Gamification can Change the Traditonal Educational System*

Sachin Chhikara

April 18, 2024

In this paper a reproduction of the paper 'On the effectiveness of game-like and social approaches in learning: Comparing educational gaming, gamification & social networking' was done. We discoved that gamification while not as effective as intial thought, it has the pontential to be better than traditonal educational system especially when comes to assignments. These results are important as it suggested that students can perform even better academically.

1 Introduction

Games have been part of human history for a long time and are loved by many young and old. Games are able to keep players engage, motivate and give the feeling of accomplishment. Due to this, recently people have been studying games and trying to understand how to apply these concepts to non-game related fields, this is known as gamification. And it is used in many fields such as enterprise, sales, lifestyle, education and etc.

The focus of this paper will be an application of gamification in the education field. In this paper, we are doing a reproduction of paper called "On the effectiveness of game-like and social approaches in learning: Comparing educational gaming, gamification & social networking". In this paper, they took a control group, students using a traditional educational system and four treatment groups, which were students using different learning approaches which involved social and game elements. The focus of the study was to see if gamifcation was effective and this was achieve by looking at academic performance of the groups.And just like original paper we will focus on effective of gamifcation, using the grades of the students, which range from 0 - 100%. In addition, we will also be asking the question whether gamification works better for male or female.

*Code and data are available at: https://github.com/SachinChhikara/Effectiveness_of_Gamification_Education.

And what we found in this paper is that gamification may not be as effective for examinations, but it works well for assignment. Even then there is still room for improvement in the design and this can help to enliven student's education. And if we are to motive students this can help their lives and others, which can create a more positive impact on society.

This current section is introduction, the next section is data section where we discuss the data used and explore it. Then, we have a section where create a model out of such data and explain the reasoning behind our model design it. After that, we have results section where we talk about figures and tables that were replicate from the original paper and the model results. Lastly, we have the discussion section where discuss the results and how it connects to outside world.

2 Data

2.1 Data Source

This paper reproduces the result from called "On the effectiveness of game-like and social approaches in learning: Comparing educational gaming, gamification & social networking" (de-Marcos, García-López, and García-Cabot (2017)). This paper will provided a overview of the data used in that study. For more detailed explanation of the data please visit the original study.

2.2 Methodology

The original study used 379 undergraduates for their experiments and there were randomly assigned into 5 groups Control(N=76), Educational game(N=75), Gamification Plugin(N=77), Social networking website(N=75), Social gamified networking website(N=76). The experiment was carried out on 10-week first year undergraduate class called 'Qualification for ICT Users' with blended learning. This course is an introduction to basic computing concepts such as word processing, spreadsheets, presentations and databases.

In this experiment, before applying the corresponding learning methods, the students take a pretest, which occurs in week 1. And through week 2-9, depending on the group the student was assigned they learn using that specific method. During those weeks, they were assigned four assignments(aka post tests) on week 3, 5, 7 and 9. In addition, these assignments are on word processing, spreadsheets, presentations and databases, which tests students on practical skills. Finally, there is cumulative final exam and tests students on conceptual knowledge.

All groups had students have lecture every two weeks and must work independently rest of the work to learn then material using their respective method. In addition, Supplementary learning material was provided such as videos and communication tools in learning material platform. The control group just use traditional e-learning approach that was mention

above. Educational game used was called Ribbonhero offered by Microsoft as as free download to download Software, and player plays as virtual character Clippy and must overcome six challenges with increasing difficulty. Gamification plugin (Figure 1) is integrated in the institutional management system where there are challenges, trophies/badges and a leader board. Social Networking website was run on a open source engine, Elgg (Figure 2) and it functions very similar to social media where students can share their thoughts, make friends and discuss the material in class. Social gamified networking website use the same engine as the previous group and having similar features to Gamification plugin(Figure 3). A summary of the groups features and potential benefits(Table 1). The study highlights that the educational game is not necessarily aligned with the learning goals, while the gamification used in this study was.

