Just-in-Time Teaching in Statistics Classrooms

Monnie McGee, Lynne Stokes, and Pavel Nadolsky

ABSTRACT
Much has been made of the flipped classroom as an approach to teaching, and its effect on student learning. The volume of material showing that the flipped classroom technique helps students better learn and retain material is increasing at a rapid pace. Coupled with this technique is active learning in the classroom. There are many ways of “flipping the classroom.” The particular realization of the flipped classroom that we discuss in this article is based on a method called “Just-in-Time Teaching (JiTT).” However, JiTT, in particular, and the flipped classroom, in general, is not just about watching videos before class, or doing activities during class time. JiTT includes assigning short, web-based questions to students based on previously viewed material. Typically, Internet-based questions are constructed to elicit common misunderstandings from the students, so that the instructor can correct such misunderstandings immediately in the next class period; hence the name, “Just-in-Time Teaching.” The addition of these pre-class questions is what separates JiTT from a general flipped classroom model. Even as the research on the superiority of JiTT as a learner-centered pedagogical method mounts, aids for the instructor have not, at least not as quickly. This article is focused on the instructor—with aids to help the instructor begin using the JiTT flipped classroom model in statistics courses.

1. Introduction
Since the 1980s and 1990s, the pace of research on STEM education has accelerated and produced valuable new insights about teaching and learning at the college level. Technological advances, social pressures, rising college expenses, and, finally, the advent of massively open online courses (MOOCs), are compelling professors to question whether the traditional lecture style of teaching is still up-to-date, when students can obtain much of the same content, for free, from another source (Berrett 2012). In order to make bricks-and-mortar institutions worth the cost, such institutions need to provide the one thing that MOOCs cannot – personal access to faculty. At the same time, scientific research on learning has identified several universal practices of effective teaching that are applicable to both child and adult learners. The key findings of this research are: (1) learners can access and use their new knowledge more readily if it is integrated into their pre-existing mental structure and connected to the facts that they already know; (2) learners are more successful when they can recognize what they understand and what they do not know, which is known as “developing metacognition;” and (3) learners’ erroneous prior beliefs should be purposively identified and addressed to improve knowledge acquisition and retention (Bradford et al. 2000, pp. 14–19).

These developments lead to a revelation that the one-way lecture method is not the best way to disseminate information to students; at least, not in terms of students’ knowledge retention. Cardinal revisions in the traditional structure of lecturing, based on active learning and multi-directional interactions among all participants, have been repeatedly demonstrated to produce better learning outcomes across a variety of STEM disciplines (Kober 2015).

One way to provide more interaction with faculty, as well as to adhere to the newly recognized learning principles, is through implementing the Just-in-Time Teaching (JiTT) model. JiTT was conceived in 1996 at Indiana University, Purdue University Indianapolis, and the US Air Force Academy (Novak et al. 1999). The aims of the method are to help students (exclusively nontraditional students at the outset) to structure time out of class and to use class time for student–instructor interactions. The Internet helped facilitate this effort by increasing ease of communication outside of class. Instructors could use e-mails, at first, and then learning management systems, to make various reading materials and/or lecture videos available. Prior to class, the students were to examine the materials and come to class ready to work problems or have discussion on what was already consumed.

However, JiTT, in particular, and the flipped classroom, in general, is not just about watching videos before class or doing activities during class time. JiTT includes assigning short, web-based conceptual questions or analysis problems to be completed outside of class. The questions or the short assignments must be answered before class and serve two purposes. The first is to encourage students to read or watch the preview material prior to class time. The second, and more important, is to identify misconceptions of the students, so they can be directly addressed and corrected, as called for by the learning principle...
of identifying erroneous prior beliefs and addressing those beliefs as soon as possible. The construction of these “warm-up,” “JiTT,” or “preflight” questions is important. Research in learning suggests that students are strongly influenced by their preexisting understanding of the subject matter. When the understanding is wrong, they are often unaware of this, and their beliefs interfere with their ability to absorb new information. Educational experiments have shown that effective teaching requires that these preexisting misunderstandings be actively elicited and corrected (Bransford, Brown, and Cocking 2000). Therefore, it is not enough to ask students what they do not understand or what is unclear. The warm-up questions are designed to obtain this information.

Once again, technology facilitates this. We use our course management system (Blackboard) to post questions for students to answer prior to class. Options within Blackboard allow the professor to designate a time frame in which the questions are available. We usually close the question submission period 2 hr prior to class time, but the optimal closing time is subject to the discretion and schedule of the instructor. Since student responses are recorded online, the professor can use example responses to show particularly thoughtful answers, or to illustrate common misconceptions. Very often, the warm-up questions reveal misconceptions that an instructor could not have foreseen (see Discussion). Such misconceptions can be discussed in class before students solidify them in their minds. Thus, the identification and correction of erroneous beliefs occurs “just in time” as opposed to during a homework assignment or an exam review. In-class activities can be tailored to address the misconceptions, or to extend concepts in the preclass material, as appropriate.

By its structure, JiTT is closely related to the “inverted,” or “flipped,” classroom method, according to which learning activities that traditionally take place inside the classroom, such as learning new information for the first time, now take place outside, and vice versa (Lage, Platt, and Treglia 2000). In the most radical flipped model, all new material is presented in the form of short videos, and all class time is taken by practical exercises. However, no strict dichotomy exists between the “traditional” and “flipped” models (Gavrin 2015). The authors of JiTT point out that their method lies within the continuum between the two. The core idea of the inverted classroom is to relegate information delivery to the preclass preparation, with the goal to make in-class instruction more interactive. This is also the main objective of JiTT. Although the preclass preparation of students in JiTT may sometimes be lighter (e.g., it may require them only to read the assigned textbook material), it is still more rigorous than in the traditional model. Our realization of JiTT also assigns preclass videos and in this sense is even closer to the “flipped” end of the spectrum. Therefore, we will describe it as a “JiTT flipped classroom method.”

