Learning from Errors: An Empirical Study on the Impact of Gamification on SQL Query Formulation

Anonymous Authors
abc@abc.com
Anonymous University
Anonymous, Anonymous

ABSTRACT

The application of gamification is becoming increasingly popular throughout all areas of education claiming to impact student motivation and learning experience positively. With the aim of providing empirical evidence on the effects gamification has on students, this paper examines how gamification influences the errors students make when learning Structured Query Language (SQL). An empirical study was carried out with 133 students enrolled in a Database course from the Anonymous University, where participants were split into three groups (gamified, competitive, and non-gamified) to understand if the environment influences the number and types of errors made by students. After analyzing 1,009 answers submitted by the students, it was observed that (i) students participating in the gamified environment tended to make fewer semantic errors; (ii) the competitive environment yielded better results reducing the prevalence of syntax errors; and (iii) the most common errors are concentrated in the SELECT and WHERE clauses. Overall, while the observed differences between non-gamified and gamified environments were minimal, it was concluded that gamification (and competitiveness) positively influences specific aspects of students' learning.

CCS CONCEPTS

• General and reference → Empirical studies; **• Social and professional topics → Software engineering education**; **Computer science education**; **• Information systems → Structured Query Language**.

KEYWORDS

Gamification, Game-based learning, Databases, Computer Science, Empirical Study

ACM Reference Format:
Anonymous Authors. 2023. Learning from Errors: An Empirical Study on the Impact of Gamification on SQL Query Formulation. In *Proceedings of 28th annual ACM conference on Innovation and Technology in Computer Science Education (ITiCSE'23).* ACM, New York, NY, USA, 7 pages. https://doi.org/XXXXXXX.XXXXXXX

ITiCSE'23, July 10–12, 2023, Turku, Finland

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

1 INTRODUCTION

Defined as the application of game-based elements and principles in non-game contexts [20], gamification has been growing in interest across all levels and areas of education. Mathletics (maths) [9], Reading Eggs (literacy) [3], and Code Avengers (programming) [2] are some examples of resources used in education that implement game-based elements such as leaderboards, points, and badges to motivate students to acquire and develop new knowledge and skills. While it is expected these game elements improve student motivation, enthusiasm, and fluency in the taught subject, there is a lack of evidence to support these claims, hindering its acceptance as an effective learning and teaching strategy [6, 13].

Gamification is not a stranger to computer science, CS Unplugged [1] is a free open-source resource that utilizes games and puzzles to teach computer science fundamentals. In the particular case of Databases, SchemaVerse [4] is an online space-based strategy game in which players compete against each other by using Structured Query Language (SQL) statements to command their fleet.

SQL is a programming language that became the global standard for storing, manipulating, and retrieving data in relational databases. Relational databases and SQL are fundamental topics in computer science and software engineering education, resulting in it being taught in almost all university-level database courses. Despite SQL being the standardized language for relational databases, it has relatively little attention in educational research [19].

With the aim of understanding and demonstrating the effects of gamification on students' learning processes, this paper analyzes the impact gamification has on SQL education in undergraduate students enrolled in database courses. Particularly of interest for this research is how gamification influences the errors made by students when learning SQL. For this purpose, an empirical study was conducted with undergraduate students learning SQL using AnonymousTool (AT-AT) [7, 8], a gamified application specifically designed for learning and practicing SQL.

This paper is organized as follows: Section 2 presents an overview of the related work around gamification and error classification in SQL. Section 3 describes the empirical study carried out. The results are presented in Section 4 followed by its discussion in Section 5. Lastly, the conclusions and future work are presented in Section 6.

2 RELATED WORK

2.1 Gamification and SQL

Gamification has become widely used throughout education with limited empirical evidence of its impact on learning [7]. In [7], authors conducted an empirical study into the impact gamification has on student performance, motivation, and user experience in a database course. Their results showed that gamification improved student performance and increased motivation but slightly decreased user experience. This paper also collated a literature review (up to 2020) on databases and gamification highlighting that research around the use of game-related methods applied to databases is evident and that previous studies around these topics focus highly on psychological areas such as motivation and engagement rather than the measurable aspects such as performance.

