DataRPG: Improving student motivation in data science through gaming elements

Azim Abdool, Daniel Ringis, Aniel Maharajh, Lynda Sirju and Hannah Abdool
Department of Electrical and Computer Engineering
University of the West Indies
St. Augustine, Trinidad and Tobago
Email: {azim.abdool, daniel.ringis, aniel.maharajh, lynda.sirju, hannah.abdool} @sta.uwi.edu

***Abstract*—The amount of data is exploding, and by all estimations will continue its exponential growth. This has led to an increased demand for engineers and data scientists. However, courses imparting crucial ideas from data design, administration and maintenance to data analytics and knowledge discovery remain under-subscribed due to perceived difficulty, low interest and motivation. This motivation can be represented by three constructs: interest in the topic, personal control and difficulty of the exercise. Role-playing games (RPGs) have been proven to stimulate students interest in analysis and problem solving when faced with varying tasks. This paper investigates the integration of specific game-based elements and pop-culture references into course delivery and assessment to boost interest and learner motivation in an elective postgraduate database systems course at The University of the West Indies. The experimental setup for assessment of students' progress in this course uses an experience point (XP) level-based system adapted from popular RPGs. To further stimulate interest, pop-culture references to specific animated series' and fantasy game genres are used. A student can choose to take one of multiple paths for their learning, representing different tools/platforms that can be used in implementation and analysis. Students can attempt more than one path for more experience and higher levels, although it is not expected. Completion of each path indicates formative "mastery" of the selected data science tool/platform. Summative assessments uses a "raid-party system" where students who have been encouraged along different paths are now asked to join together to 1. analyse given datasets and capture their findings in the format of an academic paper, and 2. build a complete database application for a particular scenario. Formative feedback is done in groups to leverage scaffolding. Measurement of the students' activity and interest is done using automatically generated activity logs from the online delivery course system as well as administration of an online electronic questionnaire. Both of these instruments are used to determine how effective this approach is in influencing students' interest and motivation. More than 5 times the historical enrollment numbers have been observed and initial results point to improvement in students' attitudes toward the content. This action research would be used to inform content delivery and further investigations would determine the impact on student performance.**

***Keywords*—gamification, flipped classroom, active learning, formative summative assessment**

I. Introduction

There is currently a problem with STEM in the Caribbean region. The highest ranked country within the region on the Global Innovation Index is Trinidad and Tobago (80th) with Jamaica afterwards (96th) [1]. This is also evident with the Gross Expenditure on Research and Development (GERD) metric being less than 0.33% in both Jamaica and Trinidad and Tobago. Science, Technology and Innovation (STI) have been touted by regional leaders as "the engine of social and economic prosperity" [2]. Regionally, it has become less feasible to keep up with the explosion of available data. With a sharp increase in data available for traffic, sport, crime, economics and energy, the demand for trained data scientists has increased. Therefore, teaching database fundamentals has grown in importance and its scope must now not only incorporate ideas from design, administration and maintenance but from data analytics and knowledge discovery. Unfortunately, this particular subset (areas of computer programming and data science) of STEM is severely under-subscribed not only in the region, but globally.

There are high attrition rates in the computer science and computer engineering [3]. This has been attributed to a disenchantment with programming which stems from a perceived learner difficulty and a lack of motivation for individual practice [3]. There is a direct correlation between the learner's interest in a topic and their motivation in a course.Therefore, an increase in student interest in the topics of data science and programming should lead to improved motivation and possibly performance.

In discipline-specific skills such as data science and computer programming, continuous practice is required [4]. For this, active-learning project-based activities are usually required. Prior research has shown that active learning methodologies are beneficial to the learner provided that the methodology does not become a hindrance to the actual lesson to be learnt.

Our study looked at an alternative course delivery method for an elective postgraduate data science and computer programming course. This course was delivered at the University of the West Indies in the MASc. Electrical and Computer Engineering programme. This study did not investigate student performance but investigated student interest and motivation in the course. This is part of a larger study in revamping the content, delivery and assessment of computer systems engineering courses at the Department of Electrical and Computer Engineering.

II. Literature Review

Gamification, the process of applying social gaming elements (e.g. role-playing, scoring systems) to real-world situations, has long been used in the education system to promote

978-1-5090-5920-1/17/$31.00 ©2017 IEEE

participant engagement by striking a balance between strictness and leniency. [5] suggests that the idea may have surfaced since the days of early Communism and further developed around the Soviet era, where it became favored over money as a work incentive. It then resurfaced in the U.S. around the eighties, and was utilized in many contexts, both academic and organizational, with promising results up until the twenty-first century.