The general flipped classroom method has been shown to be effective at improving student attitudes and student learning in statistics courses. Camp, Middendorf, and Sullivan (2010) explain how use of flipped classroom techniques motivated student learning in a statistics course with 300 students. Windquist and Carlson (2014) showed that psychology majors at Valparaiso University taking a flipped version of an introductory statistics course not only learned the material better initially, but also retained that material for 20 months better than students in a lecture-based section. Retention in the flipped classroom based statistics course was also better than retention in other lecture-based core courses within the department (e.g., Introduction to Psychology, Developmental Psychology). Earlier work (Carlson and Windquist 2011) showed that, in a workbook-based introductory statistics course that is similar to the flipped classroom, students demonstrated improved attitudes toward statistics, as measured by the Survey of Attitudes Toward Statistics (SATS) (Chiesi and Primi 2009).

The JiTT approach was originally invented to facilitate flipped structure in physics courses by combining preclass reading assignments and other preparatory work with simple assignments that students complete online before discussion in class. We have used the JiTT method of implementing the flipped classroom in two different types of statistics courses. The first course is an introductory course for prospective business students. Students meet twice a week in a large lecture (up to 100 students), and once a week in smaller labs of 25–30 students. The lab component of the course is meant to introduce the students to real-data analysis using Excel. The second course is an intermediate-level two-course sequence in statistical methods required of all statistics majors, minors, and professional masters students. The same course is often taken by graduate students in other departments. The upper-level course focuses on experimental design and linear models, using SAS as the main software package for data analysis. In the past, class sizes in the second course ranged from 8 to 20 students, depending on the semester. In addition, colleagues in physics regularly use this approach, and colleagues in mathematics have also experimented with it. One of the authors has been implementing flipped instruction based on similar principles in the SMU Department of Physics for several years, and some ideas were exchanged across the disciplines (Hake 1998).

Furthermore, colleagues in the humanities have been using this approach for decades (or perhaps centuries). In most humanities courses, students are required to read material before coming to class and discuss the already digested material in class. When we first mentioned flipping the classroom to colleagues in the humanities, they said, “We don’t call that flipping the classroom. We call that good teaching.” In this article, we concentrate on using JiTT in an introductory statistics course, since this is the course for which we have the most students, and therefore, the most data. However, the knowledge we have gained in applying JiTT in the introductory course is applicable to all level of statistics courses (and STEM courses, in general).

As instructors of statistics, we want our students to internalize certain statistical concepts and methods, regardless of the level of the course. Research for the introductory statistics course is definitive that active learning and JiTT techniques improve student learning of statistical concepts (Steen 1992; Zappe et al. 2009; Windquist and Carlson 2014) and attitudes toward statistics courses (Carlson and Windquist 2011). However, mastering these strategies can be daunting at first, especially to those instructors whose own education followed the traditional format. In short, it is difficult to teach in a way in which we were not instructed.

In this article, we discuss the components of JiTT and describe how they address the principles of learning science
spelled out above. We further outline how an instructor using the JiTT technique can introduce a lesson in linear regression. We also give examples of curricular materials and discuss how to develop them so that they adhere to these learning principles. Finally, we discuss methods of grading and assessment, and how instructors at all levels (and across all disciplines) can work together to develop resources for implementing the JiTT technique.

2. Components of Just-in-Time Teaching

There are three keys to implementing a successful JiTT classroom. We list them here, and each will be explained in detail in the following.

- Prepare lecture material for students to view prior to class.
- Prepare activities or problems to be solved in class.
- Create questions over precourse materials that are to be answered prior to class time.

The first component is to assign material for the students to preview prior to a given class period. The material can be in the form of reading from the textbook, completing an activity in a workbook (e.g., Carlson and Windquist 2011), reading articles discussing new research findings or applications, simulating processes using web-based applets (e.g., http://www.rossmanchance.com/applets/), or watching videos that preview the content for the upcoming course period. This step aims to ensure that, when the students arrive to the classroom, they already have a framework within which to place the material to be covered, as required by the first key scientific finding discussed in the introduction.

For example, before introducing simple linear regression, a precourse video (or a text document) reviewing only the basics of a linear function may be required viewing for students. In that way, all will have the same background for the new idea of summarizing data with a line. The objective of this approach is not to simply gain more instructional time, but rather to provide connection points consisting of previously learned or recently observed phenomena (e.g., a surprising simulation), to which the new concepts presented in the classroom will be related.

Another example would be to assign the student to examine an applet that introduces the susceptibility of Pearson’s correlation measure to outliers in the dataset.

There are many relevant, well-produced videos available through YouTube for all the concepts covered in introductory and intermediate level courses. Several publishers also have applets and videos associated with their texts. However, we found that students prefer even a less professionally developed video made by their own professor. The reasons for this are varied. One reason is that videos from other sources often use different notation than is used in the textbook for the course, which can be confusing to students. Another reason is that instructor-produced videos make the students feel like the instructor is present and engaged in the course. From our own student evaluations, we have also seen that students perceive the instructor as “lazy” if solely YouTube or publisher-provided videos are used to provide course content (Stokes and McGee 2014). There are a variety of software choices for screen capture with voiceover that make the task of recording and posting lectures to a course YouTube account or to a course management system relatively simple. There is also the software Zaption (https://www.zaption.com/) that allows professors to record material and insert questions into the recorded material for students to answer before they can proceed with the next topic. Most of our videos were created on the instructor’s laptop in Camtasia; however, there are also freeware choices, as reviewed at http://download.cnet.com/windows/video-capture-software/.

The second component of the JiTT model is activities or problems to be solved during class. These can be homework questions from the text, or one of many published activities that are meant for in-class use. The Journal of Statistics Education, and many websites and textbooks (delMas, Garfield, and Chance 2003; Rossman and Chance 2008; Rhem 2010), provide ample resources for interesting activities. The purpose of this component is to improve the metacognition of students, as called for in the second key scientific finding from the introduction. The instructor can help the students develop problem-solving strategies by directly asking them, one-to-one, to explain their reasoning. When the student is stymied, the instructor can provide suggestions, either directly or with leading questions. This approach has been successfully implemented with one instructor in a classroom of 40–50 students.

