Gamification Effectiveness on Education*

Sachin Chhikara

April 17, 2024

First sentence. Second sentence. Third sentence. Fourth sentence.

1 Introduction

Games have been part of human history for a long time and are loved by many young and old. The design behind why game are so fun have investigated in recent time and how to apply those element that are fun and to apply them to other fields, this is known as gamification. Gamification is used in fields such as enterprise, sales, lifestyle and education. In this paper we will focusing on application of education, we be using data from Duolingu, training model and making inference off, real life application of gamification to an individual productive.

In this paper, we are doing replication of a study … , which has a control and treatment where the treatment is applied gamification, while the control is using an ordinary education system. Yeah there benefits, but there are downside when comes to how the gamification as it might discourage learning. Furthermore, an wide use application of gamification is a language learning app called Duolingo, which has … active users and this app learning is recognized in certain Universities as legitimate English learning.So we examine game design that are use. Ways to apply gamification in daily lives, like this app called Habitica: Gamify your tasks.

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

*Code and data are available at: LINK.

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.

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%.

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

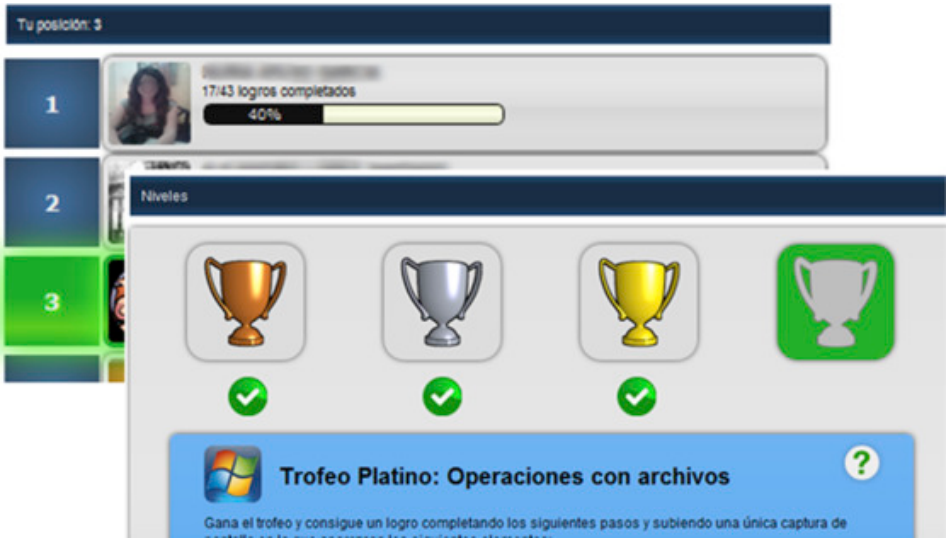


Figure 1: Gamifcation Plugin. Leaderboard and challenges.

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 and diagnostics are included in **?@sec-model-details**.

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.

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

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
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

Table 3: The results of the posttest.

Evaluation_item	Group	N	Mean	Std_err	Std_dev
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 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.

