

## Online Simulations in the Classroom

Eric Cox  
Associate Professor  
Department of Political Science  
Texas Christian University  
[e.cox@tcu.edu](mailto:e.cox@tcu.edu)

*Prepared for Presentation in Buenos Aires, Argentina  
FLACSO-ISA Joint Conference  
July 23-25 2014*

Do simulations provide an effective method to improve student learning? This paper attempts to provide a partial answer to this question by examining results of teaching two sections of the same introduction to international politics course in the same semester, one of which used a traditional research paper and one of which used the online Statecraft simulation as a supplemental assignment.<sup>1</sup> The results presented in this paper suggest that while students engaging in the simulation do not perform any better or worse than students who completed a traditional research paper, the students completing the simulation were more satisfied than those who wrote the research paper even though the simulation required more actual work than the research paper.

This paper will proceed by providing a brief review of the literature concerning the use of simulations and active learning in the political science classroom, followed by a discussion of the teaching methods used in each of two sections of my international politics class. It will then briefly discuss the research method employed by this paper, followed by a presentation of results and the conclusion.

### ***Literature Review***

While it seems that active learning is growing in popularity, little research has been published in political science journals that demonstrates the effectiveness of active learning techniques in general or simulations in particular. Much of the existing literature focuses on a discussion of the technique used or the use of student self-evaluations to measure the impact of learning.

A wide variety of scholars have written about the potential for using simulations in the classroom to enhance learning. Most published articles regarding simulations provide advice on how to design and use a particular simulation in the classroom, as well as the author's insights as to how

---

<sup>1</sup> The Statecraft simulation can be found at <http://www.statecraftsim.com>. In the interest of full disclosure, I am currently working with Statecraft to develop a Model United Nations simulation. The research in this paper was conducted with the knowledge of Statecraft, but was in no way influenced by the company.

effective the simulation was, including reflections on how to improve the simulation. A quick survey of the literature focused on international relations simulations suggests that simulations can be used in a wide variety of capacities including the study of coalition building (Shellman 2001), the UN Security Council (Chasek 2005), multilateral negotiations regarding the environment and other topics (Schnurr, de Santo and Craig 2013; Crossley-Frolick 2010), game theory (Rothman 2012), national security processes (Garrison, Redd and Carter 2010), and inter-state diplomacy including roles for many actors (Shellman and Turan 2006; Butcher 2012; Shaw 2004).

Many of these studies do take seriously the merits of active learning and provide detailed reviews of the debate surrounding the efficacy of active learning techniques. For example, Taylor (2013) effectively compares a number of different types of simulations, weighing the positives and negatives of them as teaching techniques. Some of the studies do engage in substantial evaluation of the simulations used by using surveys of students to evaluate student perceptions of their learning as a result of their surveys. For example, Shellman and Turan (2006) provided a comprehensive survey to students measuring whether or not students believed their knowledge to have been enhanced as well as whether or not the simulation increased their interest in an international relations major. They reported generally positive results. Similarly, Shaw (2004) conducted a survey of students after two simulation activities, which demonstrated that students appreciated the simulation and reported to have gained from it.

Fewer scholars have compared classes which used simulations to classes which have not in an attempt to measure student learning in the presence of simulations. While Garrison, Redd and Carter (2010) engaged in a comparative study, their work compared the outcomes of the same simulation in three different sized classes at different universities. Crossley-Frolick (2010) compared simulation results in classes that met for different lengths of time, but again did not compare classes with a simulation to those without. Frederking (2004) in a US Government class did compare six sections of an Introduction

to American Politics class that used a Senate simulation to six sections that did not, comparing the classes a variety of learning outcomes. He found that students in classes with the simulation did better on subsequent exams in the class. He does note that while the classes were largely similar, there were differences, and that the classes were not taught simultaneously.

