Using a Novel Gravity Model for Ranking and Assessment of Educational Games

Qichao Wang, Virginia Tech

Qichao Wang is a PhD student in the Transportation Infrastructure and Systems Engineering program at Virginia Tech. He holds a Bachelor of Engineering in Traffic Engineering from Nanjing Tech University, P.R.China (2014). His research interests include Game-aided Pedagogy, 3D visualization, traffic control, multi-agent system, machine learning and optimization.

Dr. Montasir Abbas, Virginia Tech

Dr. Montasir Abbas is a faculty member in the Transportation Infrastructure and Systems Engineering at Virginia Tech. He holds a Bachelor of Science in Civil Engineering from University of Khartoum, Sudan (1993), a Master of Science in Civil Engineering from University of Nebraska-Lincoln (1997), and a Doctor of Philosophy in Civil Engineering from Purdue University (2001).

Dr. Abbas has wide experience as a practicing transportation engineer and a researcher. He was an Assistant Research Engineer and the Corridor Management Team Leader at Texas Transportation Institute (TTI), where he has worked for four years before joining Virginia Tech. Dr. Abbas conducted sponsored research of more than $720,000 as a principal investigator and more than $750,000 as a key researcher at TTI. After joining Virginia Tech, he has conducted over $2,400,000 worth of funded research, with a credit share of more than $1,750,000.

Dr. Abbas is an award recipient of $600,000 of the Federal Highway Administration Exploratory and Advanced Research (FHWA EAR). The objective of the FHWA EAR is to “research and develop projects that could lead to transformational changes and truly revolutionary advances in highway engineering and intermodal surface transportation in the United States.” The award funded multidisciplinary research that utilizes traffic simulation and advanced artificial intelligence techniques. He has also conducted research for the National Cooperative Highway Research Program on developing “Traffic Control Strategies for Oversaturated conditions” and for the Virginia Transportation Research Council on “evaluation and recommendations for next generation control in Northern Virginia.”

Dr. Abbas developed Purdue Real-Time Offset Transitioning Algorithm for Coordinating Traffic Signals (PRO-TRACTS) during his Ph.D. studies at Purdue University, bridging the gap between adaptive control systems and closed-loop systems. He has since developed and implemented several algorithms and systems in his areas of interest, including the Platoon Identification and Accommodation system (PIA), the Pattern Identification Logic for Offset Tuning (PILOT 05), the Supervisory Control Intelligent Adaptive Module (SCIAM), the Cabinet-in-the-loop (CabITL) simulation platform, the Intelligent Multi Objective Control Algorithms (I-MOCA), the Traffic Responsive Iterative Urban-Control Model for Pattern-matching and Hypercube Optimal Parameters Setup (TRIUMPH OPS), the Multi Attribute Decision-making Optimizer for Next-generation Network-upgrade and Assessment (MADONNA), and the Safety and Mobility Agent-based Reinforcement-learning Traffic Simulation Add-on Module (SMART SAM). He was also one of the key developers of the dilemma zone protection Detection Control System (D-CS) that was selected as one of the seven top research innovations and findings in the state of Texas for the year 2002.

Dr. Abbas served as the chair of the Institute of Transportation Engineers (ITE) traffic engineering council committee on “survey of the state of the practice on traffic responsive plan selection control.” He is also a member of the Transportation Research Board (TRB) Traffic Signal Systems committee, Artificial Intelligence and Advanced Computing Applications committee, and the joint subcommittee on Intersection. In addition, he is currently a chair on a task group on Agent-based modeling and simulation as part of the TRB SimSub committee. He also serves as a CEE faculty senator at Virginia Tech.