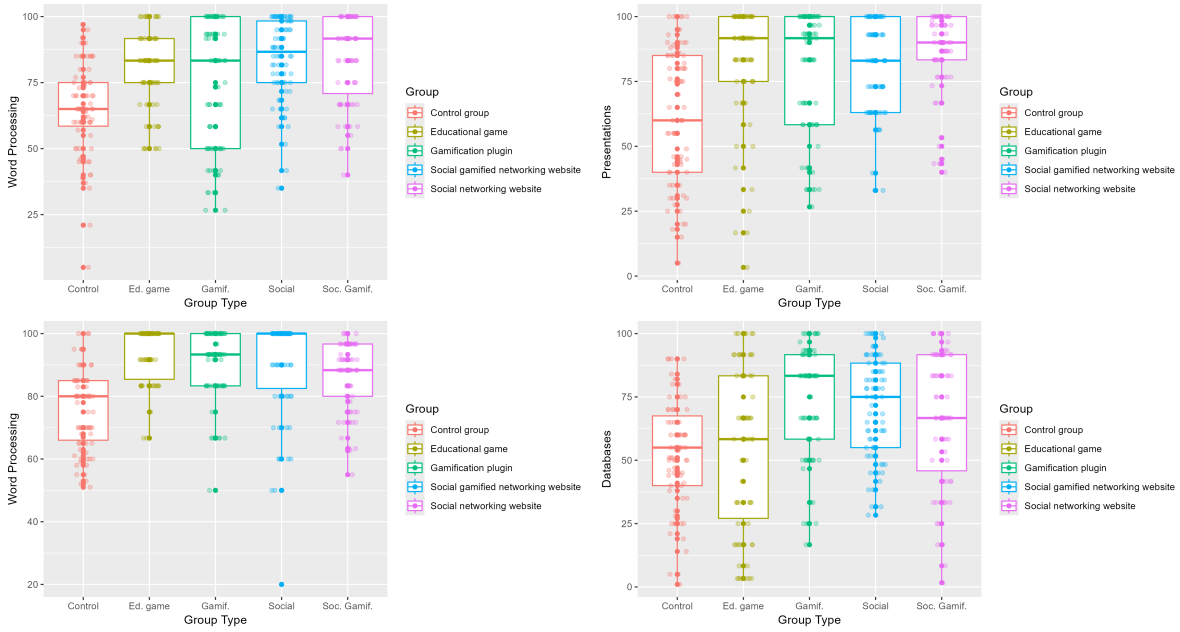


Figure 4: Boxplot showing the results of the post test for control and experiment groups

5 Discussion

5.1 Is Gamification effective? Does it have pontential to change the educational system?

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

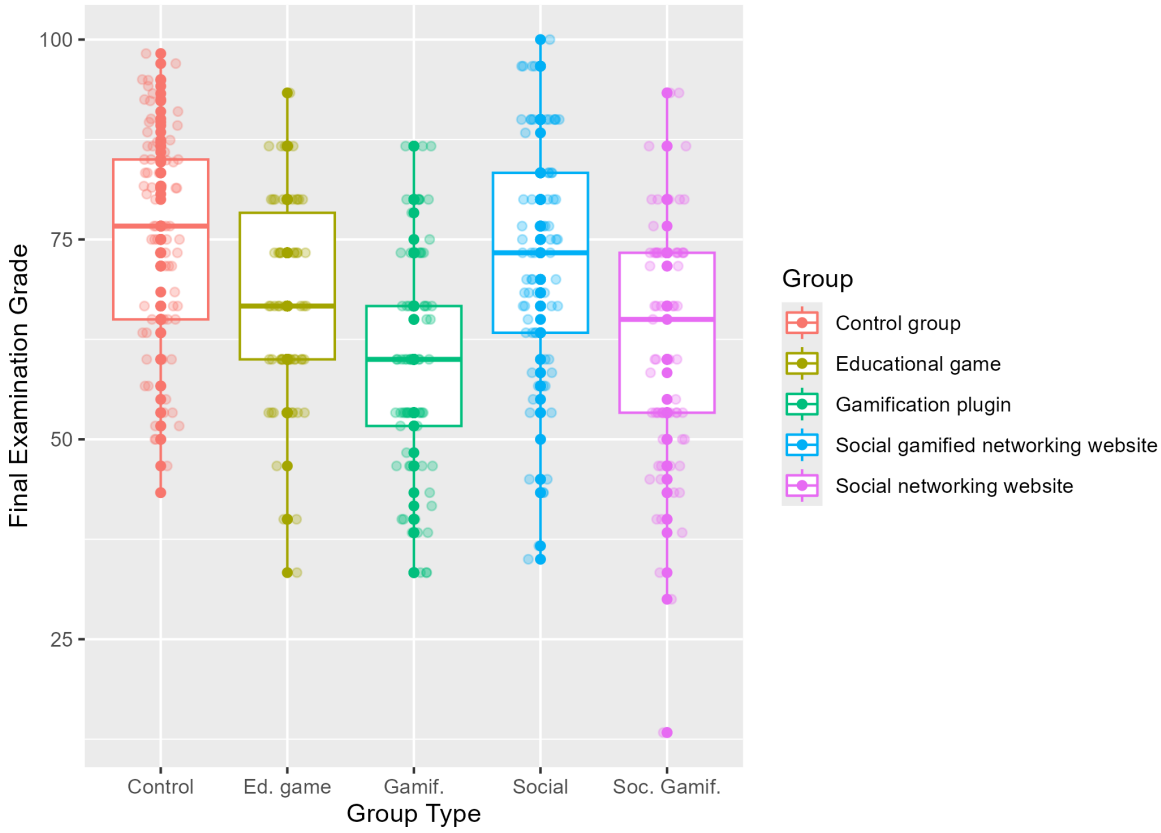


Figure 5: Boxplot showing the results of the final examination for each group that particapte in this experiment

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

5.2 Design of Gamification

5.3 Student's input

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

In **?@fig-ppcheckandposteriorvsprior-1** we implement a posterior predictive check. This shows...

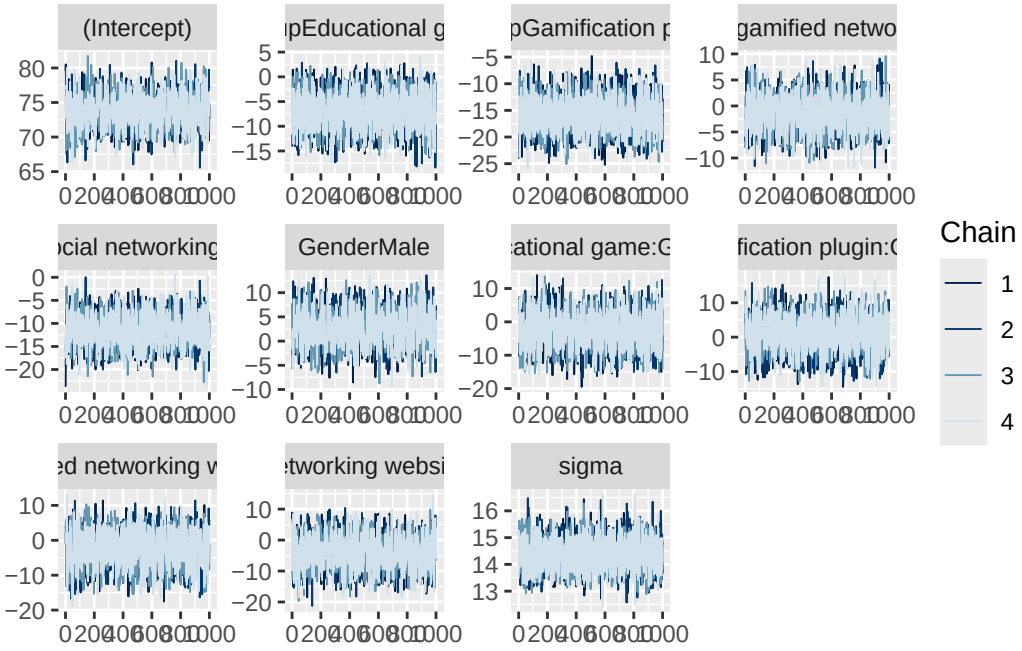
In **?@fig-ppcheckandposteriorvsprior-2** we compare the posterior with the prior. This shows...

Examining how the model fits, and is affected
by, the data

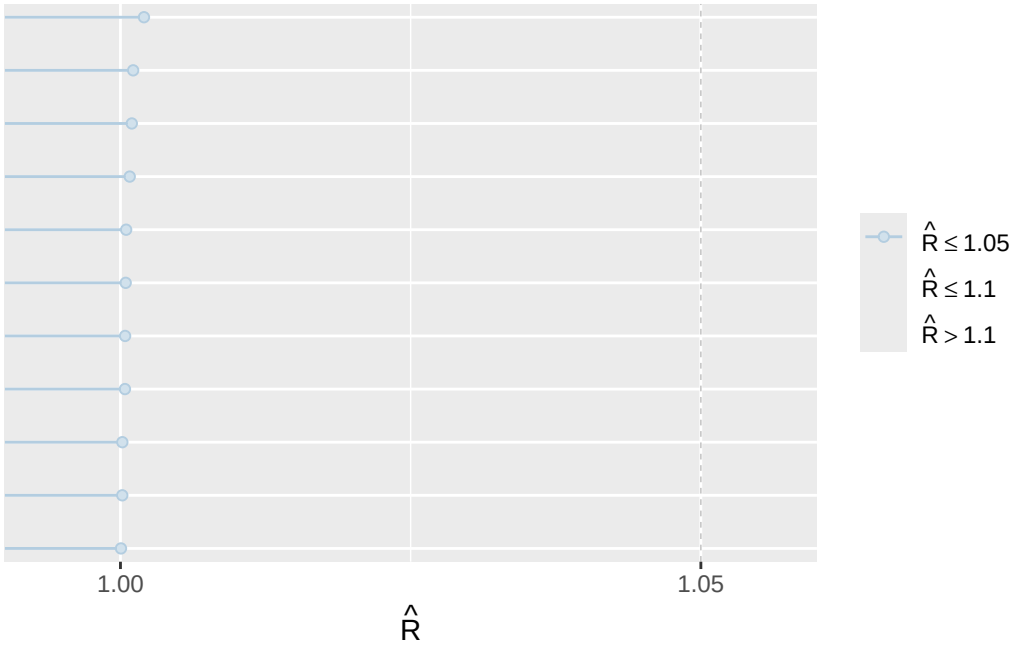
B.2 Diagnostics

Figure 6a is a trace plot. It shows... This suggests...

Figure 6b is a Rhat plot. It shows... This suggests...



(a) Trace plot



(b) Rhat plot

Figure 6: Checking the convergence of the MCMC algorithm

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.

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.

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.