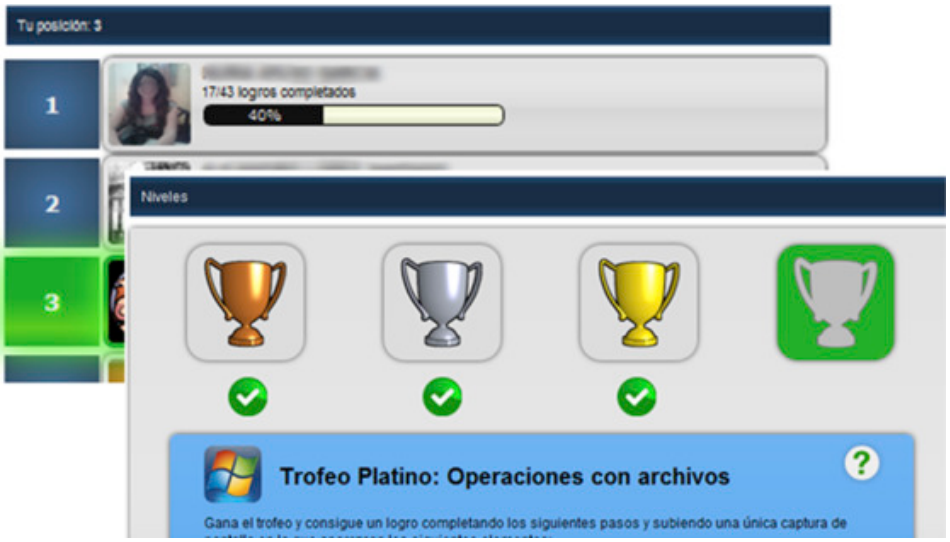


Figure 1: Gamifcation Plugin. Leaderboard and challenges.

2.3 Variables of Interest and Data Cleaning Procedure

The variables used in this study are Group, Gender, pretest and post test for Word Processing, Spreadsheets, Presentation, Databases and FinalExamination. Pretest, post test and FinalExamination use the measurement of a grade, which is from 0 - 100%.

The R package used for cleaning data were In this paper, the analysis is carried out using the statistical programming language `R` (R Core Team 2023), with the help of extra packages including `arrow` (Richardson et al. 2024), `kableExtra` (Zhu 2021), `gridExtra`(Auguie 2017), `knitr` (Xie 2014), `modelsummary`(Arel-Bundock 2022), `tidyverse` (Wickham et al. 2019).

For our replication, we drop 20 rows of data from the original dataset provided since those rows contained missing data on grades of some assessments and final examination. The age

Figure 2: Social Networking Website.

Figure 3: Social gamification website: Gamification tools (left), dashboard (top-right) & leaderboard (bottom-right).

Table 1: Summary of the groups used, their features and potential benefits.

Instrument	Features	Approach & targeted benefits
Educational game	Challenges, levels, points, narrative	Not aligned with learning objectives Independent work & exploration of students
Gamification plugin	Trophies, badges, challenges, leaderboard	Competition-driven Motivate participation through comparison with peers
Social networking website	Blogging, questions & answers, liking, friends, built-in twitter, dashboard	Cooperation and communication among participants. Boost participation, collaborative work & community building Promote student-driven discussion
Social gamification website	Blogging, questions & answers, liking, friends, built-in twitter, dashboard, challenges, points, achievements, virtual currency, shop, external rewarding, personalization (status/ visibility), peer review	Competition and cooperation Boost participation, collaborative work & community building Motivate participation through comparison with peers Social interaction affords additional means to motivate participation and engagement Addressing needs of different students (player types) and widen participation

row was removed since in the original study it mention that there age range for each group but in the dataset provided they made every value in that column to 19 since that the average age in the study. To summarize, the size of the dataset we use in this replication is N = 359(Control=75, Game=66, Plugin=73, Social Network=74, Social Gamification=71).

This dataset was chosen since my many studies in gamification don't have easy access to their datasets. In addition, this dataset includes control and treatment group which is useful for examining the difference between a traditional education system and gamification.

3 Model

Here we briefly describe the Bayesian analysis model used to investigate final examination marks and how it is influence by the group chosen for the student and the gender of the student. Background details are included in Appendix 7.