Since its emergence, there have been numerous studies seeking to understand the effect of the use of gamification in curricular contexts. [6] acknowledges the potential for gaming elements to promote motivation and engagement, highlighting that the notion is seen even in more traditional contexts, such as the merit badge system used by the Boy Scouts of America. Here, goals, skill level, reputation, rank and accomplishments become motivating factors. However, participant interest must also be developed.

In some cases, as in [7], participant motivation and engagement can actually decrease over time, where gamification mechanics are supposedly improperly applied. [6] suggests that, in addition to gaming elements, game design must also be utilized to successfully capture participant interest. That is, the use of components such as leaderboards and badges alone is not enough; they must be incorporated into the course curriculum in some engaging, meaningful way to sustain student interest. It is then ultimately expected that where participants are kept both interested and motivated, course enrolment can be preserved. [8], for example, reported improved student immersion with gamification, but admitted that better incorporation of game aesthetics and an interactive game story could boost results.

Of course, the broader goal is to promote not just successful course completion but student retention and mastery of course material. [9] presents a framework describing the overall effect(s) of gamification: motivational factors effect some psychological response in the participant, which ultimately promotes a behavioral outcome. In the case of programming and software-related topics, it is expected that once students become motivated to learn the course content by gamification (motivational factor), their interest will be stimulated (psychological response), encouraging student engagement and inclination to practice course content (behavioral outcome) and eventual improved skill acquisition.

III. Experimental Setup

The postgraduate "Database Systems: Principles and Design" is an elective course in the M.ASc. Electrical and Computer Engineering programme at The University of the West Indies [10]. The course aims to introduce principles of database systems. Prior to the 2016/17 offering, the course focused primarily on a single database platform: Oracle [11] and created a stepping stone to obtaining a certification from that solution vendor. However, this course needed to be re-rationalized to align with the current local, regional and global industry needs. The new direction focuses on three Learning Outcomes (LOs) stated at the synthesis and evaluation levels of Bloom's Taxonomy, commensurate with higher cognitive expectations at the postgraduate level. By the end of the course, the student will be able to:

1) analyze, apply and design a database and DBMS for information processing in a particular application scenario;
2) understand, describe and analyze the concepts and issues; tools and technologies and current development trends related to information and database systems; and
3) extract and analyze, through design and implementation of appropriate scripts/programs, information for decision making from existing datasets/repositories.

The first LO keeps the database design and implementation principles while allowing a vendor agnostic approach. The second LO is added in recognition that a thorough understanding of the issues surrounding database systems will help the student adapt to rapidly changing needs and an ever-growing list of potential tools for use. Finally, an LO is added which takes into consideration the shift in the industry in recognizing the emerging frontier in data analytics.

In order to deliver a corpus of facts aligned with the overarching body of knowledge, the teaching strategy incorporates many existing learning objects. A semi-flipped classroom structure allows for alternating weeks of content delivered by course staff and guest speakers, as well as learner-driven exercises and status reports. Course information is disseminated via a campus-branded Moodle installation, myelearning [12]. Links to teacher and externally provided videos for self study are provided. The Animated Database Courseware (ADBC) [13] is used to generate in-class discussions on database design and implementation. Guest lectures from Microsoft and users of other tool-platforms from local industry are used to bolster interest and to discuss issues stemming from the business environment.

Typically, postgraduate students attempting this course would have an undergraduate degree from either Electrical and Computer Engineering, Computer Science, or Physics. However, undergraduate Electrical and Computer Engineering students are allowed to take this course as a final-year elective. However, this option is given to students either pursuing a Computer Systems thematic major or attempting a capstone project with database elements. It is noted no common programming language framework can be assumed. There is a large cult following for animated films (including Japanese anime), current pop-culture (eg. "Game of Thrones"), and "gaming" using RPGs as a social endeavour amongst engineering students and faculty at The UWI.