For larger classes where instructors do not have access to teaching assistants, one-to-one feedback is not feasible. However, it is still possible to provide feedback by grouping students and by taking advantage of peers as additional instructors. Two methods that can be used for this are student response systems and “think-pair-share” (Lyman 1981). Student response systems range from high-tech versions, such as “clickers,” to low-tech versions, such as variously colored index cards or even a show of hands. All of these forms have been successfully used by the authors. In one version of this method, a multiple choice question is posed, typically a conceptual question or one requiring only a quick calculation. Each student selects and submits an answer on the clicker or by holding up an index card or a set number of fingers to indicate a response. (Our experience is that students should be instructed, if using one of the low-tech methods, that their response should be held so it is not visible to their classmates, as they are reluctant to “expose” their answers otherwise, or they quickly learn to follow the most popular, even if incorrect, answer in the class.) If there is substantial disagreement in responses, then students are asked to justify their response by explaining it to a peer. The question is repeated, and students are asked to respond again. Typically, the rate of correct responses increases. A selected student or the instructor should then clarify the correct response.

The “think-pair-share” activity is similar, but it allows practice with a wider variety of question types (Lyman 1981). Students first are asked to work the problem or activity on their own (think). Then, they are instructed to pair with another student and discuss how they arrived at their answers (pair). Finally, student pairs are either called on or asked to volunteer their answers with the class (share). The instructor can also ask students who disagree with the shared answers to describe their thinking processes and solutions.
Such a mode of teaching can lead to lively class discussions and has the side benefit of students getting to know at least one other classmate.

The third component is to provide questions for the students to answer prior to class time. These questions are carefully constructed to elicit misconceptions that students might have, but would be unable to voice if they were simply asked what they do not understand. The questions encourage students to review assigned materials before class time, as well. The instructor reviews the answers to these questions prior to class time, with enough lead time to find appropriate problems that students can work in class, or an appropriate activity that undoes the students’ misconceptions. The adjustment for the misconception depends on its nature. Simple miscalculations (i.e., forgetting how to calculate a correlation coefficient) can be addressed by going over the calculation again in class, preferably with an assignment that students do in groups while the instructor circulates around the room to answer questions. More subtle misconceptions may require the instructor to find outside sources. Potential sources for such problems or activities to do in class would be the course textbook, applets, and activity-based statistics workbooks (i.e., delMas, Garfield, and Chance 2003; Rossman and Chance 2008; Rhem 2010; Schnetter 2015; West 2015).

We stress that the combination of all three components glean the most from the JiTT approach. Simply changing the structure of the course, by delivering the contents by videos rather than by in-class instruction, is less likely to lead to better learning. One reason is that there is some evidence that it is active learning by students, not merely the flipped sequence of information delivery, that is the key variable behind improving students’ learning outcomes (Jensen, Kummer, and Godoy 2015). In addition, repeated recall of information, whether in the form of homework, quizzes, or JiTT questions, aids in student learning (Karpicke and Roediger 2007). The JiTT method motivates students to engage in out-of-class preparation by reading the textbook and watching supplementary materials, and frees precious time in class for active learning interactions. The JiTT structure thus helps the instructor to cover the same syllabus as in the traditional lecture, engages the students to do their share out of the class, while allocating more in-class time for interactive instruction.

3. A Week in the Life of a JiTT Classroom Instructor

A typical week for JiTT instruction in an introductory statistics course (and assuming a Monday, Wednesday, Friday schedule) might look like the following.

- Before Monday class: Assign readings and record an overview video summarizing what will be covered in the following week. Sometimes the video gives an overview of concepts and class time will consist of examples and problems that probe the concepts in more detail. Other times, the video goes over details of calculations so that students can review it for step-by-step practice. Videos are a useful tool for those ideas that are hard to absorb just by listening and not doing.
- Monday: Go over extensions of video material. The Monday class may be lecture-based, or may include lecture interspersed with use of a group response system. In other cases, a class activity, such as a simulation or critical review of a research study in the news might be carried out for a portion of the class period.
- Wednesday: The warm-up question is due. The instructor must review submitted responses prior to class in order to select responses to display and prepare any necessary material to help students resolve misconceptions. For early morning classes, we have had success in selecting examples from among those submitted the evening before. With experience gained in examining warm-up questions, the instructor will learn what responses to expect, and the process of selection will be faster. During class, the instructor will spend 10–15 min reviewing warm-up answers. The remaining class time is typically spent on class exercises or homework problems.
- Friday: Software lab in small sections.

Our first attempts at using JiTT were not completely successful (Stokes and McGee 2014). In the Statistics department, we did not have much guidance, other than knowing that lecture material should be assigned prior to class, that we needed a “quiz” prior to class to assure students were reading the material, and that we should do homework in class. One of us implemented the method gradually, starting first with activities in class, then moving toward presenting material outside of class. The other revamped her course in one fell swoop.

The transition was straightforward in an introductory physics course taught by one of us, where the traditional lecture was replaced by JiTT in 2012 within one semester by following the guidelines that the author had learned at a training workshop organized by a physics professional society. This training was very helpful for learning about important practical aspects emphasized in this article and simplified the transition, notably by providing a set of warm-up questions that resulted in large time savings. The general lesson learned from our experiences is that personal advice from faculty with previous experience in JiTT can go a long way in facilitating the transition. This is particularly

| Do | Don’t |
|---|---|
| Be consistent throughout the semester | Have warm-ups due before every class |
| Be explicit about the method | Cover new material exclusively outside of class |
| Explain the educational research supporting JiTT | Have videos or reading exclusively replace lectures |
| Adjust with the pace of the class | Be degrading toward wrong answers |
| Rehearse class activities prior to class | Assume students retain all information from the earlier parts of the course |
| Include warm-up questions requiring verbal responses, and possibly estimation and multiple-choice questions | Assume students and instructor concur on all conventions |
| Include a feedback form in the warm-up every week | Give up |
| Associate a small percentage of the grade with warm-up questions |
true of colleagues in physics departments, since several authors of the original JiTT approach were physicists.