Dr. Abbas is a recipient of the Oak Ridge National Lab Associated Universities (ORAU) Ralf E. Powe Junior Faculty Enhancement Award and the G. V. Loganathan Faculty Achievement Award for Excellence. ©American Society for Engineering Education, 2016
in Civil Engineering Education. He is also a recipient of the TTI/Trinity New Researcher Award for his significant contributions to the field of Intelligent Transportation Systems and Traffic Operations.
Using a Novel Gravity Model for Ranking and Assessment of Educational Games

Abstract
Teaching introductory transportation engineering subjects can be very challenging. These courses usually include diverse topics and can involve students from different background and interest levels. Keeping students engaged and focused requires changes in traditional delivery methods. It also requires the design of exercises that are especially targeted to address certain concepts. Game-aided pedagogy has been proposed to stimulate students’ interest and increase the efficiency of their learning. Our research team developed multiple games that were designed to target different concepts in the transportation fields. These games deliver appropriate amount of information density and accessibility, and utilize multimedia and hypermedia contents. We have developed a novel gravity model to relate students learning to information density, ability of students to absorb knowledge, and difficulty of delivery, and have previously demonstrated the model with one of the games.

In this paper, we expand and illustrate the use of the educational gravity model to assess and compare different games used in game-aided teaching. A description of two different games is included in this paper. Each game uses refined 3D vivid scenes to attract and stimulate students. Gameplay data collected include students’ responses in each game level. Both games use client-server architecture to interact with students and store their gameplay data to assess the students’ learning outcomes. We capture the effectiveness of each game by calibrating the gravity factor in each model. Each game has a naturally different gravity factor that could be associated with the game’s appeal and capability to transfer knowledge. We attempt to shed more light into this concept and the potential for its use in ranking and evaluating newly developed games in terms of their pedagogical value.

Keywords:
Game-Aided Pedagogy, Gravity Model, Learning Outcomes

1. Introduction
Studies from 1990s showed active learning is an effective way of teaching [1-4]. Active learning focuses on students’ responsibility for learning more than instruction based delivery. One approach of active learning is to immerse students in the problem environment and let them explore and learn the concepts on their way while trying to solve the problem. The large-scale, costly and complex nature of transportation system does not lend itself to students’ exploration of the real world. However, some concepts can only be appreciated and better understood in a large scale system. This usually calls for the use of simulation to reconstruct the reality and make it accessible for students [5-7].
Game-aided Pedagogy in transportation classes use simulation-based games to improve students’ understanding of certain abstract concepts. Our past studies showed that educational games can significantly improve students’ learning outcomes \cite{8-11}. Compared to pure simulation-based teaching, games can stimulate students’ engagement in learning and extend students’ learning activities beyond the class boundaries. Wang et al. proposed a gravity model to describe the engagement of students in such educational game \cite{9}. It was found that final scores (which serves as a surrogate measure of students’ capabilities) and the level number in a multi-level games (which indicates the distance between the student and the game) can significant affect the students’ engagement. However, the sample of this study was limited and only one game was evaluated. To show what kind of games can improve students’ engagement, a comparison between games is needed. The concepts presented in this paper can be used to guide the development of future educational games and their implementation strategies.

The purpose of this work is therefore focused on the following research question: what kind of multi-level educational games can stimulate students’ engagement more? Another related question is whether different levels of the same game can stimulate students’ engagement differently.

To answer these questions, the following work has been done:

1. We developed two games targeting two different concepts groups. The games were developed using the same design framework (Unity 3D game engine). The games can be easily accessed by students and can store and report the students behavior to a central server.
2. We used the games in a class. The games were used in one class so that the students’ properties can be measured under the same conditions.
3. We tested the students’ understanding of the related concepts before and after the gameplay to validate the games’ effects of improving students’ learning outcomes.
4. We calibrated a gravity-based model \cite{9} and compared the two games based on the model’s parameters.

2. Method

This work adopted the gravity model and customized it into one model that describes the engagement level of the students. Two educational games were designed targeting two different topics in transportation. Two experiments were conducted using the two games to calibrate the model and evaluate the games, respectively. The customized gravity model was the evaluation metric used in the experiments. The games were the evaluation subjects and the tools to collect the students engagement measures.
2.1 The Gravity Model

In analogy to the gravity model in physics, we use a gravity model that describes the attraction between the knowledge source (i.e., the game) and each student. The attraction (engagement) is described in equation 1.

\[ E = C e^{\frac{P_{source} \times P_{student}}{D}} \]  

(1)

Where:

- \( E \) = student’s engagement,
- \( C \) = engagement coefficient factor,
- \( P_{source} \) = properties of knowledge source,
- \( P_{student} \) = properties of a student, and
- \( D \) = distance between knowledge source and the student, which is the new knowledge.

