Just-in-Time-Teaching (JiTT) Improves Students' Performance in Classes – Adaptation of JiTT in Four Geography Courses

Wei Luo — Department of Geography, Northern Illinois University, DeKalb, IL 60115

ABSTRACT

By constructing "treatment" and "control" groups from existing grade rolls of courses taught with Just-in-Time Teaching (JiTT) method and traditional lecturing in recent years, this study demonstrates that JiTT improves student performance in classes. First, two groups of students were selected from each of the four different courses taught with JiTT such that the "treatment" group regularly completed warm-up exercises whereas the "control" group did not, but both groups had similar first exam scores. Statistical t-tests show that the "treatment" group had significantly higher overall course grade (by one letter grade on average) than the "control" group. Second, statistical t-test also indicates that students' relative improvement between the last and the first exams in a general education course using JiTT method is significantly higher than that of the same course using traditional lecturing prior to adopting JiTT method. In addition, qualitative surveys of students demonstrate that the majority of the students thought that doing Warm-Up exercises helped their learning, because Warm-Ups made them think about the material before going into lecture and thus made them more aware of and keen to the topics that would be discussed. These consistent results in four different courses over five semesters indicate that JiTT improves students' performance.

INTRODUCTION

Just-in-Time Teaching (JiTT), first developed for introductory physics courses and now gaining wider acceptance, is a teaching and learning strategy based on the interaction between web-based study assignments and an active learner classroom (Novak et al., 1999). The web-based study assignments, also called Warm-Up exercises, are carefully constructed and are due shortly before class. The students' responses are then used to inform and modify classroom activities in the upcoming class (Novak et al., 1999). By examining student responses to Warm-Up questions submitted electronically "just-in-time" before each class, instructors can determine the level of understanding, prior knowledge, and misconceptions that students bring to class. Selected student Warm-Up responses are shown and discussed in class, allowing classroom time to be spent addressing misconceptions and relating to prior knowledge while discussing that day's course content. Exactly how the classroom time is spent in the active learner classroom component of the JiTT depends on a variety of issues such as class size, course contents, classroom facilities, and student and instructor personalities. It can range from pretty much the same as traditional lectures, except now more focused toward addressing students' misconceptions, to mini-lectures interspersed with demonstrations, classroom discussion, worksheet exercises, hands-on mini-labs, or interactive activities using Classroom Performance System (commonly know as "clicker" technology). Regardless,

the common key is that the classroom component, whether interactive lecture or student activities, is informed by an analysis of various student responses. The heart of JiTT is the short time frame "feedback loop" formed by the students' outside-of-class preparation that fundamentally affects what happens during the subsequent in-class time together (Novak et al., 1999).

JiTT has become popular recently and been adopted in a variety of fields beyond physics, such as biology (Marrs et al., 2003; Marrs and Novak, 2004), economics (Simkins and Maier, 2004), sociology (Howard, 2004), and computer science (Gavrin et al., 2004). Surprisingly, there are very few quantitative, objective studies (as opposed to qualitative, subjective surveys based on students' self-assessment such as Marrs and Novak (2004)) that demonstrate the benefit of JiTT in terms of students' achievement and performance in classes. Linneman and Plake (2006) compared two sections of a large enrollment introductory Geology course, taught by two different instructors, one section using JiTT method and the other using traditional lecture and they found no statistically significant difference between the two sections.

The purpose of this paper is to test if JiTT helps improve students' performance in class from statistical analysis of "treatment" and "control" groups constructed from the data accumulated over recent years (after course completion). It is hoped to demonstrate, quantitatively and systematically, that JiTT improves student performance in all of the four different courses over five semesters.

BRIEF DESCRIPTION OF THE FOUR COURSES AND ADAPTATION OF JITT

The author first learned JiTT in a National Science Foundation (NSF) Course, Curriculum, and Laboratory Improvement (CCLI) program conference in 2004 and has adapted JiTT in four of the courses over the past two and half years (one of them repeated). Below is a brief description of each course.