3.1 Model set-up

Define y_i as the the final examination mark, which is a number 0 and 100. $beta_0$ is the intercept and it is the expected value of the control group. $group_i$ is defined as the experiment groups: educational game, gamification plugin, social network and social gamification. $gender_i$ is defined is binary value that is either male or female.

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \beta_0 + \beta_1 \times \text{Group}_i \times \text{Gender}_i \tag{2}$$
$$\beta_0 \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta_1 \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\sigma \sim \text{Exponential}(1) \tag{5}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We used normal distribution for this model since our sample size was big enough(over 300 sample points), this allows to use the central limit theorem. We pick variable Group because we want to see how each experiment group hold up against our control group and we want to see if the game elements and social elements improve academic performance. We pick the gender variable since we want to see if there is a learning method that is more preferred by one gender over the other.

4 Results

Table 2: The results of the pretest.

Evaluation_item	Group	N	Mean	Std_err	Std_dev
Word Processing	Control	75	47.76	1.80	15.63
Word Processing	Ed. Game	66	45.83	2.53	20.56
Word Processing	Gamification	73	44.90	2.06	17.56
Word Processing	Social	74	45.19	1.80	15.51
Word Processing	Social Gamif.	71	44.77	2.50	21.06
Spreadsheets	Control	75	50.45	1.92	16.67
Spreadsheets	Ed. Game	66	51.50	1.90	15.46
Spreadsheets	Gamification	73	53.34	2.07	17.65
Spreadsheets	Social	74	51.30	2.16	18.54
Spreadsheets	Social Gamif.	71	52.37	1.87	15.72
Presentations	Control	75	44.11	1.61	13.99
Presentations	Ed. Game	66	45.33	1.45	11.82
Presentations	Gamification	73	45.21	1.60	13.67
Presentations	Social	74	44.59	1.39	11.92

Table 2: The results of the pretest.

Evaluation_item	Group	N	Mean	Std_err	Std_dev
Presentations	Social Gamif.	71	45.39	1.66	14.03
Databases	Control	75	51.80	2.09	18.06
Databases	Ed. Game	66	57.18	2.29	18.56
Databases	Gamification	73	54.86	2.03	17.38
Databases	Social	74	51.28	2.02	17.40
Databases	Social Gamif.	71	55.17	2.15	18.11

Table 3: The results of the posttest.

Evaluation_item	Group	N	Mean	Std_err	Std_dev
Word Processing	Control	75	64.37	1.95	16.90
Word Processing	Ed. Game	66	79.54	1.88	15.27
Word Processing	Gamification	73	74.11	2.81	24.05
Word Processing	Social	74	84.23	1.78	15.30
Word Processing	Social Gamif.	71	84.20	1.91	16.06
Spreadsheets	Control	75	61.74	3.04	26.32
Spreadsheets	Ed. Game	66	82.50	2.87	23.35
Spreadsheets	Gamification	73	79.93	2.75	23.50
Spreadsheets	Social	74	80.43	2.15	18.51
Spreadsheets	Social Gamif.	71	86.10	1.93	16.25
Presentations	Control	75	75.76	1.46	12.61
Presentations	Ed. Game	66	93.56	1.06	8.62
Presentations	Gamification	73	90.91	1.25	10.65
Presentations	Social	74	90.27	1.86	15.96
Presentations	Social Gamif.	71	86.78	1.27	10.68
Databases	Control	75	52.04	2.43	21.08
Databases	Ed. Game	66	55.48	3.94	31.97
Databases	Gamification	73	73.22	2.67	22.82
Databases	Social	74	70.86	2.41	20.75
Databases	Social Gamif.	71	66.85	3.15	26.57
Final Examination	Control	75	74.95	1.58	13.65
Final Examination	Ed. Game	66	66.87	1.55	12.59
Final Examination	Gamification	73	59.47	1.62	13.88
Final Examination	Social	74	72.43	1.78	15.34
Final Examination	Social Gamif.	71	61.88	1.84	15.54

The results from pretest(Table 2) are pretty such the same for each group which aligns with the fact that this study was given before the experiment began. When looking at the post

test results(Figure 4), the control group is lagging behind the other groups, especially in word processing, spreadsheets and databases.

In word processing post test(Figure 4), the group that perform the best were the social network and social gamifcation and roughly having the same performance. And then followed by educational game and gamification plugin. So it seems for word processing that groups with social elements are performing better than the group with game elements.

In spreadsheets post test(Figure 4), the group that perform the best was the social gamification, while the other three groups perform about the same and was only little bit behind social gamification.

In presentations post test(Figure 4), the group that perform the best was the educational game and not far behind was social network and gamification plugin. Surprisingly, social gamification was behind those three groups, as it would seem that the combination of social and game elements would make better than social network and gamification plugin.

In database post test(Figure 4), the educational game performed as worse the control group. And similarly to the presentation posttest the social gamification group does worse than the social network and gamification plugin. And gamification plugin does slightly better than social network.

For the final examination(Figure 5), interestingly enough, the group with the best results was the control group. Followed closely, by the social network group and then Educational game group. So the groups that performed the worse was gamfication plugin group and social gamfication group.

Our model results are summarized in Table 4. Social gamification group grades are almost comparable to the control group as on average they tend to be lower by 1-2%. While, a student in the educational game group, their grades on average are 6-7% lower than the control group. And then, a student in social network group, their grades on average 11-12% lower than the control group. Lastly, a student in the gamification plugin group, their grades on average are 15-16% lower than the control group.

Males overall on average did better than females by 2-3%. Also, males in the gamification plugin group on average did better that females by 1-2%. Interestingly, males do 2-4% worse than females in educational game, social gamification and social network group.

Table 4: Explanatory models of Final Exam marks based off of the experiment group and gender

	Final Exam Marks
(Intercept)	73.68
	(2.20)
GroupEducational game	−6.80
	(3.38)
GroupGamification plugin	−15.63
	(3.02)
GroupSocial gamified networking website	−1.27
	(3.16)
GroupSocial networking website	−11.30
	(3.10)
GenderMale	2.63
	(3.27)
GroupEducational game × GenderMale	−2.81
	(4.78)
GroupGamification plugin × GenderMale	1.22
	(4.83)
GroupSocial gamified networking website × GenderMale	−2.65
	(4.80)
GroupSocial networking website × GenderMale	−3.96
	(4.77)
Num.Obs.	359
R2	0.168
R2 Adj.	0.110
Log.Lik.	−1461.407
ELPD	−1471.2
ELPD s.e.	12.6
LOOIC	2942.4
LOOIC s.e.	25.2
WAIC	2942.3
RMSE	14.12

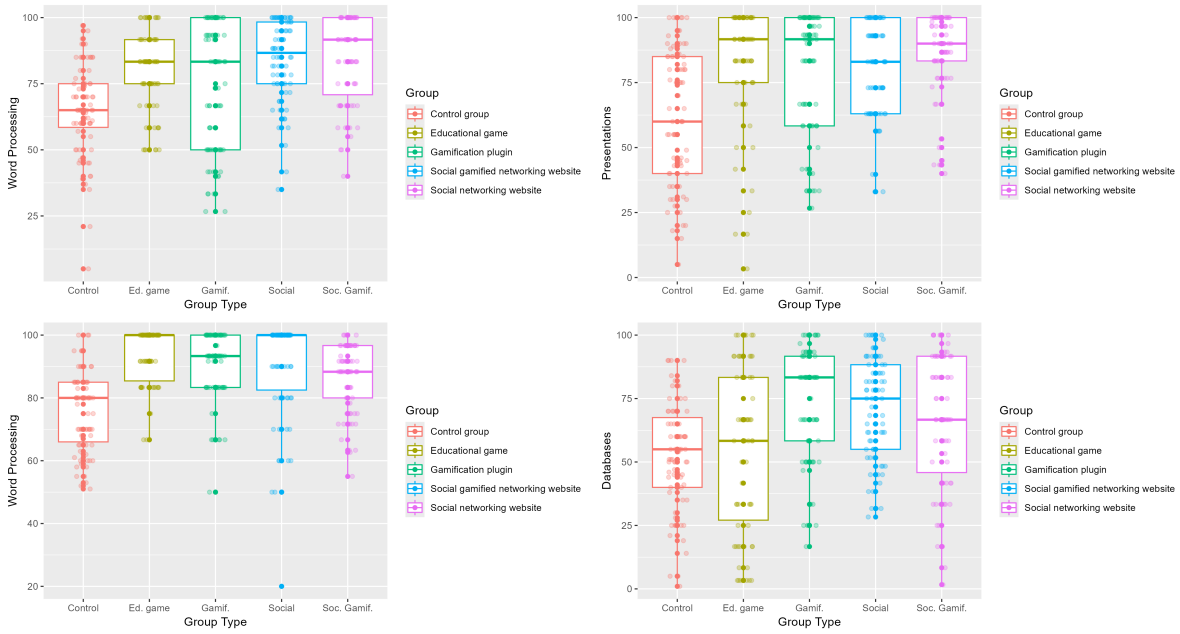


Figure 4: Boxplot showing the results of the post test for control and experiment groups

5 Discussion

5.1 Is Gamification effective? And is it better for male or females?

First, let us take a look at post test(Table 3) table again, recall that social gamification and social network does the best at word processing. Furthermore, social gamification does the best at spreadsheets as well. In addition, educational game does the best in presentations and gamification plugin does the best in databases. Notice that for post test, the best groups in each assessment seems to be related to groups that have game elements. Another thing to highlight is that social network group when comes to post tests tend to always be in top three groups. And social gamifcation is the best in 2/4 post tests. So it so seems that groups with social elements also perform well.

Gamification is not effective as we anticipated as seen in the final exam marks where the control group did the best. But, when comes to post tests as we note earlier groups with game elements were doing the best. And even for final examination social gamification was doing almost as good as the control group. From this study, we can conclude that gamification may not be good as we hope but it has potential to be better than traditional education. While, from our study we found males perform slightly better than females, the difference is pretty small(2-4%), which is good since both males and females can benefit equally from gamification. Furthermore, in order to improve gamification we will need to look at two things the design and student's feedback.

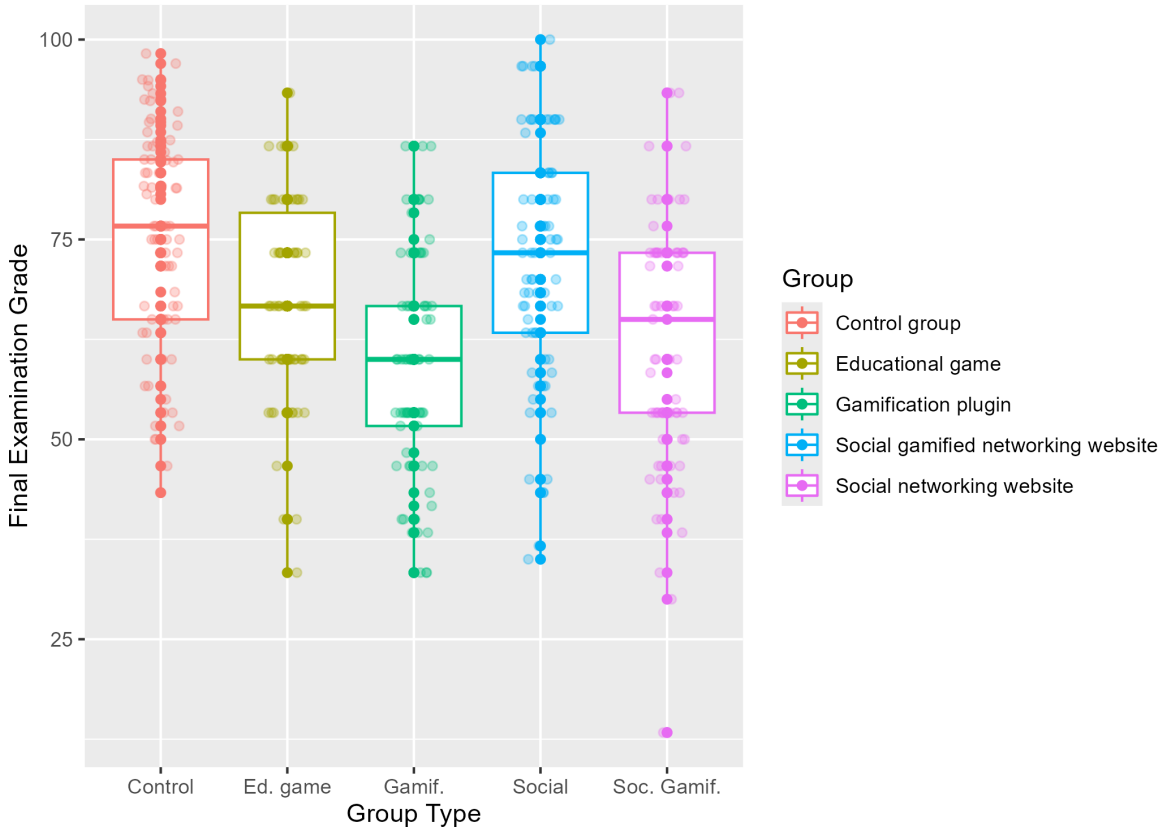


Figure 5: Boxplot showing the results of the final examination for each group that particapte in this experiment

5.2 Design of Gamification

Like I mentioned before gamification has to potential to change the traditional educational system, but in order to do that need to design it in a way that truly benefits the students. A leading expert in gamification and behavioral drive, Yu-kai Chou, has a create framework for gamification called the Octalysis Framework(Figure 6).Looking at the framework there are eight core drives: Epic Meaning, Accomplishment, Empowerment, Ownership, Social Influence, Scarcity, Unpredictability, and Avoidance. For even more in depth explanation, please visit Yu-kai website:https://yukaichou.com/gamification-examples/octalysis-complete-gamification-framework/.

Let us see how this framework applies to our study, for this example we will using social gamification group, since it one of the better performing groups with game elements. Epic Meaning category it is might but some students as they believe if they do well in this course, it could lead to some future goal like their dream job. Accomplishment is achieved by the meaningful challenges provided in each module and achievements. Empowerment is achieved by peer reviews and feedback from the teacher as well. Ownership is achieved as they can customize their profiles and status. Social Influence is achieved by being able to add friends and being able to features similar to Twitter(known as X now). This study has 5 out of 8 core drives mention from the framework.

Another example that we want to apply this framework is on a popular gamification language learning app, called Duolingo. Duolingo is a free app that has millions of user ranging from public students in third world countries to even Hollywood stars. Yu-kai Chou website provides an analysis for Duolingo, which has 6 out of 8 elements of the framework and here I am going to highlight some of the elements. Firstly, Epic Meaning is to learn a language and the app even asks the user when creating an account what is their purpose for learning new language. Secondly, Ownership, users have a currency call Lingots which allows players to buy Power-Ups(ways to make the player stronger) such as streak freeze and to buy clothes for their player character. Lastly, it has Scarcity and Avoidance as player have a limited amount of lives and if they lose then they can no longer play for a certain of time.

5.3 Limations and Weaknessess

Limitations of this dataset is that doesn't tell us how student's satisfaction, retention, motivation and engagement. This is would useful information since it will paint us better picture on the effectiveness of gamification. In addition, if students don't really like then it would be indicator whether gamification is not working or that something missing that could improve it. According to study about academic boredom (Özerk (2020)), some aspects of it comes from lack of motivation and learning methods, so if gamification can solve this would a huge win for students and for society at a whole. Furthermore, if students really enjoy gamification then this means this a potential valid learning method, which could convince others to fund and study further.