Since the goal is to evaluate and compare the different games, there is no need to have the coefficient factor separated from the \( P_{source} \). Thus we combine and simplify the model into equation 2.

\[ E = \frac{P_{source}^e \times P_{student}}{D} \]  

(2)

Where:

- \( P_{source}^e \) = properties of knowledge source in terms of engagement

The equation can be linearized if we take all the factors as exponential function of corresponding underlining measures, i.e.:

\[ \exp(e) = \exp(P_{source}^e) \times \exp(P_{student}) / \exp(d) \]  

(3)

From equation 3, we can get:

\[ e = P_{source}^e + P_{student} - d \]  

(4)

We make the assumption that all the terms in equation 4 can be a function of certain measurable parameters. In this case, we use students’ final score to indicate the students’ properties, level number to indicate the new knowledge, the number of gameplays to indicate the engagement, i.e.,

\[ e = f_e(number of gameplays), \]

\[ P_{student} = f_{ps}(student's final score). \]
\[ d = f_d(\text{level number}) \].

Another assumption is that the relationship between the terms in equation 4 and the measurable parameters are linear, which means

\[ f_i(x) = a_i \times x + b_i \]

Where

\[ i = \text{function group indicator}, \text{ could be } e, ps, d, \text{ and } \]

\[ a_i, b_i = \text{parameters for function } f_i() \].

Then equation 4 can be written as:

\[ f_e(\text{number of gameplays}) = p^e_{\text{source}} + f_{ps}(\text{student's final socre}) - f_d(\text{level number}) \]

\[ \Rightarrow \]

\[ a_e \times (\text{number of gameplays}) = p^e_{\text{source}} + a_{sp} \times (\text{student's final socre}) - a_d \times (\text{level number}) + b_{sp} - b_e - b_d \]

Let \( \theta^e_1 = a_{sp}/a_e, \theta^e_2 = a_d/a_e, \) let the constant parameters be \( \theta^e_3 \), we can get

\[ (\text{number of gameplays}) = p^e_{\text{source}}/a_e + \theta^e_1 \times (\text{student's final socre}) - \theta^e_2 \times (\text{level number}) + \theta^e_3 \]

For comparison purposes, \( p^e_{\text{source}}/a_e + \theta^e_3 \) as a whole can be used to represent the game’s properties, therefore we denote it as \( p^e_{\text{game}} \). Then we get the customized model to describe the engagement shown in equation 5.

\[ (\text{number of gameplays}) = p^e_{\text{game}} + \theta^e_1 \times (\text{student's final socre}) - \theta^e_2 \times (\text{level number}). \] \( (5) \)

### 2.2 Game Design

The games should be easily accessible, user friendly, and have the ability to record and upload the users’ gameplay data. To achieve these properties, Unity 3D game engine was selected as the developing tool associated with Microsoft Visual Studio. The games were designed in a form of web-games. Web-games are easier to deploy and maintain. The web-games published by Unity can be accessed by multiple internet browsers with a plugin from Unity. All the games were developed under a browser/server architecture. The architecture can be found in our past work [11].

Each game should target one specific topic. Since the games would be used by the same group of students, the difference in the topic between the two games should be maximized to eliminate the
impact on the second experiment from the first game. So the selected topic were dilemma zone protection in traffic signal control and curve design in highway design.

2.2.1 DZ-man

Dilemma zone is a zone where the drivers are in a dilemma whether to continue or to stop at the onset of the yellow indication at a traffic signal. Drivers who decide to stop may risk being rear-ended by the following drivers that might decide to continue. Drivers who continue might risk running the red light. A common way of protecting dilemma zone in the real world is using green extension systems to extend the green of the subject approaches until no vehicle is present in dilemma zone.

