaching: planning backwards in a college genetics course. Biosci Educ 13(1):1–9. http://dx.doi.org/10.3108/beej.13.4.
15. Davidovitch N. 2013. Learning-centered teaching and backward course design from transferring knowledge to teaching skills. J Int Educ Res 9:329–338.
16. Jensen JL, McDaniel MA, Woodard SM, Kummer TA. 2014. Teaching to the test, or testing to teach: exams requiring higher order thinking skills encourage greater conceptual understanding. Educ Psychol Rev 26:307–329.
17. Linder KE, Cooper FR, McKenzie EM, Raesch M, Reeve PA. 2014. Intentional teaching, intentional scholarship: applying backward design principles in a faculty writing group. Innov High Educ 39:217–229.
18. White PJT, Heidemann MK, Smith JJ. 2013. A new integrative approach to evolution education. BioSci 63:586–594.
19. Pfund C, Miller S, Brenner K, Bruns P, Chang A, Ebert-May D, Fagen AP, Gentile J, Gossens S, Khan IM, Labov JB, Maidl Pribbenow C, Susman M, Tong L, Wright R, Yuan RT, Wood WB, Handelsman J. 2009. Summer institute to improve university science teaching. Science 324:470–471.
20. Ebert-May D, Weber RP. 2006. FIRST—What's next? CBE Life Sci Educ 5:27–28.
21. Andreason AR. 1985. 'Backward' market research. Harvard Bus Rev, May 1985.
22. Pearson RW. 2010. Statistical persuasions. SAGE Publications, Inc., Thousand Oaks, CA.
23. Dolan EL. 2015. Biology education research 2.0. CBE Life Sci Educ 14:ed1,2.
24. Lawson AE. 2010. How 'scientific' is science education research? J Res Sci Teach 47:257–275.
25. Orgill M, Bodner GM. 2007. Locks and keys: an analysis of biochemistry students' use of analogies. Biochem Molec Biol Educ 35:244–254.
26. National Research Council. 2012. Discipline-based education research: understanding and improving learning in undergraduate science and engineering. Singer SR, Nielsen NR, Schweingruber HA (ed). The National Academies Press, Washington, DC.
27. Maslow AH. 1943. A theory of human motivation. *Psychol Rev 50*:370–396.
28. Maslow AH. 1954. *Motivation and personality*. Harper and Row, New York.
29. Lawson AE, Banks DL, Logvin M. 2007. Self-efficacy, reasoning ability, and achievement in college biology. J Res Sci Teach 44:706–724.
30. Johnson MA, Lawson AE. 1998. What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes? J Res Sci Teach 35:89–103.
31. Lawson AE, Clark B, Cramer-Meldrum E, Falconer KA, Sequist JM, Kwon YJ. 2000. Development of scientific reasoning in college biology: do two levels of general hypothesis-testing skills exist? J Res Sci Teach 37:81–101.
32. Theobald R, Freeman S. 2014. Is it the intervention or the students? Using linear regression to control for student characteristics in undergraduate STEM education research. CBE Life Sci Educ 13:41–48.