In the field of international relations in particular, a few scholars have directly measured the impact of active learning techniques as compared to more traditional learning. Krain and Lantis (2006) used an experimental research design, including pre- and post-tests and simulations run in each of these classes while the other served as a control to determine whether student learning improved when using a simulation. They found that students learned different things from simulations and traditional formats, but not that one method was necessarily “better,” but that they gained different insights from the simulation and generally felt positively about the activity. In studies of other active learning techniques, Krain and Shadle (2006) found students participating in a hunger week event had a better understanding of world hunger issues than similar students who were exposed only to lecture. In a result similar to Krain and Lantis (2006) Powner and Allendoerfer (2008) did not find that students in sections with active learning techniques necessarily did better overall than students who participated in traditional discussion sections, but each group of students did perform differently on essay versus multiple choice questions.

While this literature review is not exhaustive, there is a paucity of research comparing active learning techniques, simulations in particular, with more traditional assignments. Two of the studies cited above (Krain and Shadle 2006; Krain and Lantis 2006) both look at active learning techniques related to a very specific set of questions; a search of the literature does not reveal a comparative study of classes using traditional assignments with one using an extended, multi-week simulation. This paper contributes to the existing literature on active learning in international relations generally and that related to the use of simulations more specifically by using an experimental design to assess the learning

outcomes of students in a traditional international politics class supplemented by a multi-week online simulation with students in another section of the same class, taught by the same professor, who wrote a traditional research paper.

### **Setting up the Classes**

In spring 2012, the author taught two sections of Introduction to International Politics. Both classes were taught on a Tuesday / Thursday schedule with 80 minutes each day. The classes were similar in size; one class had 30 students, the other 34. While the two classes met in different rooms, the layout of each was similar and both had access to the same technology. In teaching the two classes, the instructor used the same texts and supplemental readings, as well as the same PowerPoint lecture notes. The only difference between the two classes was that one used the Statecraft simulation, while the other used a more traditional research paper. To determine which class would do the simulation and which would do the research paper, the author used a coin flip (conducted with the TCU Political Science Department's administrative assistants) to assign each class to a treatment. The morning class (taught from 9:30-10:50) was assigned the Statecraft simulation, while the afternoon class (taught 2:00-3:20) used the traditional research paper.

### *Statecraft*

Statecraft is an online simulation that requires the professor to set up an account and create the parameters for the simulation. Students create their own accounts and are assigned to countries either by the professor or by a foreign policy attitude test. For this class, students took the foreign policy attitude test. The simulation assigns students to countries in the Statecraft world; 4-6 students are assigned to each country. Each country is able to select certain attributes (such as the economic and government structures, the name of the country and major cities). With each turn of the simulation,

countries are able to implement projects, purchase military equipment, and engage in foreign relations with other countries. Countries earn resources each turn, and the simulation tracks domestic approval ratings, faction strength, overall social and economic welfare, as well as international performance. A series of events occur in the simulation, including terrorism and environmental problems that countries can tackle individually or use mechanisms in the game to tackle collectively.<sup>2</sup>

The instructor has a great deal of flexibility in setting up the Statecraft simulation. For this class, the simulation played out over eight weeks (or turns). Students met in class with their “countries” for 25-30 minutes each Tuesday in class. During this time, they were able to make plans for their domestic projects, military acquisitions, and any international actions they desired to take. Each week, each student had to prepare a policy brief of 250 words submitted directly to the Statecraft system. At the conclusion of the simulation, students completed a four page paper using material learned in class to analyze the simulation and compare it to a real-world event. The simulation grade was determined by two quizzes over the simulation mechanics (10%), completion of memos (30%) and the final paper (60%).

### *Research Paper*

The research paper assignment was a traditional open ended assignment. The base assignment was as follows:

You will choose a research question concerning some issue or event dating from the end of World War II to the present. You will develop a hypothesis (hypotheses) that you will test based on the historical record of the event / issue of interest to you. Your paper will be 1600-1800 words, double spaced. The word count will exclude references. 10 points will be automatically deducted for any paper that falls outside of the word range.

Your paper should follow this basic outline:

- I. Introduction to the question
- II. Theoretical explanation
- III. Empirical support for explanation

---

<sup>2</sup> For a more detailed review of Statecraft, see Cox 2014.

- IV. Alternative explanation
- V. Conclusion

The paper requires a minimum of 8 sources. Unless special permission is given, the eight sources must come from peer reviewed journals or from books in the library. I recommend your using the library website as a starting point for your research.

Additionally, in order to balance the time lost in class to the Statecraft simulation and to improve the paper itself, three different classes were dedicated to either learning about how to do a research paper or in conducting peer review. On January 31, the first day Statecraft teams met, the afternoon class had a general lesson on how to construct an international relations research paper. Later in the semester, one entire class period was dedicated to students discussing their research questions and presenting their preliminary findings in peer groups. Finally, one day in class was dedicated to peer review. Ultimately, the two classes lost similar amounts of class time overall; however, the research paper class missed one full day each prior to each of the three exams. The Statecraft class missed the majority of its class time prior to the second exam, with some time missed prior to the first exam as well.

The project was submitted to the Institutional Review Board at TCU and given an exemption in compliance with TCU policy. Students were informed on the first day of class that they were involved in a pedagogical study and were told in detail both the assignment that they would have and the assignment that the other class would have; no students switch between classes, though both classes did have minor attrition over the course of the semester.

### **Comparing the Classes**

Two types of evaluations were used to compare the assignments. First, student learning was measured by comparing results on in-class exams. Second, students completed a four question survey on the final full day of class (before exam week) administered by a neutral party. Students were informed the survey would not be seen by the instructor until after grades had been submitted.

### *Comparing Exam Performance*

The first measure used in class exams. To ensure that the students were tested identically, the exam questions in each class were the same, though they were presented in different order.<sup>3</sup> Each exam contained a combination of multiple choice (30) and short answer (5) questions. The final exam added a single essay question. Rather than simply comparing final test averages in the two classes, I used linear regression to test differences between the classes on each of the three exams. Rather than simply using one independent variable (Statecraft), I also created a dummy variable indicating whether or not the student had a political science major or minor versus a minor outside of the department. I also included student classifications – interestingly, no students were classified as freshman under the hour-based classification system that the school uses. Finally, I created a dummy variable for students who were taking the class on a pass/no credit basis under the belief that such students may not try as hard for a good grade as other students.

The class utilizing Statecraft had 3 students taking the class pass/no credit (requiring a minimum grade of C- to pass), 13 students with one of three political science majors or a political science minor, and 17 students who were not part of the political science department. The class was younger; 16 students were classified as sophomores, or second year students, 10 as juniors and 4 as seniors. The research paper class was similar overall, but was a bit older. Like the Statecraft class, 3 students were taking it pass/no credit, 17 had a major or minor in political science, and 17 did not. 14 students were sophomores, 12 juniors and 8 seniors.

I constructed six models to look at the exams. Models 1, 3, and 5 included the dummy variables for pass/no credit and political science or not, as well as the students' classification. Models 2, 4, and 6 included these variables as well as the dummy variable for Statecraft or research paper. Models 1 and 2

---

<sup>3</sup> This may have introduced the possibility of cheating if students from the first class informed students in the second class of material on the exam. However, as the classes were only separated by three hours, I determined that the necessity of giving students identical measurements outweighed the risk of cheating.



used the score on the first exam as the dependent variable; Models 3 and 4 used exam two, and Models 5 and 6 used exam 3.<sup>4</sup>

As will be seen below, only one model was significant overall, and that was significant only at  $p < .1$ , and only one variable in that model attained significance at  $p < .05$ . One additional variable in Model 5 was significant.

### *Models 1 & 2*

Model 1, including graded versus not graded, political science or not, and student standing, had an  $R^2$  of .018. As seen in Table One, no variable was significant. Model 2 added the Statecraft variable. The overall model had an  $R^2$  of .040. Neither model was significant. As seen in Table Two, again no variable was significant.<sup>5</sup> Importantly, however, the non-standardized coefficient for Simulation or not (Sim) was positive. Students in the research paper class scored about 3 points higher on average than the Statecraft class.

**Table One – Exam One**

Model 1		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	81.363	4.904		16.592	.000
	Graded	1.568	4.724	.045	.332	.741
	PoliSci Major-Minor	-2.695	2.789	-.131	-.966	.338
	Standing	.893	1.759	.066	.508	.613

<sup>4</sup> Students were able to complete bonus assignments for the first exam that were largely completion grades. To eliminate the effect of the bonus assignments, the test scores without bonus were used for this test.