- GEOG 101 (3 credits): Survey of Physical Geography, a general education course that satisfies the science requirement and serves the whole university and an entry course to most geography courses. This course introduces students to natural environmental systems, with emphasis placed on the interrelationship and interaction among different subsystems such as hydrology, soils, landforms, vegetation, and geologic materials and the processes involved; the impacts of human activities on natural environmental systems are also discussed. There is a separate companion 1 credit lab course GEOG 102. This is a large enrollment class (n > 200), in which students are assessed on the basis of four exams.
- GEOG 303 (3 credits): Water Resources and the Environment, an elective course with no prerequisite. This course aims to provide students with a broader understanding of water as a natural resource and its importance to our lives and the Earth's complex

environment. It covers physical properties of water, the hydrologic cycle, and social, economical, and environmental issues related to water distribution, water usage, and water management. Students are assessed on the basis of three exams, one term paper and other assignments.

- GEOG 359 (3 credits): Introduction to Geographic Information Systems (GIS), a required course for the GIS certificate. This course introduces the basic theory and concepts of GIS and offers hands-on experience of working with major commercial GIS software packages. There is a lab component within this course. Students are assessed based on two exams, ten labs, and other assignments.
- GEOG 460 (3 credits): Remote Sensing of the Environment, a required course for the GIS certificate. This course is co-taught by two professors. The author covers the basic principles, concepts, theory of remote sensing, with an emphasis on image processing and interpretation. The other professor's part emphasizes application of remote sensing in meteorology. JiTT was only used in the author's part of the course. The rest of discussion regarding this course only applies to this part of the course. There is also a lab component within this course. Students are assessed based on two exams, labs, and other assignments.

Following the JiTT pedagogy (Novak et al., 1999), Warm-Up exercises were designed based on the main points or concepts to be covered in the lecture. The format of the lecture was not drastically changed, except that now with JiTT the main points or concepts were introduced by first presenting selected students' responses to Warm-Up questions. Since students have already tried to answer those questions and have thought about them to some degree, they should thus be more attentive to the lecture and better able to retain the correct formation. From the instructor's perspective, he or she can be more focused on addressing students' misconceptions shown in their answers by adjusting the time and depth of discussion on different concepts/topics.

For this study, all Warm-Up exercises were implemented in the BlackBoard Learning System, a popular on-line course management system adopted by Northern Illinois University, and were worth 5% to 10% of the course grade. They were due 1 or 2 hours before class. After the due time, the Warm-Ups would automatically disappear from the students' view in BlackBoard. Each exercise was composed of 2-4 short answer questions. This type of question was chosen because it allowed students to express their ideas more freely and the instructor could get more information from students than from multiple choice questions. Students were told in the first class that their Warm-Ups would be graded based on their effort and not on the correctness of their answer. In other words, if they made their best effort to answer the questions, they would get the full credit even if the answer was wrong. For large enrollment classes, they were only graded for completeness (otherwise it is not feasible to do). The instructor (author) went through students' answers before lecture and incorporated selected, representative answers anonymously in the appropriate place of the lecture PowerPoint presentation. (Due to the time limit, the actual grading was done after the lecture.) During the lecture, instead of passively listening, students participated in a guided discussion that began with their own preliminary understanding of the material.

The Warm-Up questions were designed to be thought provocative or in a way that students could easily relate to their prior knowledge or everyday life experience. Here are some example Warm-Up questions:

For the Physical Geography course, when introducing the concept of weathering and groundwater, the following Warm-Up questions were developed:

1. The words carved on stone walls of ancient buildings often became blurred after hundreds of years. What do you think caused this phenomenon?
2. Most rivers have water flowing in them even during droughts (i.e., no rainfall at all). Where do you think the water is coming from?

In the Water Resources course, when talking about the properties of water, the following Warm-Up questions were devised:

1. In winter people often put salt on driveway to melt ice. Why can salt melt ice?
2. In a sunny hot summer day, if you put a cup of sand and a cup of water under the Sun for a long time, which one will be hotter, the sand or the water? Why?

In the Introduction to GIS course, when discussing Map Projection and Map scale, the following questions were designed:

1. The following map shows the flight path between New York (JFK) and Los Angeles (LAX). Why is the path curved?
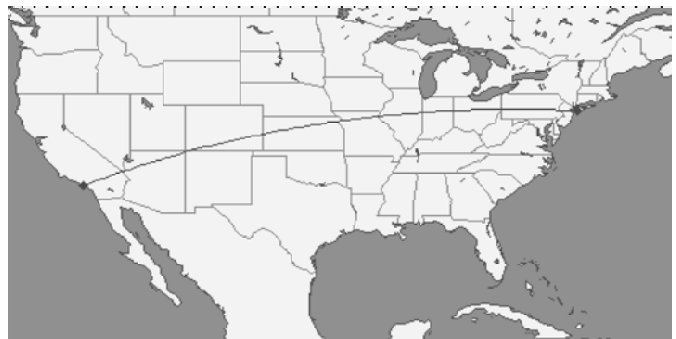
2. Compare the two maps below, what are the differences between them?