Table 1 shows what we have learned about the “Do’s” and “Don’ts” in implementing the JiTT flipped classroom method (Stokes and McGee 2014).

Some of these do’s and don’ts are self-explanatory. For example, students should know why they are expected to do more work outside of class, otherwise students will tend to see the extra work out of class as a shift of the “burden of teaching” from the professor to the student (Novak and Patterson 2010). They should be able to expect consistency from the course, even though the material and the pedagogical method may be unfamiliar. However, students should not be expected to digest all new material outside of class, completely on their own. That is why the instructor gives a short lecture in each class period (or spends one class period per week explaining material). The idea is to repeat (briefly) and extend what the students should have already seen in the preclass material. Activities done in class should also apply or extend any preclass material.

The idea of rehearsing activities prior to class has come from several experiences where the activities contained mistakes, had missing information, or were simply too lengthy for class time. In one instance, the professor assigned an activity where the students were to calculate a correlation coefficient for a small dataset using the standard formula, only to find out that the students did not remember how to calculate the standard deviation, which had been covered three weeks earlier. For the record, the purpose of the activity was not to have students memorize the formula. It was to show that the correlation is unit-less and the same regardless of the units of $X$ and $Y$. A copy of this exercise is available in Appendix C (the Appendices are available in the online supplementary information). In this case, the in-class activity revealed that students had not retained information about the standard deviation (hence, do not assume students retain all information), and a quick refresher was needed. However disconcerting that may seem, it is better for the instructor to know about such a deficiency before an exam than to discover it afterward.

The format of the warm-up that we describe is not unique and can vary depending on the learning objectives, curriculum, initial preparation and size of the class. For example, the warm-up assignment can be due at the beginning of each week and can include multiple-choice, conceptual, and “back-of-an-envelope” estimation questions. Alternating between different formats of the questions helps to engage the students and provide “hooks” to initiate various types of activities during the class. The multiple-choice questions on the warm-ups are the simplest; almost all students will answer them to get some credit. The learning management system can quickly count percentages of the multiple-choice answers; the instructor can follow up with a clicker question in the class and compare these percentages before and after the instruction. One of the authors has had success with including a feedback form in each warm-up. Using the form, the students can communicate about any issues with the instructor. The form helps elicit comments from students who might not otherwise approach the instructor.

The warm-ups are efficient when each student answers them individually and generates varied responses. In reality, some students will try to save time and copy the answers from elsewhere. A number of steps can be taken to reduce such copying. JiTT emphasizes the importance of grading the warm-up questions predominantly on the students’ effort and originality of answers, much less on correctness (Novak and Patterson 2010). An example of a grading rubric for warm-up questions, adapted from Novak and Patterson (2010), is given in Appendix A. One of the authors of this article allows students up to five points for submitting a relevant and original answer by the warm-up due date. Students can review their answers after the first submission and in-class discussion and get two additional points if their answers are correct when they are graded a week later. Students are told from the start that they may be asked to discuss their submission with the whole class and that they should be prepared to defend their answer. Every student is thus expected to participate in the class activities, with the understanding that imperfections will happen and will be fully acceptable during the lectures, but not on the tests. This promotes responsible completion of the warm-ups.

In order for the warm-ups to be most successful, students should perceive that they are an important part of their learning process, and not just another component of their grade or a course task to be dispensed with as quickly as possible (Novak and Patterson 2010). Consistency in use of warm-up questions in the classroom helps to develop this as part of the course culture. For example, the one-day-per week class when JiTT questions are due can always be started with discussing the questions. Having students’ exact words and explanations displayed in class as models, with their names shown for exemplary answers, also reinforces the idea that teaching and learning are a shared responsibility in the classroom. It also personalizes the classroom experience and reinforces student accountability, both of which are hard to do in large classes. The instructor can create this feeling of individual attention by showing exemplary answers from as many different students as possible over the course of the term.

4. An Example Lesson

In this section, we give a detailed example of teaching a particular topic using the JiTT method. The topic is descriptive statistics for two quantitative variables, as given in Chapter 2 of the text that is used in our introductory statistics course (Moore et al. 2011).

First, the students are asked to preview a video. The preview video is 26 min long, which is longer than usual. Most videos are 10 min long. This particular video is a voice-over PowerPoint slides, and was recorded using Camtasia on a desktop PC. The preview video covers the following, as given on the first slide of the video.

In this video, you will get a preview of our next week’s topics:

- The most common tools for examining and describing the relationship between 2 numerical variables. You can do this…
- Graphically: scatterplot
• With summary statistic: correlation
• Be introduced to one of the most important tool in statistics: regression

Subsequent slides in the video illustrate a scatterplot, correlation, and regression with a dataset that was collected from the students on the first day of class. The students were asked to report their weight and desired weight (no names were attached to the survey). The video covers usual lecture material on these topics, including the notion of least squares, and interpretation of scatterplots, correlation, and regression coefficients for a simple linear regression.

The instructor also provides the students with links to other videos on these topics. The assignments page on the course LMS contains the following information:

• scatterplots: https://www.youtube.com/watch?v=PE_BpXTyKCE
• correlation: https://www.youtube.com/watch?v=372iaWfH-Dg
• intro to regressions: https://www.youtube.com/watch?v=ocGEhiLwDVc
• what is least squares: https://www.youtube.com/watch?v=jEEJNz0RK4Q

On the first day of class that week, the students spend most of the time on a class exercise that introduces and motivates the concept of \( r^2 \), as well as asks them to interpret the parameters of an estimated regression model (see Appendix B). Experience suggests that just explaining \( r^2 \) is not enough to remove the mystery associated with it. The students need to actually work with the data to have a chance at understanding it. This exercise takes most of the 50-min class period. At the end of the class period, the instructor does a debriefing, in which answers are provided. The students turn in their exercises at the end of class.