Formative assessments are in total given 30% of the final course grade, whereas two summative coursework projects make up the remaining 70%. While gaming, especially in MMORPGs, a player must "grind" in order to "level-up" their character by gaining experience points (XP). After the character has gained enough experience, the player can join a party (a group of similarly levelled characters, not necessarily with the same abilities) in order to challenge a "raid quest" or "boss fight". Raid quests are usually the real point of a game's plot, while grinding is a means to gain experience in order to attempt that challenge. This aligns with the formative vs. summative schemes found in education literature. During delivery students can be reminded what the true point of the game/course is - the skills gained and showcased in attempting the two summative

assessments, rather than the disproportionate amount of time spent grinding on formative assessments. This may improve student focus on the overarching goals/learning outcomes rather than getting lost in the details.

It follows that the formative assessments are individual, while the two summative assessments are group-based. The first summative assessment entails the design and build of a database system for a chosen application scenario, worth 30% of the final grade. The group selects a scenario where they must investigate the requirements of the system and end with a complete business application. They are expected to produce a report, worth 20%, detailing the:

- scenario: description with evidence of users'/customers' feedback/requirements elicitation
- design: design decomposition and diagrams (such as DFD's, ER-diagrams, logical schema/tables, story-boarding/GUI layouts)
- implementation and user manual, including an overview of tools used
- test cases and results of testing the system
- a discussion involving how salient issues, such as security, performance and storage, are accommodated

The remaining 10% is given for the implemented artefact (application source files, DB etc.). The party can select any platform (typically Azure w/SQL Server, MySQL, MS Access, or Oracle) for implementation. This coursework is assessed by the course lecturer - i.e. "the final boss". This particular assessment covers the first LO completely, and partially covers the third.

The second raid quest is worth 40% of the final grade. The party first finds an open dataset at the beginning of the semester in an interest area and then utilizes the tools and techniques learned about during the course to analyze that dataset. The idea here is to:

- collect (using a myelearning activity plugin) summaries of studies done in the topical area of the dataset. The summaries are done in a given template format (for scaffolding) - worth 5%.
- write a paper in a given journal style (10-15 pages) using ShareLaTeX or Google Docs worth 20%. The party must critically assess current trends in the literature review, present a methodology as per classwork, and interpret results.
- demonstration results (5%) and conduct an oral presentation (10%).

The course staff acts as non-player characters (NPCs) to assist with the two raid-quests. For example, the course TA acted as a "helper bot" whose elven character, "Hanneth", would log on to the party's document sharing system (Google Drive, ShareLaTeX, OneDrive, Dropbox etc.) and leave "helpful suggestions" at least once weekly. Most inter-communications of this nature were riddled with old-English speak through e-mail, myelearning, Facebook, WhatsApp, other social media and "sending a raven".