Following the publication of this paper, new empirical studies around gamification in database education have been completed, and a systematic literature review to complement the one presented in [7] was carried out for this paper. The review showed that, since 2020, gamification and SQL have been continuously studied from different perspectives. In 2021, Indriasari et al. [14] studied the impact gamification had on students' behavior in peer code review activities. This study resulted in the gamified group submitting more peer reviews and elevated student motivation. Zorrilla Pantaleón et al. [21] carried out a study focused on extrinsic motivation and identified what gamified elements out of personal rewards or leaderboards had a more significant impact on participant performance. Also, they found that gamification resulted in extrinsic motivation for participants, and leaderboards significantly impacted performance but negativity affected motivation.

Lastly, in this review, as mentioned in [7] the combination of points, badges, and leaderboards tends to be the most commonly used subset of game design elements [12].

2.2 Complications and causes behind SQL errors

SQL is a widely taught database query language in computer science, data science, and software engineering programs which is highly expressive but challenging to learn for novices [16]. Understanding how students learn SQL is a nascent research topic, unlike the wealth of research on students' difficulties in learning imperative, object-oriented, or functional languages, there has been little research on how students learn SQL [17].

After performing an ad-hoc literature review, previous related work was identified. A search on "sql AND errors" in Scopus retrieved 739 results of which 35 candidate papers were filtered through pre-defined inclusion and exclusion criteria. This search resulted in six primary papers being identified as relevant to this research.

In 2022, Miedema et al. conducted an analysis to understand how students deal with complexity during query formulation [15]. The authors examined query responses to six SQL exercises from 104 students. SQL submissions were categorized as correct, semantic errors, syntax errors, and timeouts. It was found that students mismanage complexity in SQL by trying to persevere through syntactic and semantic errors, leading them to overly complex answers. This resulted in SQL educators being urged to integrate SQL problem decomposition in their lectures.

Taipalus analyzed 12,000 SQL errors collected over four years from undergraduate students enrolled in an introductory database course [18]. It was found that logical errors are the most common cause of query formulation failures in novice learners. The results showed that missing expressions, extraneous or omitted grouping columns, incorrect comparison operators, missing joins, and missing ordering columns are the most common errors that novices are unable to fix.

Derived from the aim to identify the sources of complications of learning SQL, researchers have come up with several categorizations of errors; the most common and top-level classes are the semantic and syntax types of errors. Semantic errors are defined as a legal SQL query that does not (or not always) produces the intended results and is therefore incorrect for the given task [10]. Syntax errors, on the other hand, are defined as errors where the SQL engine was not able to execute the query and thus returned an error code [17].

In 2018, Taipalus et al. categorized SQL query errors that students made in an undergraduate database course [19]. The study analyzed 30,000 SQL queries and defined three categories of SQL errors: syntax, semantic, and logical. The authors concluded that error categorization should be used by educators to help understand problems students face when learning SQL.

A similar study to [19] was published by Poulsen et al. in 2020. Their research analyzed 286 students' SQL homework submissions to understand the difficulties students have in learning SQL [17]. SQL errors were defined and categorized as syntactical and semantical. It was observed that syntactic errors were the majority and that errors relating to GROUP BY and correlated subqueries were the most difficult for students to overcome. Like [19], Poulsen et al. concluded that instructors can gain insight to understand and address common students' misconceptions.

A significant contribution to this area that many studies build on is Brass and Goldberg's categorization into classes of SQL queries that are syntactically correct but not correct for the task the query was written for [10]. This research defined and provided a detailed list of 43 semantic errors in SQL queries.

Miedema et al. identified 14 misconceptions that fall under four categories that cause SQL errors [15]. It was concluded that interference of previous course knowledge, generalizing answers to questions where they do not fit, differences in language usage between SQL and natural language, and incomplete or incorrect mental models are the causes of SQL errors for novices. The analysis split semantic errors into seven sub-categories, "incorrect or missing table/column", "incorrect or missing keyword", "returning incorrect results", "issues with the schema", "alias problems", "contractions", and "complications".

In 2019, Casterella et al. investigated the impacts that two different intervention techniques had on undergraduate students learning SQL [11]. This paper identified SQL query errors as "Query Structure Errors" and mapped these by which clause the error occurred. This mapping was done to compare the impact on the different types of interventions and tasks applied.

For the purposes of this paper, errors were classified into semantic and syntax as most of the identified previous work did [15–19]. In addition, as in [11], a structural approach was used to identify which the most "problematic" SQL clauses are for students.