A web-game named DZ-man was designed to mimic the control logic of the green extension system. The Graphic User Interface (GUI) is shown in Figure 1. After logging in, the player will see an intersection with a signal indicator. The player can start the game when they are ready by clicking the “start” button. After the game starts, vehicles in the queue before the intersection will start moving and new vehicles will arrive. Players can click the “stop” button to end the green after a minimum green duration. All the vehicles that passed their dilemma zone would continue and all the vehicles that haven’t reached the dilemma zone would stop. For these vehicles, their color will turn green indicating they are out of dilemma zone and are therefore safe. For the vehicles within the dilemma zone, some of them would stop and the rest would continue. All the vehicles in dilemma zone will turn red as an indication of danger. Players strive to end the green without catching vehicles in dilemma zone. This game had 5 levels. Players can replay any level as many times as they want.
2.2.2 Angry Curves

The design of curves can significantly affect the safety of a segment of highway. Engineers usually design the curves to connect given tangent lines representing the connected parts of the highway. Other relevant information includes the coefficient of side friction, designed speed, area map, and budget and space constraints. The design parameters are the curve radius and the super-elevation.

A web-game named Angry Curves was designed to illustrate the concepts in this module of a highway design class. The GUI of Angry Curves is shown in Figure 2. It has similar user control layout to DZ-man. In this game, the player is given a set amount of budget. Different design will have different costs. The objective is to design a safe curve (without making it angry) and minimize the cost. This game has two levels and an introduction chapter. After user login, the game will go into the introduction sequence, provide the background information, and provide general guidance of how to play the game. In the introduction chapter, the coefficient of side friction, design speed and relationship between cost and design parameters are given. At the first level, players design the first curve without space constraints. If the design is safe, players can see the animation of vehicles passing the curves safely. If the design is not safe, players will see the curves shaking and when vehicles drive on the curve, the shaking curve will throw the vehicles away. Players can click the “switch camera” button to switch between first person view...
and third person view. Only a safe design in the first level can let the player go to the second level. In the second level, the focus of the camera will move to the location where a second curve would be designed. With other settings being the same, this time big trees are set nearby as design constraints. Each level can be replayed as many times as the players want.

Figure 2. Game User Interface for Angry Curves

2.3 Experiment

The games were given to students as assignments during the class of Introduction to Transportation in Fall 2015. The class was taught as flipped classroom and 49 students were formally enrolled. The experiment was approved by the Institution Review Board (IRB).

The games were wrapped into modules associated with corresponding chapters. One module included a prerecorded video lecture, class discussion (with Teaching Assistant and professor’s help), a pre-quiz, the gameplay, and a post-quiz. Lecture and class discussion were the normal teaching content. The pre-quiz, gameplay, and post-quiz were the package of game-aided teaching. In this part, students took the pre/post-quiz in “scholar” (a website where the teaching materials and notices regarding the class were hosted). Students did not know their quiz scores before finishing the whole package. This was intentionally done so that the change in understanding level could only be affected by the gameplay in between the quizzes. The pre- and post-quiz were untimed.
There were 10 questions associated with the DZ-man game, and 7 questions associated with the Angry Curves game. All the questions were designed in the forms of multiple choice so that the grading of the quiz would be objective. An example of the question is shown in Figure 3.

![Post-Game Quiz](image)

**Figure 3. Screenshot of One Question**

3. Results and Discussion

328 and 174 valid data points were collected from students’ gameplays in DZ-man and Angry Curves, respectively. 32 students finished the full package of DZ-man, and 17 students finished the full package of Angry Curves. 12 students finished both packages. The gameplay data were processed to represent the number of gameplays in each level for each student. The scores from the gameplay were not taken into account because the game playing skills were not the focus of this research. The gameplay data was linked to the quiz data and the students’ final grades for this class and reformatted for the analysis in SAS JMP software [12].