<sup>5</sup> Though the variable was not significant, I was pleased to see that the coefficient on political science students was negative, indicating that political science students did do better than non-political science students.

**Table Two – Exam One**

Model 2		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	80.467	4.947		16.265	.000
	Graded	1.629	4.709	.046	.346	.731
	PoliSci Major-Minor	-2.400	2.792	-.117	-.860	.393
	Standing	.553	1.777	.041	.311	.757
	SIM	3.125	2.658	.152	1.176	.244

*Models 3 & 4*

Models 3 and 4 focused on the second exam. Model 3, including the pass/no credit, political science, and standing variables had an  $R^2$  of .046, and was not significant. Model 4, with an  $R^2$  of .123 was significant at  $p < .1$ . In that Model, the Simulation variable was significant, showing an effect of almost six points different on the exam. Tables Three and Four show the results for the individual variables.<sup>6</sup>

**Table Three -- Exam Two**

Model 3		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	76.281	4.903		15.558	.000
	Graded	-.579	4.723	-.016	-.123	.903
	PoliSci Major-Minor	-4.418	2.789	-.212	-1.584	.118
	Standing	.878	1.759	.064	.499	.620

**Table Four – Exam Two**

Model 4		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	74.599	4.798		15.547	.000
	Graded	-.465	4.567	-.013	-.102	.919
	PoliSci Major-Minor	-3.864	2.708	-.186	-1.427	.159
	Standing	.239	1.724	.017	.138	.890
	SIM	5.862	2.578	.282	2.274	.027

---

<sup>6</sup> Again, political science students did better than non-political science students.

### Models 5 & 6

Models 5 and 6 focus on exam three in the class. Model 5 includes the pass/no credit, political science, and standing variables, while Model 6 adds the simulation variable. Neither Model 5 nor Model 6 were significant;  $R^2$  for Model 5 was .063, and  $R^2$  for Model 6 was .095. As can be seen in Table Five, the political science major/minor or not variable was significant in Model 5. It continued to be significant in Model 6, albeit at a slightly reduced level. The simulation variable was not significant in Model 6. Tables Five and Six present results for the individual variables in Models 5 and 6 respectively.

**Table Five – Exam Three**

Model 5		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	86.047	4.284		20.087	.000
	Graded	3.055	4.379	.092	.697	.488
	PoliSci Major-Minor	-4.728	2.402	-.262	-1.968	.054
	Standing	.045	1.541	.004	.029	.977

**Table Six – Exam Three**

Model 6		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	84.989	4.309		19.723	.000
	Graded	3.376	4.346	.101	.777	.440
	PoliSci Major-Minor	-4.430	2.390	-.245	-1.853	.069
	Standing	-.268	1.543	-.022	-.174	.863
	SIM	3.286	2.289	.182	1.436	.157

### Discussion

No variable consistently had a significant effect on student exam scores. Though students in the simulation class did score about 3 points worse on exams one and three, the effect was not significant, and the margin of error was somewhat large. On the second exam, however, students in the simulation class did much worse; the raw difference was 6.18 points worse. This difference was significant. While

not testable, I speculate that this difference was because students in the research paper class had more lecture time than students in the simulation class. Had the students been tested on concepts covered more specifically by the simulation in that class, I suspect the grades would have been different.<sup>7</sup>

Interestingly, class standing had no real effect on student performance; younger students did just as well as older students. Additionally, political science majors and minors did consistently better than non-majors and minors. Though the direction and approximate size of the effect was similar across models, it was only significant in one model.

### *Student Reaction*

As noted above, students in each class were given a simple four question assessment on the last day of class. Each question was answered on a scale of 1-5, with 1 being *Strongly Disagree*, and 5 being *Strongly Agree*. The students in each class were read a description of the assignment the other class completed. The questions were as follows:

Q1: I enjoyed the time spent on the [research paper/Statecraft] assignment.

Q2: The [research paper/Statecraft] assignment improved my understanding of class material.

Q3: The time spent in class on the [research paper/Statecraft] assignment was a good use of time.

Q4: Instead of the research paper, I would have preferred to have completed the [research paper/statecraft] assignment completed by the other section of this class.

Tables Seven and Eight present the results for the Statecraft and research paper classes respectively. As can be seen in these results, a higher percentage of students in the statecraft class either agreed or strongly agreed that they enjoyed the assignment (79%) versus those who did not (7%). More tellingly, 80% of students in the Statecraft class disagreed or strongly disagreed with the statement that they would rather have done the research paper. Interestingly, if one adds up the research memos and the

---

<sup>7</sup> As I could not rule out that the treatment led to lower grades on this exam, I did curve the exam scores for the simulation class.

final paper, the Statecraft students did a great deal more work on their assignment. Further, the overall Statecraft assignment grades were almost identical to the final research paper grades (86.03 for Statecraft versus 86.7 for the research paper).

**Table Seven: Statecraft class responses**

	<b>Strongly Disagree</b>	<b>Disagree</b>	<b>Neutral</b>	<b>Agree</b>	<b>Strongly Agree</b>
<b>Q1: Enjoyed time</b>	1	1	4	14	8
	3.57%	3.57%	14.29%	50.00%	28.57%
<b>Q2: Improved Understanding</b>	2	4	9	10	3
	7.14%	14.29%	32.14%	35.71%	10.71%
<b>Q3: Good use of time</b>	0	1	8	9	10
	0.00%	3.57%	28.57%	32.14%	35.71%
<b>Q4: Would rather have done other assignment</b>	18	4	3	1	2
	64.29%	14.29%	10.71%	3.57%	7.14%

In contrast to the Statecraft class, while a slightly higher percentage of students agreed or strongly agreed that the research paper helped them to better understand course material (55% in the research paper class versus 46% in the Statecraft class), fewer students enjoyed the assignment and more expressed interest in completing the other assignment. Whereas only 11% of students in the Statecraft class would have preferred to have done the research paper, 55% of students in the research paper class would have preferred to have participated in Statecraft.

**Table Eight: Research paper class responses**

	<b>Strongly Disagree</b>	<b>Disagree</b>	<b>Neutral</b>	<b>Agree</b>	<b>Strongly Agree</b>
<b>Q1: Enjoyed time</b>	3	6	10	10	0
	10%	21%	34%	34%	0%
<b>Q2: Improved Understanding</b>	3	7	3	12	4
	10%	24%	10%	41%	14%
<b>Q3: Good use of time</b>	0	4	7	9	9
	0%	14%	24%	31%	31%
<b>Q4: Would rather have done other assignment</b>	4	4	5	6	10
	14%	14%	17%	21%	34%

## Conclusion

The results of this study are somewhat mixed and do not lead to easy answers. In this conclusion I will discuss both ways to improve the study as well as to improve the integration of Statecraft into the class.

While it is important to note that Statecraft students were only worse on a statistically significant level in one model, and it was on the exam prior to which students spent the most time in class on Statecraft. One could interpret this result in one of two ways. A negative interpretation would be that Statecraft negatively impacted student learning; students did not acquire the same knowledge as students in the other class; one should avoid using such an extensive simulation. An alternative reading would be that assessments and lectures need to be designed to address the different things learned in Statecraft. Additionally, students completing the research paper did not do significantly better than Statecraft students on the final exam – in that regard, a traditional research assignment did not appear to have an advantage in overall student learning.

Further, I did not fully use the simulation as intended in the semester in which I conducted the study. So as to not overly bias the results, I was careful to use the same basic lecture material in both classes. I did not supplement my lectures with extra discussions of how Statecraft related to the class. Neither could I add Statecraft questions and keep the instruments the same. In subsequent classes, I have more intentionally incorporated Statecraft into my lectures and have revised assignments to require students to draw more heavily on Statecraft in analyzing historical and current events.<sup>8</sup> One area in which the Statecraft class did suffer was in discussing supplementary readings; the research paper class received more in depth discussion of articles that appeared on the class.

In Statecraft's favor, students did enjoy the exercise and they believed that it was useful. While the exams may not have measured the knowledge gained, students completing the simulation did believe it improved their understanding of international affairs. In their comments, reflecting the sentiments above, many students did encourage the professor to do more to integrate the simulation into the theories and examples discussed in class.