In the Remote Sensing course, when covering the topic of supervised classification and unsupervised classification, the following Warm-Up questions were asked:

1. In the picture below, how would you group the objects based on shape? How would you group them based on color?

2. Suppose you have a Martian visitor who has never seen an apple, a banana, or a pear. How would you teach him to identify these three objects? Now to test if he really understands, you give him an apple (which is slightly different from the one you used to teach him). How do you think he would tell this is an apple?

In addition to these questions that acted as prompts to naturally introduce the main points/concepts of that day's lecture, the instructor also added one question related to the previous lecture to test students understanding of topics covered before. This gave students a timely reinforcement of the concepts covered previously and also provided the instructor with some much needed quick feedback on how the students grasped the concepts and allowed the instructor to make remedies if necessary. Periodically after each chapter or logical unit, online practice quizzes were offered on BlackBoard that can be repeated as many times as needed and were not part of the students' course grades. In this way, students could put their understanding of the material to test in a less "threatening" environment. Finally, the instructor always had an optional question in each Warm-Up exercise inviting students to give any comments/suggestions/complaints or ask any questions. This gave the instructor an additional means of getting students feedback in a timely fashion.

RESULTS

Statistical t test - There are many factors that may influence students' overall performance in class, e.g., students' prior knowledge, students' effort in studying, the difficulty level of the exam, the teaching style of the instructor, the lecture material used, or even the classroom environment, etc. To establish causal relationship between the use of JiTT method in class (i.e., the treatment) and its effectiveness in enhancing students' overall performance (i.e., the response), one needs to conduct a treatment and control experiment by teaching one class with JiTT (the treatment group) and the other with traditional method (the control group) while keeping all other factors the same for both groups. Ideally, only when all the other factors (except for treatment or teaching method) are kept the same can the difference exhibited in the response (overall students' performance) between the two groups be reasonably attributed to the treatment (teaching method) (Reiter, 2000 and references therein). In reality, this can be accomplished by randomly assigning students to each group. Because of the randomness, these other factors (both known and those not yet to be identified) between the two groups can be considered approximately the same or closely balanced (Reiter, 2000).

However, in a public education environment, conducting treatment and control experiment with randomization may be considered unethical because as educators we are obligated to offer the best education to all students using best practices available. So this study will not be a standard treatment and control experiment, in which students were randomly assigned to treatment and control groups at the beginning, but rather an observational study (Reiter, 2000), in which "treatment" and "control" groups were constructed from existing grade rolls after the completion of courses by applying the principles of a standard treatment and control experiment as described above.

Difference Between Two Groups within Classes using JiTT - In order to ensure all other factors are as similar as possible, "treatment" and "control" groups were constructed from existing grade rolls of courses taught with the Just-in-Time Teaching (JiTT) method based on two criteria: (1) whether students regularly completed Warm-Up exercise, and (2) whether they had similar first exam scores. The "treatment" group regularly completed Warm-Up exercises whereas the "control" group did not, but both groups had similar first exam scores (i.e., they were at similar level early in the semester). Because the Warm-Up exercises were graded based on effort, the first criterion is implemented as follows: students whose Warm-Up scores were greater than 1 standard deviation above the mean belong to the "treatment" group; those with Warm-Up scores that were less than 1 standard deviation below the mean belong to the "control" group. The initial temporary "treatment" and "control" groups were formed after applying the first criterion. Statistical t-tests were conducted on the two temporary "treatment" and "control" groups to see if they had statistically different first exam scores (second criterion). If they did, students with exceptionally high or low first exam scores were removed from the groups. This process was repeated until the two groups had similar first exam scores and the final "treatment" and "control" groups were formed. The sample sizes of the final groups that satisfy both criteria (n_0 in Table 1) were reduced (by half in one case) from those of the initial groups that satisfy only the first criterion (n in Table 1). The difference of the first exam scores between the final groups was statistically insignificant (Table 1), i.e., they had similar first exam scores. Thus the final "treatment" and "control" groups constructed this way ensured that students in both groups were approximately at the same level in the beginning part of the course.