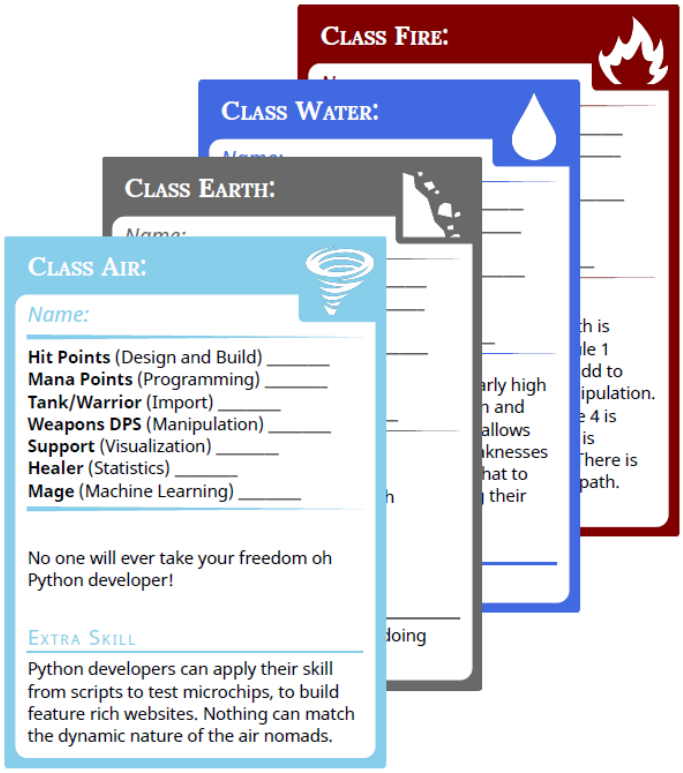

Fig. 1. Figure showing RPG cards

The 30% formative assessments must therefore support leveling up for the raid quests - 20% focused on skills for analyzing datasets while the remaining 10% focuses on skills in building databases. The raid quests must be attempted before the semester ends, leaving little time to grind over learning all concepts, issues and potential tools. Therefore, within the 5-6 member party, each specializes in one of four paths (as an "Avatar: The Last Airbender" [14] pop-culture reference) representing four major tools/languages: Water (SQL - high HP and MP), Earth (R - strong and dependable back-end), Fire (Azure ML Studio - those who seek powerful tools) and Wind (Python - freedom and flexibility of front-end). Each path is represented in each party. Students are mandated to become a master of at least one path for the course, however, students are further encouraged as part of continuing professional development to master all paths to become "The Avatar". Bonus marks are awarded as encouragement in Snape (Harry Potter) style at the discretion of the course lecturer.

RPG cards are generated with open fan-generated images to the back as shown in Figure 1. These cards serve as a mechanism to track the each character status. The parties, at regular points during the semester, verify that members are leveling at an appropriate rate. This encourages use of leaderboards and spurs competition. Feedback is taken, independent of the particular tool/path chosen for:

- Hit Points/HP - a metric indicating the ability of the character to design and build database schemas
- Mana Points/MP - indicating the ability to utilize programming languages (R or Python)

Factor	Mean
Resilience	4.08
Expectations	3.66
Effort	4.04
Time Management	4.46
Utility	3.5
Accomplishment	3.38
Attitude	4.22

Items - 1 - strongly disagree; 3 - neutral; 5 - strongly agree

TABLE I. SURVEY ANALYSIS: MEAN SCORE BY UNDERLYING FACTOR

- XP surrounding character development

In a time-constrained semester, it is difficult for a student to learn all data analysis functions prior to undertaking the raid quests. Therefore, character classes, adapted from "Log Horizon Elder Tales" [15], were created surrounding independent processes in data analysis: data importation (tank/warrior), manipulation (weapons DPS), visualization (support), statistics (healer), and machine learning (mage). Each member selected a different character class at the beginning to maximize early coverage of topics.

Formative assessment metrics/points are collected using staff generated quizzes on myelearning (re: database design), exercises on Datacamp [16] (students fill in R and python snippets in an online IDE which marks the submission and reports to the myelearning GradeBook module via an API), labs done using Azure ML Studio through the Microsoft Virtual Academy (MVA) programme course on Data Science and Machine Learning Essentials [17], and SQLite tutorials on Khan Academy [18].

In order to investigate the students' experience in the course, a survey, adapted from [4], was conducted when teaching concluded. Nineteen Likert scaled statements were developed with answers scaling from "Strongly Disagree" to "Strongly Agree". The factors which were investigated were the:

- expectations of DATARPG (4 statements)
- effort in DATARPG use (2 statements)
- time management in DATARPG use (4 statements)
- benefits perceived in DATARPG (2 statements)
- feeling accomplished in DATARPG use (4 statements)
- attitudes toward DATARPG (3 statements)

In order to ensure result validity, questions were checked for internal consistency by reversing some statements. Each response was coded 1 to 5, where 5 was a strong positive reaction and 1 a strong negative reaction with blank responses not allowed. There was also one open-ended question and one question to determine student resilience. Responses to the open question were coded and classified under one of the three motivation constructs: Personal Control, Ownership and Interest.

IV. INITIAL RESULTS

Qualitative feedback shows significant increase in student interest, compared to previous offerings. The average prior student enrollment is around 6 students per year (with a variance of about 3), while 28 students are enrolled in the current offering. Student performance has been shown to not take any significant deviation from prior offerings, despite the new delivery methods, content and larger class size.

	2010	2011	2012	2013	2014	2015	2016	**2017**
Enrollment	9	3	9	11	11	7	2	**28**
Pass Rate %	78	33	89	100	100	100	100	**100**
Average Grade %	66.25	28.6	61.6	76	78	70.7	76.5	**77.9**
Standard Deviation	11.57	26.2	20.8	14	9	10.8	12.5	**10**

TABLE II. STUDENT PERFORMANCE. 2017 (BOLD) USED THE DATARPG SYSTEM

V. DISCUSSION

It can be seen from Table I that students have a high positive reaction across all of the underlying motivational factors. The students perceive that this delivery method was of benefit. In particular, students indicated a very strong positive feeling towards how DATARPG system has affected their time management, and their attitude towards the underlying concepts in database systems. Also, from the open ended questions, students have given responses linked to motivational constructs of personal control and interest. However, there were no responses which linked to the construct of ownership.