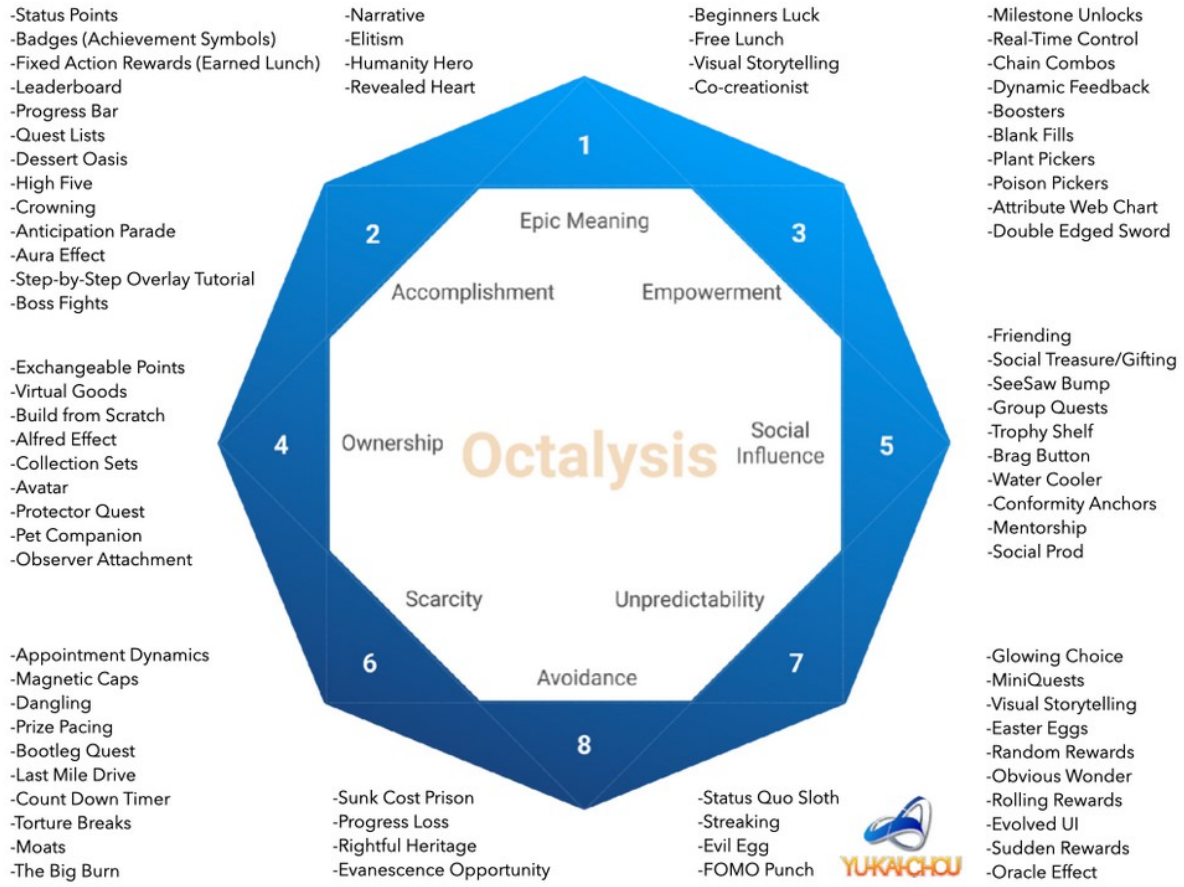


Figure 6: The Octalysis Framework for Gamification & Behavioral Design

Weakness of this dataset is that age column take the average age, which is 19 years in this case and gives every student that age. If the dataset gave us every student's age or even age range this could useful in determining which age group was benefiting more from gamification. It could that told us younger students better with gamification, but old students preform better with the traditional learning system.

Another weakness of this study is that study it reproduce is about 10 years old. This mean that data may not hold up as well. But there are not many gamification studies that compare treatment to a experiment group, and with open source data to boot. Also, Duolingo is popular language learn app, which use gamification and it has millions of active users which suggest that gamification is not a dead field and it is clearly effective for learning languages.

5.4 Future Studies

Like I mentioned before, this paper shows that gamification has potential to surpass traditional educational system. In future studies, more attention should be put into designing the gamification aspect. The Octalysis framework should be used and the design should it is best to meet all 8 core drives. Furthermore, Yu-kai is open to collaboration and consulting so that could make the design even better as there is more than just this framework that he talks about on his website. In addition, there should focus on social elements as well since that experiement did well in post tests. Lastly, this design is probably going to need multiple iterations and constant feedback from the students to further improve it and to allow it eventually surpass the traditional education system.