Statistical t-tests were then conducted between the two final groups on their overall performance (overall score minus the Warm-Up component, hereafter referred to as overall score) in each course using SAS software. If JiTT improves students' class performance, then one

	Enroll-ment	Group	n_0	First Exam					Overall Score			
				n	mean	Std dev	t	$p > \lvert t \rvert$	mean	Std dev	t	$p > t$
Geog 101 Fall 05	300	Treatment	61	30	74.2	11.1	*-0.47*	*0.64*	65.9	15.3	-2.52	*0.007*
		Control	30	18	75.7	10.2			76.2	10.4		
Geog 101 Spr 07	228	Treatment	29	24	71.7	13.1	-0.52	0.60	61.2	17.7	-2.85	0.003
		Control	38	24	73.3	7.0			72.4	7.5		
Geog 303 Fall 04	25	Treatment	4	3	72.0	14.2	*-0.81*	*0.46*	75.0	9.6	-2.72	0.021
		Control	5	4	78.5	7.1			88.9	3.7		
Geog 359 Fall 06	62	Treatment	11	11	70.3	12.7	*-1.03*	*0.31*	71.9	14.3	-1.98	0.034
		Control	20	9	75.4	8.2			81.4	6.3		
Geog 460 Fall 06	56	Treatment	7	7	61.7	13.7	0.03	0.97	66.6	3.3	-2.09	0.038
		Control	28	15	61.5	6.8			73.6	5.5		

Table 1 Statistical t-test results between two groups within classes using JiTT. Note: Italicized numbers indicate t-test based on equal variance; the rest are based on unequal variance. (All results obtained in SAS). n_0 is the number of samples in the initial groups that satisfies the first criterion; n is the number of samples in the final groups that satisfies both criteria.

	n	mean	Standard deviation	t	$p > t$
Non-JiTT(03, 04)	100	2.747	14.762	-3.07	0.0012
JiTT (05, 07)	100	8.570	11.908		

Table 2 Statistical t-test on improvement from exam 1 to exam 4 between Non-JiTT and JiTT classes based on 100 random samples.

would expect the "treatment" group to have a higher overall score at the end of the course than that of the "control" group. So the null hypothesis (H_0) is that the means of overall scores of the "control" group ($\mu_{control}$) and the "treatment" group ($\mu_{treatment}$) are the same: H_0: $\mu_{control} = \mu_{treatment}$ and the alternative hypothesis (H_1) is that the overall mean score of the "treatment" group ($\mu_{treatment}$) will be higher than that of the "control" group ($\mu_{control}$): H_1: $\mu_{control} < \mu_{treatment}$. Results are shown in Table 1 and indicate that for all of the courses, the means of overall scores in "Treatment" Groups (students who completed Warm-Ups routinely) are significantly higher than those in "Control" Groups (students who did not regularly complete Warm-Ups). Thus H_0 can be rejected and H_1 accepted at the standard 95% confidence level. The difference of the overall score between the two groups ranges from 7.0 to 13.9 with an average of 10.4 and standard deviation of 2.5, or roughly one letter grade. Please note that the differences on their first exams were statistically insignificant, i.e., these two groups can be considered at the similar levels in the early part of the course. In addition, they were taught by the same instructor, using the same material, in the same classroom, and tested on the same exams. The only obvious difference between the two groups is their Warm-Up scores, which reflect whether or not they have regularly completed the Warm-Up exercises. Thus the results suggest that Warm-Up exercises do help improve students' overall performance in classes, by roughly one letter grade on average.