The increase in enrollment is perhaps the best initial result. As can be seen, historical enrollment has been poor, with a dramatic increase in the current offering. Students are always given an opportunity to experience the course for at least three weeks before committing to completing it. This gave them exposure to the teaching methodology used. Further tracking would be required to determine if this was a "one-off" occurrence or a consistent improvement because of the alternate teaching method. At the time of writing this paper, the course has not yet concluded. Many points for future work remain incomplete. In particular, 1. collection of past motivation scores from prior offerings for comparative analysis, 2. re-collection this data at the end of the course may be of particular interest to see if there is any shift in attitude, 3. student feedback collection at the end of the course with their suggested methods for improvement, and 4. comparing across cohorts to understand the impact on student performance (in terms of final marks/grades).

The Department is also hosting an external examiner for QA feedback. After showcasing the course to her, one valuable feedback under consideration for incorporation during the next iteration is in adopting a short timed exam to better cover the third LO, and for proof of character levelling. This third individual quest is suggested to be worth 20%, taking 5% from formative, 5% from build a database raid quest, and 10% from the paper raid quest.

VI. CONCLUSION

In structuring the delivery and assessment of a traditionally low enrollment, elective STEM course to include elements of role playing games and pop-culture references seems to have global benefits to student motivation and interest. Further study is needed to ascertain whether this translates as an impact on end performance, and to rule out cohort specific effects.

REFERENCES

[1] G. I. Index, "Global innovation index 2015," 2015. [Online]. Available: https://www.globalinnovationindex.org/userfiles/file/reportpdf/GII-2015-v5.pdf

[2] R. Kalloo, "Relationships between primary teachers' and students' attitudes towards science, and students' attitudes and achievement of science process skills," Ph.D. dissertation, 2008.

[3] T. Koulouri, S. Lauria, and R. D. Macredie, "Teaching introductory programming: A quantitative evaluation of different approaches," *Trans. Comput. Educ.*, vol. 14, no. 4, pp. 26:1–26:28, Dec. 2014. [Online]. Available: http://doi.acm.org/10.1145/2662412

[4] J. Persad, D. Ringis, and C. Radix, "Leveraging student motivation in engineering skills acquisition," in *2014 IEEE Frontiers in Education Conference (FIE) Proceedings*, Oct 2014, pp. 1–4.

[5] A. Iosup and D. Epema, "An experience report on using gamification in technical higher education," in *Proceedings of the 45th ACM technical symposium on Computer science education*. ACM, 2014, pp. 27–32.

[6] S. Deterding, "Gamification: designing for motivation," *interactions*, vol. 19, no. 4, pp. 14–17, 2012.

[7] M. D. Hanus and J. Fox, "Assessing the effects of gamification in the classroom: A longitudinal study on intrinsic motivation, social comparison, satisfaction, effort, and academic performance," *Computers & Education*, vol. 80, pp. 152–161, 2015.

[8] M. Sillaots, "Gamification of higher education by the example of computer games course," in *The Seventh International Conference on Mobile, Hybrid, and On-line Learning (eLmL)*, 2015, pp. 62–58.

[9] J. Hamari, J. Koivisto, and H. Sarsa, "Does gamification work?– a literature review of empirical studies on gamification," in *System Sciences (HICSS), 2014 47th Hawaii International Conference on*. IEEE, 2014, pp. 3025–3034.

[10] "Master of applied science in electrical and computer engineering (masc)." [Online]. Available: http://sta.uwi.edu/eng/electrical/pgrad/masc.asp

[11] "Oracle integrated cloud applications & platform services." [Online]. Available: https://www.oracle.com/index.html

[12] "myelearning." [Online]. Available: http://myelearning.sta.uwi.edu/login/index.php

[13] M. C. Murray, "Animated database courseware." [Online]. Available: http://adbc.kennesaw.edu/

[14] "Avatar: The last airbender." [Online]. Available: http://www.imdb.com/title/tt0417299/

[15] "Log horizon elder tale character classes." [Online]. Available: http://loghorizonfanon.wikia.com/wiki/Classes

[16] "Datacamp." [Online]. Available: https://www.datacamp.com/

[17] "Data science and machine learning essentials." [Online]. Available: https://mva.microsoft.com/en-US/training-courses/data-science-and-machine-learning-essentials-14100?l=l90d5BXdB$_9$705050723

[18] "Khan academy." [Online]. Available: https://www.khanacademy.org