Appendix

A Additional data details

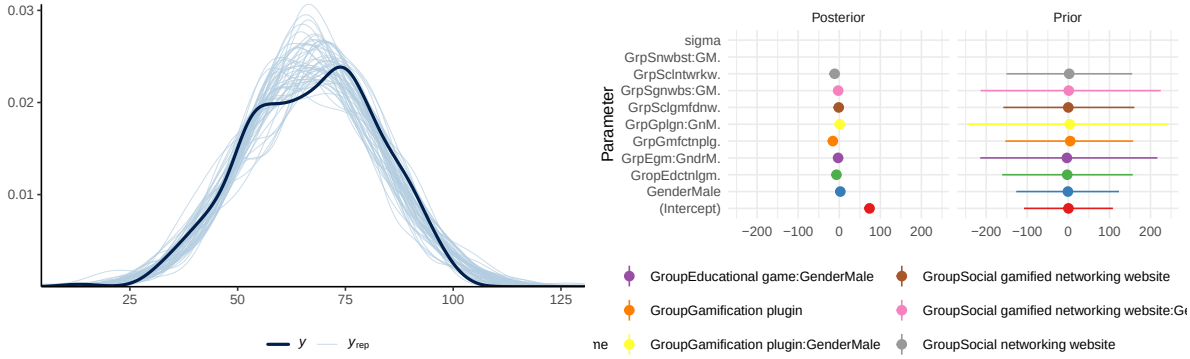
B Model details

B.1 Posterior predictive check

In Figure 7a we implement a posterior predictive check.

In Figure 7b we compare the posterior with the prior.

These two graphs to do a good job of showing that the data is properly fitted.



(a) Posterior prediction check

(b) Comparing the posterior with the prior

Figure 7: Examining how the model fits, and is affected by, the data

‘‘‘

References

Arel-Bundock, Vincent. 2022. "modelsummary: Data and Model Summaries in R." *Journal of Statistical Software* 103 (1): 1–23. https://doi.org/10.18637/jss.v103.i01.

Auguie, Baptiste. 2017. *gridExtra: Miscellaneous Functions for "Grid" Graphics*. https://CRAN.R-project.org/package=gridExtra.

de-Marcos, Luis, Eva Garcia-Lopez, and Antonio Garcia-Cabot. 2016. "On the Effectiveness of Game-Like and Social Approaches in Learning: Comparing Educational Gaming, Gamification & Social Networking." *Computers & Education* 95: 99–113. https://doi.org/https://doi.org/10.1016/j.compedu.2015.12.008.

de-Marcos, Luis, Eva García-López, and Antonio García-Cabot. 2017. "Dataset on the Learning Performance of ECDL Digital Skills of Undergraduate Students for Comparing Educational Gaming, Gamification and Social Networking." *Data in Brief* 11: 155–58. https://doi.org/https://doi.org/10.1016/j.dib.2017.01.017.

Gebru, Timnit, Jamie Morgenstern, Briana Vecchione, Jennifer Wortman Vaughan, Hanna Wallach, Hal Daumé Iii, and Kate Crawford. 2021. "Datasheets for Datasets." *Communications of the ACM* 64 (12): 86–92.

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

Horst, Allison Marie, Alison Presmanes Hill, and Kristen B Gorman. 2020. *Palmerpenguins: Palmer Archipelago (Antarctica) Penguin Data*. https://doi.org/10.5281/zenodo.3960218.

Özerk, Gül. 2020. "Academic Boredom: An Underestimated Challenge in Schools." *Lnternational Electronic Journal of Elementary Education* 13 (October): 117–25. https://doi.org/10.26822/iejee.2020.177.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Richardson, Neal, Ian Cook, Nic Crane, Dewey Dunnington, Romain François, Jonathan Keane, Dragoș Moldovan-Grünfeld, Jeroen Ooms, Jacob Wujciak-Jens, and Apache Arrow. 2024. *Arrow: Integration to 'Apache' 'Arrow'*. https://CRAN.R-project.org/package=arrow.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.

Xie, Yihui. 2014. "Knitr: A Comprehensive Tool for Reproducible Research in R." In *Implementing Reproducible Computational Research*, edited by Victoria Stodden, Friedrich Leisch, and Roger D. Peng. Chapman; Hall/CRC.

Zhu, Hao. 2021. *kableExtra: Construct Complex Table with 'Kable' and Pipe Syntax*. https://CRAN.R-project.org/package=kableExtra.