Difference Between JiTT and Non-JiTT Classes - One additional way to test whether Warm-Up exercises enhances students' class performance is to compare the students' improvement from the first to last exams in classes taught with JiTT and that in the same classes taught with traditional method (before the adoption of JiTT). For this purpose, GEOG 101 is chosen as an example because it is a large general education course (thus a large population size to draw random samples). This course was taught in the spring of 2003 and 2004 using traditional lecturing method, and in spring of 2005 and 2007 using JiTT. Since the exams do change over the years, the grades may not be directly comparable. However, each student's improvement over the course of a semester (as represented by the difference between the last exam and the first exam) is comparable between different semesters of the same course. (In general, students' last exam grades are higher than their first exam as they learn more and are more experienced with the exam format. At least this is true for GEOG 101.) To further minimize the effect of other factors, such as student's prior knowledge, difference in exam questions, etc., that might also contribute to the difference in scores, the following data processing is done: (1) the scores were normalized to between 0 and 100 (by dividing each student's exam score by the maximum of that exam in class); (2) 50 students were randomly selected from each course so that the non-JiTT group (2003 and 2004) and JiTT (2005 and 2007) group each had 100 samples; (3) statistical t-test was conducted on the improvement from exam 1 to exam 4 (i.e., exam 4 - exam 1) between the non-JiTT classes (control group) and JiTT classes (treatment group). The null hypothesis (H_0) is that the means of last and first exam differences between non-JiTT (non-JiTT) and JiTT (JiTT) classes are the same: H_0: non-JiTT = JiTT and the alternative hypothesis is that mean of last and first exam differences for JiTT classes (JiTT) will be higher than that for non-JiTT classes (non-JiTT): H_1: non-JiTT < JiTT. The result (see table 2) indicates that students' improvement over the semester in JiTT classes are significantly higher (by almost 6 points) than that of the non-JiTT classes. Thus, H_0 can be

What do you think about the effectiveness of Warm-up exercises in helping you learn?	Geog 359 (N = 45)	Geog 460 (N = 38)
They don't help me understand the lecture material at all. I don't like doing them.	13.3%	2.62%
They help me understand the lecture material a little.	44.4%	57.9%
The help me understand the lecture material a lot.	33.3%	31.6%
They help me understand the lecture material greatly. I enjoy doing them.	8.9%	7.9%
Unanswered.	0%	0%

Table 3 Survey of students' opinion about Warm-Up exercises in Fall of 2006 (GEOG359 and GEOG460) (not all students participated the survey).

	Statement 1	Statement 2
Strongly agree	20.62%	21.13
Agree	48.45%	48.97%
Neither agree nor disagree	17.53%	11.34%
Disagree	12.89%	11.86%
Strongly disagree	0.52%	5.57%
Not applicable	0%	1.03%
Not answered	0%	0%

Table 4 Survey of students opinion about Warm-Up exercise in Spring of 2007 (GEOG 101) (N=194, not all students completed the survey). Statement 1: The Warm-Up exercises help prepare my mind for the lecture and thus enhance my learning of the lecture material. Statement 2: I often read textbook, lecture notes, or look up things when I answer the Warm-Up questions.

rejected and H$_1$ accepted at 99% confidence level (Table 2).

Student Feedback - In the fall of 2006, a simple anonymous survey was conducted in the middle of the semester in GEOG 359 and GEOG 460 to get students' opinion on the effectiveness of Warm-Up exercise in help with their learning. The first question was a multiple choice question and the result is summarized in Table 3. An overwhelming majority of those who participated in the survey thought that Warm-Up exercises helped their learning at least a little (87%, 97%). About 40% of the participating students thought that Warm-Ups helped their learning a lot and enjoyed doing them. In the spring of 2007, a similar survey was also conducted in GEOG 101 at the end of the semester. The result is summarized in Table 4. Nearly 70% of the students agreed or strongly agreed that Warm-Up exercises helped prepare their minds for the lectures and, thus, enhanced their learning of the lecture material. Interestingly, 70% of the students also agreed or strongly agreed that they often read textbook, lecture notes, or looked up things when they answered the Warm-Up questions (even though this was not required). In both surveys, an open ended question was included asking them to provide any comments/suggestions (either positive or negative) about Warm-Ups and how to make them better in the future. Most students liked Warm-Up exercises in that the exercises made them think about the material before going into lectures and thus made them more aware of and keen to the topics that would be discussed during lectures. Students also liked the Warm-Ups because they did not take long to complete and could give students some "easy" points. Students in the more technically oriented courses (GEOG 359 and GEOG 460) found that the Warm-Up questions that related more difficult concepts with their every day life experience were extremely helpful. The complaints were generally about forgetting to complete the exercises, sometimes not getting the correct answers (even though they were discussed in class) and computer or Internet connection problems. Some suggested to reduce the number of Warm-Up exercises to once per week (instead of once per class period) but may have more questions in each Warm-Up exercise. This suggestion was implemented in 2007 and students were happy about that. They also suggested that the correct answers to Warm-Up questions be given on BlackBoard after each Warm-Up is due and use more of them in exams.