The first analysis needed to be done was to test the effect of the gameplay for improving the students’ learning outcomes. Matched pairs t-tests were conducted for both gameplay experiments. Figure 4 shows the results for the tests. Figure 4 a) shows the mean difference between the scores of pre- and post- quiz for DZ-man is 0.9375 at 0.0006 significant level; Figure 4 b) shows the mean difference between the scores of pre- and post- quiz for Angry Curves is 1.05882 at 0.0006 significant level. The results show both games can improve the students’ understanding of the related concepts significantly.
Next, we looked at the features that can affect the students’ engagement of the games. The number of gameplays were used as an indicator of engagement to show how interested the students were in each game. Figure 5 a) shows the average number of gameplays in each level. We can see a decreasing trend of engagement as the game level goes up, except for level 5. Then we look into how they are distributed in Figure 5 b). It can be clearly seen that there is an outlier in level 5 who played this level more than 25 times, so we excluded this data point. The revised number of gameplays in each level is shown in Figure 6. Now we can see a clearer trend of decreasing of number of gameplays in higher levels.

**Figure 4. Matched Pairs T-Test for Both Games’ Pre/Post Scores**
For the Angry Curves, there are only two levels, which provides us a more detailed way of visualizing it. Figure 7 shows each player’s number of gameplays in each level. The horizontal axis shows the number of gameplays in level 1 and the vertical axis shows that in level 2. Each point shows one player’s number of gameplays in each level. The red line indicates where the number of gameplays in level 1 equals to the one in level 2. It can be seen that most players play level 2 more than level 1. This is because comparing to DZ-man, Angry Curves is harder in term of understanding the gameplay (that is also the reason why Angry Curves needs an introduction chapter to guide the players). The gap between players and the game is larger in the first level.
The engagement gravity model was calibrated with the formatted data. Figure 8 and Figure 9 show the results of the calibration. The intercept represent the $p_{game}^e$ in equation 5. Because the purpose was to evaluate the games for comparison, the calibrated $p_{game}^e$ should represent how an average student would engage in the game as a whole. So we shifted the final grade reference point to 85. Since the level number is a linear term, the parameters with the average of the numbers of the game should be able to represent the whole game’s properties, so the level number was referenced to 3 and 1.5 for DZ-man and Angry Curves respectively. By doing so, the intercept, the $p_{game}^e$ can represent the engagement to the game as a whole of a normal student with grade around 85.

**Figure 8. Calibrated Model Parameters for DZ-man**
Figure 9. Calibrated Model Parameters for Angry Curves

We can see that the students are more interested in the lower level of DZ-man. The sign of parameter for level is negative, which indicates in DZ-man, the higher level it is, the more new knowledge would be. While in Angry Curves, the sign of parameter for level is positive, which indicate the second level includes less information than level one. The reason is that in the first level, students need to get used to the gameplay and learn all the parameters for designing the curves, while in level two, only the constraint was added as a new information.

The different sign of the parameter for the final grade shows the difference of the fittest students group for the two different games. DZ-man can interest higher grade students more, while Angry Curves attracts lower grade students more.

Comparing the two games, Angry Curves has higher $p^{game}$ than DZ-man. This means the design of the game Angry Curves is better than DZ-man in terms of students’ engagement. The reason could be the better 3D modeling and more appealing UI in Angry Curves.

4. Conclusion and Future Work

In this work, two different games were designed and developed for two different topics within transportation engineering education. Two experiments were conducted with the two games. We also presented and customized a gravity model for evaluating the engagement of the students for the educational games. We found that different games have different properties in terms of attracting students’ engagement. In this specific case, Angry Curves can attract students more than DZ-man. This also indicate that the UI design is an important factor in the educational game design—a point that should be paid more attention by the educators. Within the same game, it was found that students’ engagement are different in different levels.
The significance of this work does not stop at comparing these two different games. It was found that the games can significantly improve the students’ understanding of the targeting concepts. There should be more educational games developed for transportation education in the near future. With more development teams contributing to the educational game libraries, there will be more than one game designed for the same topic. This ranking and assessment framework can be used in education practice to select the most suitable game for the targeted group of students.

The future work needs more tests in a larger scale. Cross institutional collaboration and testing will be considered and pursued in the near future.

Acknowledgement

This work was funded by NSF-TUES-Type 1 grant: Game-Aided Pedagogy to Improve Students’ Learning Outcomes and Engagement in Transportation Engineering. Grant number 1245728.

Reference: