Going Beyond the Existing *Consensus*: The Use of Games in International Relations Education

Michael Lee  
Hunter College, CUNY

Zachary C. Shirkey  
Hunter College, CUNY

Despite the popularity of using games to teach international relations, few works directly assess their effectiveness. Furthermore, it is unclear if games help all students equally or if certain students are more likely to benefit. Finally, how closely the game must mirror the concept being taught to be an effective pedagogical tool has received scant attention. We address these points by discussing the use of an updated version of the classic American electoral game, *Consensus*, to help illustrate the role of domestic political coalitions in an international political economy course. Assessing the performance of students on a pre- and post-quiz, we find evidence that student performance improved overall, particularly among students that played games frequently.

One challenge of teaching political science is that the experience of powerful political actors is distant from that of undergraduates. Games that place students in the role of decision-makers can overcome this obstacle. Surveys have shown students believe games enhance their understanding of abstract concepts and increase their interest in politics and international relations (Dougherty 2003; Shellman and Turan 2006). The use of games also aligns with findings that students are more likely to learn when instructors use techniques that engage students (Dorn 1989; Endersby and Shaw 2009; Loggins 2009). Students are more likely to retain knowledge through activities which combine doing and saying rather than through other types of learning activities such as hearing, reading, or even doing and saying separately (Boyer el al 2000). Games can demonstrate how theories “work”, illustrate how institutions functions, and get students to understand the emotions of actors in situations which are alien to them (Asal 2005; Stover 2007). Finally, politics makes the most sense if experienced or “played” rather than
just discussed (Asal and Blake 2006). Not surprisingly, in recent years a variety of articles have appeared on the use, construction, and integration of games and simulations in international relations instruction (Arnold 2015; Asal 2005; Asal and Blake 2006; Boyer, Trumbore, and Fricke 2006; Dougherty 2003; Haynes 2015; Simpson and Kaussler 2009; Starkey and Blake 2001; Wheeler 2006).

That said, much of the evidence that games increase student learning has been impressionistic (Shellman and Turan 2006; Wheeler 2006) or has relied on indirect assessments such as student surveys and evaluations (Krain and Lantis 2006). Existing direct assessments of games largely rely on natural experiments. For example, Frederking (2005) reported that in years he included a simulation in his Introduction to American Politics classes, student scores improved significantly in exams offered after the simulation compared to those offered before the simulation. In years when no simulation was included, exam scores did not significantly improve.

An exception to the lack of direct assessment of the impact of simulations and games is Krain and Lantis (2006). They examined the effectiveness of the Global Problems Summit simulation by comparing pre- and post-tests results for students who participated in the simulation to those who were taught the material without incorporating the simulation. They found that scores improved roughly equally in both the control and test populations, but that the areas of greatest improvement varied across groups. This suggests that simulations may impart certain types of knowledge better than traditional instruction methods and vice versa. Such direct assessments have remained rare. Equally important, little has been done to build upon Krain’s

---


We aim to add to these findings in three ways. First, we directly assess the impact of participating in a game on student learning. This allows us to determine how large the effect of the game is on student learning. Knowing the size of an effect is vital as games are time-consuming. Such large investments of time are more defensible if the impact on student learning is also large. Second, we employed a game that did not directly mirror the course material to determine if games that illustrate similar concepts and mechanisms to the material being covered can be effective pedagogical tools even though they do not directly reflect the substantive material. This is an important question. Unfortunately, though many games exist, there is not a game for every topic. Given instructors may lack the time or expertise to construct their own game, the ability to use ready-made games that only imperfectly fit the material being covered but which illustrate similar mechanisms and concepts would be a significant boon. Further, by modeling the impact of a less directed gaming session we may gain some insight into what students learn when they play thematically relevant games outside class. Finally, we would like to know whether games benefit some students more than others. Perhaps avid gamers learn more because games illustrate concepts in a manner that is familiar to them. Alternatively, perhaps those who do not regularly play games would benefit more as they would be exposed to a new way of seeing concepts. Either result would have considerable implications for how instructors incorporate games into the classroom (e.g., using classroom time to play games versus providing students with supplementary materials). Likewise, we are interested in discovering whether the learning impact of games varies across student ability. Perhaps stronger students would gain
more due to greater motivation. Or maybe weaker students would benefit more because games present an alternative to traditional teaching techniques that they are less able to follow.