DISCUSSION AND CONCLUDING REMARKS

This observational study constructed the "treatment" and "control" groups from completed courses taught with JiTT method, based not only on whether students completed the Warm-Ups regularly (represented by high and low Warm-Up scores), but also on whether they had similar first exam scores. The second criterion generally ensures that the two groups had approximately the same other factors that might also influence students' performance such as students' prior knowledge and being responsible in studying course materials. Pre-informed students may be more likely to complete their Warm-Ups (because they are not intimidated by the topics) and pre-informed students also do better on their exams (independent of instruction). Responsible students tend to complete Warm-Ups regularly and responsible students also study for their exams thus do better in exams. The similar scores on the first exams indicate that the two groups were at similar levels at that point in class, in addition to being taught by the same instructor, using same material and same exams. The overall performance of the "treatment" group were significantly better than the "control" group (roughly by one letter grade on average), which can reasonably be attributed to the only factor that was different: their Warm-Up scores, i.e., whether they regularly completed the Warm-Up exercises or not. In addition, the fact that the results are consistent in all four courses over five semesters (of which GEOG 101 is repeated in spring 2007) added more support to this claim, in contrast to the one semester study in Linneman and Plake (2006). Further, comparison of students' relative improvement during a semester from the first to last exams in GEOG 101 based on random sampling shows that students taught with JiTT method improve almost 6 points more than those taught with traditional lecturing prior to the adoption of JiTT and the difference is statistically significant.

In addition, qualitative surveys of students demonstrate that the majority of the students thought that doing Warm-Up exercises helped their learning, because Warm-Ups made them think about the material before going into lecture and thus made them more aware of and keen to the topics that would be discussed. The majority of the students also reported that they often read textbook, lecture notes, or looked up things when they answered the Warm-Up questions (even though this was not required).

Thus consistent results from both quantitative analyses of students' performance and qualitative feedback from students in different courses taught with JiTT method and traditional lecture over five semesters demonstrate that JiTT improves students' performance. The results of this study supports the previous findings of (Novak et al., 1999) that: (1) JiTT helps student to be better prepared for class, leading to overall better attentiveness and engagement with course concepts; (2) the short time frame feedback loop of JiTT helps instructors understand students thinking process prior to class, allowing faculty to make more effective and efficient use of precious class time; and (3) JiTT may also help change students' study habits by forcing them to be thinking about or even reading course material outside of the classroom on a continuous basis (see Table 4 statement 2 statistics).

ACKNOWLEDGEMENTS

REFERENCES

Gavrin, A., Watt, J. X., Marrs, K., andBlake, R. E., 2004, Just-in-Time Teaching (JiTT): Using the Web to Enhance Classroom Learning, Computers in Education Journal, v. 14, p. 51-60.

Howard, J. R., 2004, NOTES - Just-in-Time Teaching in Sociology or How I Convinced My Students to Actually Read the Assignment, Teaching Sociology, v. 32, p. 385-390.

Linneman, S., and Plake, T., 2006, Searching for the Difference: A Controlled Test of Just-in-Time Teaching for Large-Enrollment Introductory Geology Courses, Journal of Geoscience Education, v. 54, p. 18-25.

Marrs, K. A., Blake, R. E., and Gavrin, A. D., 2003, Web-Based Warm Up Exercises in Just-in-Time Teaching - Determining Students' Prior Knowledge and Misconceptions in Biology, Chemistry, and Physics, Journal of College Science Teaching, v. 33, p. 42-48.

Marrs, K. A., and Novak, G., 2004, Just-in-Time Teaching in Biology: Creating an Active Learner Classroom Using the Internet, Cell Biology Education, v. 3, p. 49-61.

Novak, G. M., Gavrin, A., Christian, W., and Patterson, E., 1999, Just-in-time teaching: blending active learning with web technology, Upper Saddle River, NJ: Prentice Hall, 188p.

Reiter, J., 2000, Using statistics to determine causal relationships, American Mathematical Monthly, v. 107, p. 24-32.

Simkins, S., and Maier, M., 2004, Using Just-in-Time Teaching Techniques in the Principles of Economics Course, Social Science Computer Review, v. 22, p. 444-456.