Before the second class meeting that week, students answer warm-up questions. Usually two questions are asked, but there is a repertoire of more than two so that answers cannot be so easily passed from one student to another. Here are four examples of warm-up questions:

1. Watch the TED talk on women and babies here http://www.ted.com/talks/lang/en/hans_rosling_religions_and_babies.html
   This is a remarkable illustration of how a simple tool like a scatterplot can be used to understand a phenomenon. Discuss the relationship between income and fertility rate (children per woman) using the four features of the scatterplot we discussed: form of relationship, direction of relationship, strength of relationship, and outliers
2. The regression equation that predicts desired weight from weight for our class was (Desired wt) = 1.02wt – 5.5. The correlation between weight and desired weight was 0.928. Suppose the weight and desired weight had been reported in kilograms rather than pounds. What would happen to the correlation? What would happen to the regression equation? Explain your thinking.
3. Suppose that weight and desired weight have a correlation of 0.98. Does this mean that the distributions of weight and desired weight are nearly the same? Specifically, must the means of the two distributions be close to each other? Explain your thinking.
4. Suppose a store reduced its entire inventory to half price. What would be the correlation between original prices and sale prices? Explain your thinking.

Before the beginning of the second class period, the instructor examines the answers and chooses some good answers to display to the class. Often, there will be answers that are quite different from each other, but each hitting on an important point. The goal of showing different good answers is to show the students that they should not be desperate for the “only” right answer.

For example, when discussing the first question about the TED talk above, the following two responses were displayed in class:

1. Response from a student who did a good job of interpreting the scatterplot
   In the TED talk the speaker discusses factors that lead to women producing more babies. Is it wealth? Religion? or Something Else? The scatterplot shown in the presentation shows that as the years have gone by the influence of religion on the fertility rate does not differ between religions, however money does seem to play a role. In lower income level countries, women tend to produce more babies than in higher income level countries. Through the years the income level has played less of a role, but is still a strong influence on predicting the number of babies a woman might produce. As the income level increases the number of babies a woman has decreases. There is strong relationship with money and fertility rate, but there is not a strong relationship between religion and fertility rate.
2. Response from a student who did a good job of answering a specific question about the features of the scatterplot
   …The poorer the country, the greater the fertility rate. The direction of the relationship is negative because fertility decreases as income increases. The relationship is strong. Finally, the form of the relationship is linear but not a perfect linear. The 1960 scatterplot had quite a few outliers, most notable being Qatar.

This also provides an opportunity to explicitly correct misunderstandings that students might have. For example, for question 4 above, one student wrote the following:

The correlation between original prices should be a perfect –1 because, if every single item in the store is being decreased by a flat rate, then every corresponding new price will be lower than the original while maintaining proportions, providing a perfectly straight line with negative direction and a slope of –0.5.

Discussion of this response provided an opportunity to emphasize how useful it is to sketch the described data.
After discussion of the warm-ups, the instructor will lecture briefly on topics that are difficult for students, or ones that have not been previously covered. For example, for the regression topic, the instructor might cover the idea that prediction can be improved in some cases by allowing separate regression lines for two subgroups in the data. Then, the instructor directs the students to related homework problems, which they are encouraged to work on for the remainder of the class, while the teaching assistants and instructor are available to assist. Here are examples of problems used in the regression unit:

1. The attached file shows a scatterplot for the weight and desired weight data from a previous class. It shows two regression lines for predicting desired weight from weight, separately for males and females. Predict the expected desired weight, in pounds, of a female who weighs 140.

2. The predicted desired weight for a 140-pound female is quite different from the prediction we’ve made using a single regression line for males and females. Which prediction do you think is better, and why? Your answer should discuss statistical properties of the two regression lines, not your knowledge of how males and females are different. That is, the answer “Females want to lose weight more than males, and so it is more accurate” is not the answer I am looking for here.

Other important concepts can be presented in the same manner.

5. Advantages and Disadvantages

As mentioned in the Introduction, the advantages for active learning for the student, as manifested by the flipped classroom, are clear. The advantages for the instructor, to our knowledge, have not been discussed as much. Here, we discuss advantages specifically of JiTT, although many of them apply to other flipped classroom strategies and active learning as well.

One major advantage for the instructor is that she gets deeper appreciation of what confuses the students from extended answers on warm-ups and class activities (Marrs, Blake, and Gavrin 2003; Guertin 2010). The instructor can correct such issues before they become habits. The problem in calculating the standard deviation in the previous section is one example. We will now give examples of two more warm-up questions, used in an introductory statistics course, each of which shows how such questions can reveal misunderstandings.

**Example Question 1:** Predict what you think the shape of the distribution of desired weight for students in this class looks like. Explain why you think it will look that way.

The expected correct answers, which occurred, were that weight would be either bimodal, because male and female classmates differ in weight, or that it would be normally distributed, because weight is a “natural phenomenon” that is commonly normal. However, a misunderstanding that was not anticipated was represented by this response: “I think the shape of the distribution would be skewed to the left because people want to weigh less, ….” This remark illuminated why some students had such a hard time with understanding the meaning of skewness throughout the course. They interpret the statistical concept of skewness with the height of the curve (the vertical direction), rather than a long tail in the horizontal direction. In other words, they reverse the x and y axes in their understanding of the term skewness. This misunderstanding can then be explicitly corrected, rather than simply expecting students to take the definition of skewness and apply it themselves.

**Example Question 2:** A 95% confidence interval for the mean \( \mu \) of a population is computed from a random sample and found to be 9 ± 3. Select the best conclusion from the responses below, and explain your choice.

a. We are 95% certain that each member of the population has a value between 6 and 12.
b. We are 95% certain that the sample mean is between 6 and 12.
c. We are 95% certain that the population mean \( \mu \) is between 6 and 12.

d. The probability is 95% that if we chose a member from this population at random, its value would be between 6 and 12.

Example 2 is a multiple choice question, crafted so as to elicit certain common misunderstandings about the interpretation of confidence intervals. The correct answer is “c.” However, most students choose “d.” This reveals two misconceptions. First, sometimes students perceive that the confidence interval is about individual members of the population, rather than a parameter (or, perhaps, that students haven’t yet internalized the difference between a parameter and individual). Second, it reveals difficulties with the interpretation of the confidence level. This question can also be used in class as a “clicker” question, or as a “think-pair-share” exercise, where students discuss their answers with a partner and then share the team’s answers with the rest of the class. Note that, for a difficult concept like the confidence interval, just one such question may not suffice! Asking questions similar to this one, but in a different context, is often required to help students see the issues at hand.