Additional testing of the hypotheses using better data may alter the results as well. Based on incomplete data, I strongly suspect that the students in the research paper class have higher GPAs overall than did students in the Statecraft class. As I did not have complete data at the time this paper was written, I was unable to include GPA in my findings. However, I am still working to acquire that data to determine if prior student performance is a better predictor than any of the variables I have included.

In conclusion, the results of this study reaffirm what other studies have shown. Demonstrating the benefits of active learning and simulations on a large scale may at times be difficult. Students may be learning different things through active learning than they are in traditional classes; directly comparing the two is difficult. Nonetheless, it does seem clear that, in student eyes at least, there is a tangible benefit to continuing to engage in active learning techniques.

---

<sup>8</sup> The Statecraft website contains several resources to help professors incorporate lessons from the simulation into their classes.

## References

- Butcher, C. (2012). Teaching foreign policy decision-making processes using role-playing simulations: The case of US-Iranian relations. *International Studies Perspectives*, 13(2), 176-194. doi:10.1111/j.1528-3585.2012.00457.x
- Cox, E. (2014). Review of statecraft online simulation. *Journal of Political Science Education*, 10(2), 243-244. doi:10.1080/15512169.2014.893761
- Crossley-Frolick, K. A. (2010). Beyond model UN: Simulating multi-level, multi-actor diplomacy using the millennium development goals. *International Studies Perspectives*, 11(2), 184-201. doi:10.1111/j.1528-3585.2010.00401.x
- Frederking, B. (2005). Simulations and student learning. *Journal of Political Science Education*, 1(3), 385-393. doi:10.1080/15512160500261236
- GARRISON, J. A., REDD, S. B., & CARTER, R. G. (2010). Energy security under conditions of uncertainty: Simulating a comparative bureaucratic politics approach. *Journal of Political Science Education*, 6(1), 19-48. doi:10.1080/15512160903467653
- Krain, M., & Lantis, J. S. (2006). Building knowledge? evaluating the effectiveness of the global problems summit simulation. *International Studies Perspectives*, 7, 395-407.
- Krain, M., & Nurse, A. M. (2004). Teaching human rights through service learning. *Human Rights Quarterly*, 26, 189-207.
- Krain, M., & Shadle, C. J. (2006). Starving for knowledge: An active learning approach to teaching about world hunger. *International Studies Perspectives*, 7(51), 66.
- Krain, M. (2010). The effects of different types of case learning on student engagement. *International Studies Perspectives*, 11(3), 291-308. doi:10.1111/j.1528-3585.2010.00409.x
- Patterson, A. (2000). It's a small world: Incorporating service learning into international relations courses. *PS: Political Science & Politics*, 33(4), 817-822.



- Powner, L. C., & Allendoerfer, M. G. (2008). Evaluating hypotheses about active learning. *International Studies Perspectives*, 9(1), 75-89.
- Rothman, S. B. (2012). Developing and adapting simulations through six points of variance: An example of teaching applied game theory through international negotiations. *International Studies Perspectives*, 13(4), 437-457. doi:10.1111/j.1528-3585.2012.00466.x
- Schnurr, M. A., De Santo, E., & Craig, R. (2013). Using a blended learning approach to simulate the negotiation of a multilateral environmental agreement. *International Studies Perspectives*, 14(2), 109-120. doi:10.1111/j.1528-3585.2012.00470.x
- Shaw, C. M. (2004). Using role-play scenarios in the IR classroom: An examination of exercises on peacekeeping operations and foreign policy decision making. *International Studies Perspectives*, 5(1), 1-22. doi:10.1111/j.1528-3577.2004.00151.x
- Shellman, S. M. (2001). Active learning in comparative politics: A mock german election and coalition-formation simulation. *PS: Political Science & Politics*, 34(4), 827-834.
- Shellman, S. M., & Turan, K. (2006). Do simulations enhance student learning? an empirical evaluation of an IR simulation. *Journal of Political Science Education*, 2, 19-32.
- Taylor, K. (2013). Simulations inside and outside the IR classroom: A comparative analysis. *International Studies Perspectives*, 14(2), 134-149. doi:10.1111/j.1528-3585.2012.00477.x