In order to answer these questions, we had undergraduate students in a 200-level international political economy course play an updated version of the 1960s board game Consensus. In the game, players allocate scarce campaign resources to win over domestic interest groups in the United States. With these groups’ support, players ultimately try to be elected President of the United States by winning states worth a majority of votes in the Electoral College. The game was incorporated during a unit on the effect of domestic politics on international trade and financial policy.

**Experimental Design**

Our experiment consisted of three components: a pre-quiz; a gaming session; and a post-quiz. Students were recruited from two introductory International Political Economy classes, one an honors section (with 10 students), the other a non-honors section (with 35 students). By including both groups, we can assess the utility games to assist student learning between more and less advanced students.

Following Human Subjects research protocol, students were allowed to opt out of sharing their data for research purposes; however, participation in the simulation was mandatory. Specifically, the pre-quiz, participation in the game, and post-quiz were each worth 5% of a student’s grade, for a total of 15%. As a result, the stakes of the assignment were reasonably high. A substantial proportion of students participated in all parts of the exercise (39/45) and most of the instances of non-participation were the result of student absences. As a result, we are reasonably confident that our sample is not systematically biased by our recruitment process.
The pre-quiz consisted of two sections. The first section included six multiple choice questions on the relationship between electoral systems, domestic politics, and trade policy. Each concept drew from influential works on the politics of trade (Busch and Reinhardt 2000; Garrett, 1998; Grossman and Helpman 2002; Hiscox 2003; McGillivray 2004; Milner 1997; Olson, 1962; Rogowski 1987; Tsebelis 2002) and had been discussed in previous lectures and in the course textbook (Oatley 2012: 68–89). Students in the honors and non-honors classes were given different questions in order to prevent cheating. The quizzes are available in an online appendix. The second section inquired about student demographic information including: how frequently students played strategy board games or video games; gender; interest in the course topic; interest in international political economy; student assessment of their understanding of course material; and student assessment of their knowledge of which factors policymakers consider when they run in elections and govern countries.

Students were then divided into groups to play Consensus. Students were provided with the rules to the game in advance as well as a video demonstrating sample turns. In Consensus, candidates allocate scarce campaign hours to different states in a United States presidential election. Each turn, candidates simultaneously reveal which states they have chosen to campaign in. More populous states require more campaign hours than less populous ones. Whichever candidate has spent the most cumulative time in a given state, leads in that state. Each state, in turn, has ties to particular interest groups. For instance, Alabama has ties to the military-industrial complex, evangelical groups, and manufacturing interests. When a candidate takes the lead in a state, they gain the state’s campaigning hours toward interest groups that are influential within that state. A candidate controlling a majority of the hours for a given interest group gains that group’s endorsement, granting them additional campaign hours to spend in states where the
group is influential. A candidate wins the game by locking down a majority of the Electoral College. Our updated *Consensus* map, game rules, and interest group tables are included in an online appendix.

*Consensus* contains a number of features that are useful for explicating course concepts. The game involves an electoral contest under majoritarian rules, namely, the Electoral College. Because campaign hours are scarce, candidates’ campaigns often focus on battleground states, expending less effort in safe states. The importance of interest groups in the game also pushes students to think about parties as representing coalitions of interests, rather than simply a particular ideological position. Further, some interest groups overlap more than others. For instance, states with ties to evangelical groups tend also to have a strong gun lobby. As a result, students soon realize that only so many coalitions of interest groups are electorally viable. As a result, students experience first-hand the ways in which American electoral institutions inhibit the emergence of a multiparty system and constrain political possibilities.

In the class following the gaming session, students participated in a post-quiz. The questions in the post-quiz addressed similar underlying concepts to those in the pre-quiz, although the wording and details of each question was different. Also, students were asked if they enjoyed the game, if they studied for the pre-quiz, post-quiz, or both, and if they believed they had a good idea of which factors policymakers consider when running in elections and governing countries.

**Results**

On average, student performance improved after playing the game. Students averaged 58.5% on the pre-quiz (SD = 1.393), and 65.8% on the post-quiz (SD = 1.317). We employed a
paired sample t-test to assess whether individual student improvements were statistically significant. Paired samples difference of means reveal that for the entire sample (N = 39) the result \( t(39) = 1.716 \) was significant at the 95% level (p-value = .0471). Confining our analysis to the non-honors class (N = 32), the results were stronger \( t(32) = 1.869 \) and significant at the 95% level with a p-value of .0356. Thus, playing Consensus had a positive impact on student learning. Certainly, we would caution that our experimental design was limited. Ideally we would compare experimental results to a control group.

Additionally, student response to the game was largely positive. Many students conveyed that they enjoyed the game and participated enthusiastically. When surveyed, 64.1% said that they preferred the game to a lecture, while 33.3% were indifferent between the two, and 2.6% (a single student) preferred a traditional lecture format.

**Distributional impacts of the game**

Understanding who gains can be essential to unlocking the underlying causal mechanism by which games enhance learning. If a narrow subset of students (e.g., gamers, honors students) tends to gain from games, this might suggest alternative ways to incorporate games into classroom learning. As our dependent variable we used student performance on the post-quiz, while controlling for performance on the pre-quiz. We employ a Tobit regression to address the issue that our data is censored (see Sigelman and Zeng 1999). What we are really interested in is whether student knowledge of the domestic politics of trade policy changed. However, our post-quiz is limited in its ability to capture this. For instance, a strong undergraduate and a political science professor might both score a perfect six, although the latter is more knowledgeable.

---

2 There is a rule of thumb that sample sizes of 30 distinguish large from small sample techniques. Some works argue that even this number is too conservative (Cohen 1990). Stover (2007) uses similar techniques to us with a smaller sample size of 32.
Similarly, on the low end, two students with different (but low) knowledge of trade politics might both hypothetically score zero. Our Tobit regression analysis had a lower limit of zero and an upper limit of six, with five students hitting the upper limit.\(^3\)

We included three variables reflecting our questions about whether regular gamers tend to gain more from in-class games, our interest in the impact of games on honors and non-honors students, and our question of whether student enthusiasm predicted better results. To gauge whether students were gamers, we used the results of our survey question: “How often do you play board games, and/or strategy video games?” We employed a variable ranging from zero (a student that never played games) to three (a student that often played games) to capture whether or not students were gamers. We also included a binary variable indicating whether students were in the honors program or not. In order to gauge the effect of student enthusiasm, we included a variable indicating whether students preferred the game to a lecture (+1), a lecture to the game (-1), or were indifferent (0).

We also included a set of controls. We controlled for performance (ranging from zero to six) on the pre-quiz. We also included two binary variables pertaining to studying behavior: one indicating whether a student only studied for the post-quiz and another indicating whether or not a student studied at all. It is possible that some students saw their results improve because they only studied for the post-quiz. Additionally, if the game had some scholastic impact, its impact might be higher among students that studied the least, and thus, were not at a point where additional effort would yield diminishing returns. Finally, we included a variable for gender, which might group dynamics and assessments of what constitutes playing strategy board games “often.” Descriptive statistics of these variables are summarized in Table 1.

\(^3\) OLS and negative binomial regression produce substantively similar results.
Student performance on the pre-quiz exhibited a positive, weakly significant impact on performance on the post-quiz (see Table 2). Similarly, gamers also experienced statistically significant gains in quiz performance. In contrast, students that only studied for the post-quiz and those that enjoyed the simulation performed worse on the post-quiz, a result that was statistically significant. Gender, honors status, and not studying for either quiz all were statistically insignificant.

If we look at the substantive impact of a student being a frequent gamer, the impact of playing the game was quite large. Using model 1, holding all binary variables at zero, and non-binary variables at their mean, our model would predict a score of 3.088/6 for a student answering “never” to the gaming question, and a 4.477/6 for a frequent gamer. This relationship is summarized in Figure 1.

Our distributional results shed some light on the causal mechanisms behind student learning in games. Although many students expressed enthusiasm about the game, enjoyment did not predict better results on the quiz. Likewise, it is implausible that our results occurred simply because some students studied more for the post-quiz than the pre-quiz as students that studied only for the post-quiz did poorly. Moreover, our findings suggest that games can be advantageous both for honors and non-honors classes. Ceteris paribus, both experienced gains, challenging either the notion that games are only appropriate for advanced students able to follow complex rules, or the notion that games will have a larger impact on non-honors students.
Discussion

Our most important finding is that gamers gain the most from in-class games. We believe this result can be leveraged in a particularly useful manner. In addition to using games as a learning tool in the classroom, encouraging students to play games that are complementary to course concepts on their own may be beneficial. For instance, a class exploring balance of power theory might encourage students to play *Diplomacy*, providing students with materials that facilitate debriefing. While many instructors are not gamers and may have trouble providing suitable recommendations, we suspect many instructors do play games and could provide debriefing notes for games they do know. If, through collaboration, we could assemble a single centralized list, it would be fairly simple for instructors – whether or not they are gamers – to incorporate such a list to their syllabi. A centralized list would not only help instructors looking for viable simulations, but could also help professors make suggestions to students who are inclined to play board games or video games anyway.

Works Cited


Cohen, Jacob. 1990. “Things I have learned (So Far),” American Psychologist, 45(12): 1304-1312


List of Tables

Table 1: Descriptive Statistics for Tobit Model of post-quiz performance

<table>
<thead>
<tr>
<th>Variable</th>
<th>Min</th>
<th>Mean</th>
<th>Max</th>
<th>Std. Dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Post-quiz</td>
<td>1</td>
<td>3.949</td>
<td>6</td>
<td>1.317</td>
</tr>
<tr>
<td>Pre-quiz</td>
<td>1</td>
<td>3.513</td>
<td>6</td>
<td>1.393</td>
</tr>
<tr>
<td>Honors</td>
<td>0</td>
<td>0.18</td>
<td>1</td>
<td>0.389</td>
</tr>
<tr>
<td>Gamer</td>
<td>0</td>
<td>1.436</td>
<td>3</td>
<td>0.882</td>
</tr>
<tr>
<td>Gender</td>
<td>0</td>
<td>0.41</td>
<td>1</td>
<td>0.498</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>-1</td>
<td>0.616</td>
<td>1</td>
<td>0.544</td>
</tr>
<tr>
<td>Studied for post-quiz</td>
<td>0</td>
<td>0.051</td>
<td>1</td>
<td>0.224</td>
</tr>
<tr>
<td>Did not study</td>
<td>0</td>
<td>0.282</td>
<td>1</td>
<td>0.456</td>
</tr>
</tbody>
</table>
Table 2: Coefficients of Tobit Model for post-quiz performance

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-quiz score</td>
<td>0.342*</td>
<td>.335*</td>
</tr>
<tr>
<td>(out of 6)</td>
<td>(.175)</td>
<td>(.169)</td>
</tr>
<tr>
<td>Gamer (0 = never, 3 = often)</td>
<td>0.545**</td>
<td>.529**</td>
</tr>
<tr>
<td>Gender (1 = male)</td>
<td>-.126</td>
<td></td>
</tr>
<tr>
<td>Honors</td>
<td>.71</td>
<td>.704</td>
</tr>
<tr>
<td></td>
<td>(.434)</td>
<td>(.436)</td>
</tr>
<tr>
<td>Enjoyment</td>
<td>-1.247***</td>
<td>-1.22***</td>
</tr>
<tr>
<td>(1 = prefer game)</td>
<td>(.429)</td>
<td>(.401)</td>
</tr>
<tr>
<td>Studied post-quiz</td>
<td>-1.645**</td>
<td>-1.593**</td>
</tr>
<tr>
<td></td>
<td>(.635)</td>
<td>(.613)</td>
</tr>
<tr>
<td>No studying</td>
<td>.562</td>
<td>.528</td>
</tr>
<tr>
<td></td>
<td>(.498)</td>
<td>(.462)</td>
</tr>
<tr>
<td>Constant</td>
<td>2.66***</td>
<td>2.646***</td>
</tr>
<tr>
<td></td>
<td>(.868)</td>
<td>(.88)</td>
</tr>
</tbody>
</table>

N = 39
F-statistic = 3.89*** 4.55***

Robust Standard errors in parentheses
* p<.1, ** p<.05, *** p<.01
Figure 1: Predicted post-quiz score by frequency of gaming, with 95% confidence intervals