Other examples of warm-up questions, and the misconceptions they typically elicit, are given in Appendix C.

In addition to correcting mistakes before they become ingrained, there are other advantages for the instructor, at least in our experience. Students often start homework earlier, particularly on days when work is done in class. Fewer students turn in late homework or have last minute questions about homework on the due date. Furthermore, students engage with one another (and with the instructor) in class. Even in a large lecture class, students will often have names and contact information for classmates, which means that students will contact one another when they have questions, thus cutting down on instructor load and contributing to camaraderie in the classroom.

There are several disadvantages as well. One may have already surmised that a major disadvantage is time (Rhem 2010; Berrett 2012). Preparing warm-up questions, reviewing the answers, assigning grades to such questions, and creating in-class activities is time-consuming. As with most courses, the initial run with a different method of teaching takes much
more time than any subsequent runs. It also helps to approach the design and operation of the flipped classroom courses as a collective, rather than individual, effort. For example, several instructors can take turns preparing or updating the videos and other instructional aids to be used by the whole department for several years. This redistributes the burden of the preparation of the materials, while also involving each instructor into reviewing all materials regularly. It also prevents attrition of understanding on the part of the instructor that gradually happens if the instructor uses someone else’s videos for a long time.

With regard to the additional task of reviewing students’ answers to the warm-ups before the class, in our experience, only a random sample of about 20 answers is necessary to identify the misconceptions that most students will have (Stokes and McGee 2014). Therefore, the instructor does not need to read the entirety of the students’ responses within the limited time before class; however, eventually all responses will have to be graded.

Another disadvantage is a feeling of loss of control over the syllabus. Since it is important to adjust to the pace of the class, fitting in all of the material can be difficult, and some instructors find that they get behind with respect to the course syllabus. For an upper level course, this is perhaps not as much of an issue as it would be for an introductory course, where students are expected to learn certain material before progressing to the next level. Well-designed videos for out-of-class viewing may help to cover all material. Some typical activities done in the class, such as reviewing common solution methods or solving a sample problem on the whiteboard, can be relegated to a video to be viewed outside of class.

Finally, some students do not appreciate the extra work involved, and such dissatisfaction shows on course evaluations. Students complain of “too much work,” particularly in general education courses (Guerin 2010); this tendency broadly affects classes with active learning (Kober 2015). In our own experience, a student has written to our department chair complaining that his instructor was “not really teaching” by having students watch videos outside of class and do activities in class. The instructor was able to counteract the student’s arguments by showing to the department chair the evidence for increased student learning with the flipped classroom method in general, and further evidence of improved assessment outcomes in the particular course at the end of the semester (see the next section). However, we have also had positive comments on evaluations, such as the following:

“I didn’t mind getting up for an 8 a.m. class (as much) because there was always a lively discussion.”

Small steps toward this method can have big benefits in terms of evaluations and will not take up very much time on the part of the instructor. For example, one of the authors of this article began by recording videos of certain homework problems that students seemed to have difficulty doing. The students very much appreciated this since they could review the problem multiple times and benefit from the instructor’s clear, step-by-step explanation of the steps. Recording of videos detailing the use of software has been very successful in an upper level statistics course, and students noted the extra time the professor took to create the videos on end of semester evaluations. Students appreciated being able to refer to the video when they were stuck on an assignment requiring software. Another good place to start would be as an exam review. The instructor could record videos, or post review problems prior to a review session (or both), and have students in the review session discuss answers to the problems. Any one of these items would allow an instructor to get used to the idea of video recording while producing material that students would find helpful.

With all of the above in mind, untenured faculty who wish to use this method should consider two things very carefully: the impact of the extra time on their research (if research is a criterion for promotion and tenure), and the impact on student evaluations. Faculty need to be proactive when considering a change to the JiTT method or any active learning method. Sitting down with the department chair and explaining the concept of JiTT, its benefits, and its detractions, will allow the department chair to make appropriate arguments in favor of the faculty member when the time comes.

6. Assessment

JiTT has been used most consistently in an introductory statistics course with an enrollment of 60–80 per section. Approximately 60% of the students taking the course are business majors. Other majors represented (in decreasing order of representation) are economics, finance, advertising, psychology, journalism, mathematics, marketing, applied physiology, history, engineering, and art. Most of the students are first-years and sophomores. The course is one of six gatekeeper courses for entry into the increasingly competitive business major in our university. Students must have at least a 3.30 GPA in the six-course subset and at least a 3.30 GPA overall to gain entry into the business school.

Proper experimental designs for assessment of teaching innovations are difficult to implement in university settings due to the autonomy of both teachers and students. Though the authors have assessed the performance of our students under the standard and JiTT method, there was no randomization of students to our classrooms (when there were multiple sections), nor did we offer contemporaneous sections of our own sections with different treatments.

One of us has taught the same introductory business statistics course at least once per year for six consecutive academic years (2010–2011 through 2014–2015). Student learning in the course was assessed with the ARTIST test (delMas, Garfield, and Chance 2003) in the first and last lab of the semester each term. The JiTT method was partially implemented in this course in Spring 2013 and fully implemented in Fall 2013, and the course was taught using that method thereafter. The mean pretest and posttest scores from these sections are shown in Figure 1, along with the difference in the two. The error bars represent 95% confidence intervals.

The mean of each test was computed from the scores of students who were registered in the course, completed the assessment in the lab the day it was made available, and for whom more than 5 min elapsed between the time they opened and submitted the test. The figure shows that the pretest scores have not increased over the time period, but that the mean of
the posttest and the change in means both appear to have increased.

At best, the assessment data from these sections can be regarded as having come from a natural experiment. As with any such observational study, there are factors besides the new instructional method, such as changes in the student characteristics or behaviors that could affect the outcomes. Two concerns for this study are:

1. The average SAT scores for entering classes in the university increased from 1269 to 1308 between 2011 and 2015. This could cause the scores to rise over time due solely to more able students.
2. The proportion of students initially enrolled in the course and who dropped increased over time (See Table 2). The cause of this behavior is not clear; it could be either because more students disliked the course format, or because a greater number of students were exhibiting "grade management" behavior. If either of these reasons is the true cause, it would be expected that the average posttest score would increase due to elimination of a greater fraction of poorly performing or unmotivated students by the end of the term.

Besides these problems, there are also a substantial number of missing scores due to absenteeism. Missing pretest scores are often due to late registration or other standard beginning-of-term confusion. For example, in some spring semesters, the first day of class was a lab day before the MLK holiday and so missing data rates for the pretest were especially large. Though these missing data may plausibly be considered random, the missing posttest scores are more concerning. Students were incentivized to take the posttest (by providing extra credit for scoring above the national

---

**Figure 1.** Mean ARTIST Test scores and 95% confidence intervals in an introductory business statistics course from spring 2011 until spring 2015.

**Table 2.** Class enrollment, drop counts, and assessment score counts.

| Semester     | # Enrolled | # Dropped | # Taking pretest and posttest | # Taking pretest only | # Taking posttest only |
|--------------|------------|-----------|-------------------------------|-----------------------|-----------------------|
| Spring 2011  | 63         | 6         | 27                            | 3                     | 22                    |
| Fall 2011    | 34         | 2         | 24                            | 9                     | 0                     |
| Spring 2012  | 45         | 5         | 34                            | 9                     | 0                     |
| Fall 2012    | 51         | 7         | 35                            | 2                     | 0                     |
| Spring 2013  | 89         | 15        | 55                            | 12                    | 6                     |
| Fall 2013    | 78         | 12        | 49                            | 20                    | 1                     |
| Spring 2014  | 99         | 17        | 48                            | 19                    | 22                    |
| Spring 2015  | 78         | 22        | 32                            | 23                    | 9                     |

**Table 3.** Estimates of mean pretest, posttest, and change scores.

| Semester     | Pretest mean | Posttest mean | Mean of the pretest–posttest difference |
|--------------|--------------|---------------|----------------------------------------|
| Spring 2011  | 47.9 (1.3)   | 56.8 (1.4)    | 8.9 (1.7)                              |
| Fall 2011    | 47.3 (1.5)   | 59.0 (2.7)    | 11.6 (2.4)                             |
| Spring 2012  | 45.9 (1.8)   | 55.7 (2.1)    | 9.8 (1.7)                              |
| Fall 2012    | 46.5 (2.0)   | 57.8 (2.0)    | 11.2 (1.8)                             |
| Spring 2013  | 43.1 (1.3)   | 54.8 (1.3)    | 11.7 (1.5)                             |
| Fall 2013    | 46.3 (1.2)   | 67.1 (2.0)    | 20.8 (2.0)                             |
| Spring 2014  | 43.2 (1.1)   | 61.7 (1.2)    | 18.5 (1.0)                             |
| Spring 2015  | 40.0 (1.7)   | 59.3 (2.5)    | 19.3 (2.6)                             |
average), but not all students did so. A plausible argument could be made that either high performing (because they perceive they do not need the extra credit) or low performing (because they perceive they would not earn the extra credit) students are more likely to have missing posttest data. Table 2 shows the number of missing values for each test in each semester.

To address the problem of missing data from both drops and absenteeism, we used a multiple imputation procedure (with $m = 5$ imputations) to impute the missing score for all students who took only the pretest or only the posttest. If the students missing either test were the worse (or better) performing students, the imputation could help account for this assuming the student’s knowledge was reflected in the available score. Indeed the correlation between pretest and posttest for students taking both averaged 0.45 over the eight semesters.

The method used for imputation of the scores was the EM algorithm of Schafer (1997). The method was implemented using SAS PROC MI and analyzed with PROC MIANALYZE. The data were analyzed as a matched pairs design for each replication, and the results combined. Table 3 shows the estimates of the mean pretest, posttest, and change in the scores from this analysis, along with the standard errors of the estimates, which reflect both sampling and imputation variance. These data are plotted, using 95% confidence intervals, instead of standard errors, in Figure 1.

The data show the difference in mean ARTIST posttest and pretest scores is larger after the JiTT method was fully implemented. The Spring 2013 semester does not show evidence of an improvement in score, even though some components of the JiTT approach were used. Specifically, videos (that were not prepared by us) were provided to the students and they were asked to prepare before class. Some class time was used for problem solving. However, the warm-up exercises were not used until the following semester. We believe that both the student’s articulation of their thinking and our ability to offer immediate error correction contribute to the greater effectiveness.

7. Discussion

Since identifying typical misconceptions and designing effective questions inevitably proceeds through trial-and-error, creation of shared instructional aids and sizable databases of good questions for JiTT benefits the whole community. Physicists were instrumental in experimenting with JiTT techniques and determining what works (Novak et al. 1999). They also began a website containing resources for practitioners of the JiTT known as the Just-in-Time Teaching Digital Library (jittdl.physics.iupui.edu/jit/DL/dsp_home.php). This resource has now expanded with user input to include pedagogical aids and submitted curricular materials for instructors of courses in Biology, Chemistry, Environmental Engineering, Geoscience, Mathematics, and Physics. The contributors of materials within each discipline classify them as to the resource type (e.g., class interaction, demonstration), subject and description (free-form explanation), audience level (Secondary, Undergraduate, Graduate, Professional/Clinical), and content orientation (Subject Major, Service, Multidisciplinary).

There are useful general resources for teachers of statistics courses in the Digital Library. However, since there is not yet a specific statistics disciplinary area, there are few actual class materials for topics in statistics. We have the rudimentary beginnings of such an archive on GitHub (https://github.com/MonnieMcGee/JiTTQuestions). Our experience is that, besides the materials themselves, access to hands-on training in the method is invaluable. One of us was trained in a workshop sponsored by the American Physics Society and American Association of Physics Teachers. Participants in these two-day long workshops are trained in know-how aspects of the instructional technique that make tangible difference in the students’ learning outcomes and satisfaction. The other two of us attended training and a demonstration of the JiTT approach provided by our University’s teacher training center, which was supported by our physics colleagues.

Just-in-Time Teaching and active learning can be used at a variety of course levels and in a variety of ways. The examples and outlines of our classrooms are simply examples of ways to proceed. The guidelines in Table 1 are very general. The details of the day-to-day handling of the course, the warm-up questions, assigning credit for questions, and use of out-of-class materials are purely up to the instructor. Methods of assessment, such as the ARTIST exercises (delMas, Garfield, and Chance 2003) and shared test questions across different sections of the same course, can help faculty visualize student progress as instructors (and students) become more familiar with JiTT.

Acknowledgments

We thank Andrew Gavrin for illuminating communications about the relation between JiTT and inverted classroom methods.

Supplementary Materials

Supplemental data for this article can be accessed on the publisher’s website.

References

Berrett, D. (2012), “How Flipping the Classroom Can Improve the Traditional Lecture,” The Chronicle of Higher Education, February 19, 2012.

Bransford, J. D., Brown, A. L., and Cacking, R. R. (eds.) (2000), How People Learn: Brain, Mind, Experience, and School. Washington, DC: National Research Council, The National Academies Press. Available at http://www.nap.edu/download.php?record_id=9853.

Camp, M. E., Middendorf, J. and Sullivan, C. (2010), “Using Just in Time Teaching to Motivate Student Learning.” In Just in Time Teaching, ed. Scott Simpkins and Mark Maier, 25 - 38. Sterling, VA: Stylus.

Carlson, K. A., and Windquist, J. R. (2011), Evaluating an Active Learning Approach to Teaching Introductory Statistics: A Classroom Workbook Approach,” Journal of Statistics Education 19(1).

Chiesi, F., and Primi, C. (2009), “Assessing Statistics Attitudes Among College Students: Psychometric Properties of the Italian Version of the Survey of Attitudes Toward Statistics (SATS),” Learning and Individual Differences, 2, 309–313.
delMas, R., Garfield, J., and Chance, B. L. (2003), "An Online Resource for the Assessment of Instructional Outcomes," *Proceedings of the Joint Statistical Meetings*, San Francisco.

Gavrin, A. (2015) Private communication.

Guerrin, L. A. (2010), "Using Just in Time Teaching in the Geosciences," *Just in Time Teaching: Across the Disciplines, Across the Academy*, Sterling, VA: Stylus Publishing, pp. 101–116.

Hake, R. (1998), "Interactive Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses," *American Journal of Physics* 66, 64–74.

Jensen, J. L., Kummer, T. A., and Godoy, P. D. (2015), "Improvements from a Flipped Classroom May Simply Be the Fruits of Active Learning," *CBE—Life Sciences Education*, 14, Spring 2015, pp. 1–12.

Karpicke, J. D., and Roediger, H. L. (2007). "Expanding Retrieval Practice Promotes Short-Term Retention, but Equally Spaced Retrieval Enhances Long-Term Retention," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33, 704–719.

Kober, N. (ed.) (2015), *Reaching Students: What Research Says About Effective Instruction in Undergraduate Science and Engineering*. National Research Council, Washington, DC: The National Academies Press. Available at //www.nap.edu/download.php?record_id=18687.

Lage, M. J., Platt, G. J., and Treglia, M. (2000). "Inverting the Classroom: A Gateway to Creating an Inclusive Learning Environment," *The Journal of Economic Education*, 31, 30–43.

Lyman, F. (1981). "The Responsive Classroom Discussion: The Inclusion of All Students," *Mainstreaming Digest*, College Park, MD: University of Maryland.

Marrs, K. A., Blake, R., and Gavrin, A. (2003), "Use of Warm Up Exercises in Just in Time Teaching: Determining Students’ Prior Knowledge and Misconceptions in Biology, Chemistry, and Physics," *Journal of College Science Teaching*, 33(1), 42–47.

Moore, D. S., McCabe, G. P., Alwan, L. C., Craig, B. A., and Duckworth, W. M. (2011), *The Practice of Statistics for Business and Economics* (3rd Ed.), New York: W.H. Freeman and Company.

Novak, G., and Patterson, E. (2010), "An Introduction to Just in Time Teaching (JiTT)," *Just in Time Teaching: Across the Disciplines, Across the Academy*, Sterling, VA: Stylus Publishing, pp. 3–24.

Novak, G., Patterson, E., Gavrin, A., and Christian, W. (1999), *Just-in-Time-Teaching: Blending Active Learning with Web Technology*, New York: Prentice Hall.

Rhem, J. (2010), Foreword to *Just in Time Teaching: Across the Disciplines, Across the Academy*, Sterling, VA: Stylus Publishing.

Rossmann, A. J., and Chance, B. L. (2008), *Workshop Statistics: Discovery with Data* (3rd ed.), New York: Wiley.

Schaefer, J. (1997), Analysis of Incomplete Multivariate Data. Boca Raton: Chapman and Hall, 260–264.

Schneiter, K. (2015), Statlets. Available at http://www.math.usu.edu/~schneit/CTIS/.

Steen, L. A. (ed.) (1992), *Heeding the Call for Change: Suggestions for Curricular Action*, Washington, DC: Mathematical Association of America.

Stokes, L., and McGee, M. (2013), *Adventures in the Flipped Classroom*, Department of Statistics Colloquium Series, Southern Methodist University, Dallas, Texas.

West, R. W. (2015), "Web's Page of Statistical Applets," available at http://www4.stat.ncsu.edu/~west/.

Windquist, J. R., and Carlson, K. A. (2014), "Flipped Statistics Class Results: Better Performance Than Lecture Over One Year Later," *Journal of Statistics Education*, 22. Available at www.amstat.org/publications/jse/v22n3/windquist.pdf

Zappe, S., Leicht, R., Messner, J., Litzinger, T., and Lee, H. (2009), "Flipping the Classroom to Explore Active Learning in a Large Undergraduate Course," *Proceedings, American Society for Engineering Education Annual Conference and Exhibition*, Austin, TX: ASEE.