3 RESEARCH METHOD

3.1 Overview

The research question this paper aims to answer is: *How does gamification influence the errors students make while learning SQL?*

The motivation for answering this question is to understand if the errors made by students when learning in a game-based environment are different from those made in a non-gamified one. Identifying the potential differences will help educators to apply new or modify existing teaching strategies with the aim of (i) improving the learning outcomes of students when using gamification, and (ii) tackling more efficiently the most common SQL errors made by students.

To the best of the authors' knowledge and from the systematic and ad-hoc literature reviews carried out on gamification and SQL (Section 2.1) and SQL errors (Section 2.2), no previous publications about the impact gamification has on errors made by students learning SQL were found.

3.2 The tool: AT-AT

For the purpose of providing students with an effective environment to practice and improve their SQL abilities, AT-AT was developed based on three game elements: challenges (queries to be solved by students), points, and leaderboards. Researchers ran an empirical study in 2019 [7], showing that students using AT-AT obtained better results in motivation and user experience than students using a non-gamified version of the application. More importantly, they found that students using the gamified version doubled the differential of performance compared to the non-gamified group, i.e., gamification had a positive impact on students' performance.

In order to understand the effects of gamification and competitiveness on students, a new version of AT-AT was developed, and the empirical study presented in [7] was replicated.

The newly developed version of AT-AT allows instructors to create different environments by adding or removing the available game elements (i.e., points, leaderboards, and challenges) and enabling or disabling the mechanics of the game (e.g., timers and opening/closing times). In this way, instructors are able to set a gamified environment where points are earned but no time limits apply; or alternatively, a competitive environment in which points are shown on leaderboards and students have to submit their answers by a fixed date.

3.3 Empirical Study Design

Since the goal of this research was to understand the effects of gamification on the errors students made when learning SQL, three environments were set in AT-AT splitting the students into groups: gamified (G), competitive (C), and non-gamified (N). These environments are described in Table 1.

The study ran between May and June 2022. Participants for this study were students enrolled in a Database course at the Anonymous University. The course had 161 students enrolled and it was taught by one lecturer with extensive experience teaching the course and three casual tutors. Students were randomly split into three groups (see Table 1) and data was collected while they used AT-AT during a four-phase study described below.

Table 1: Empirical Study Groupings

Group	AT-AT Environment
G	Queries are shown to students and they earn points if their answers are correct.
C	In addition to earning points, a leaderboard presents the current place the student sits in the group.
N	A plain instance where queries were treated as "homework" and students got feedback if their answers were correct or not.

Replicating the original study described in [7], the empirical study was designed to have four major phases: a pre-test, a core stage, a post-test, and a survey. All four stages were successfully completed, however, for the purposes of this paper, the pre-test and core stages will be described.

- **Pre-test:** A 'test' all participants sat containing 10 SQL queries. The queries were ordered by difficulty (0 – extremely easy to 9 – extremely hard) and were designed to identify the participants' current SQL skill level. The participants were given one hour to solve them.
- **Core stage:** It consisted of three instances, one per environment: G, C, and N. Each instance was composed of three rounds. In each round, the students had to answer three queries. The first round included queries from levels 1 to 3 (Easy), the second round from levels 4 to 6 (Medium), and the third round from levels 7 to 9 (Hard). The rounds ran during three consecutive weekends and were open for 48 hours (one round per weekend). The expected structure of the queries is presented in Table 2.

Table 2: Expected query structure per level

	SQL keywords
L1	SELECT FROM WHERE
L2	SELECT FROM GROUP BY
L3	SELECT FROM WHERE (complex JOINs and conditions)
L4	L3 plus L2 as nested query
L5	L3 plus UNION plus L3
L6	L3 plus GROUP BY and ORDER BY
L7	L3 plus GROUP BY and HAVING
L8	L7 plus L4 as a nested query
L9	L8 plus L7 as a nested query

This research was approved by the Ethics Committee of the Anonymous University (Reference: ABCD 2022/00/00).

3.4 Data collection and analysis

Students' answers were stored in AT-AT's database once the student clicked 'Save' on the web application. Information such as time opened, time submitted, and the student's response was recorded for each answer.

Once the pre-test or a round closed, the marking of the student answers was automated through a script. This script marked each submission against a model answer by comparing the query keywords and the expected and returned result sets.

For the purposes of this paper, the answers marked as incorrect are the unit of study. An answer was considered incorrect if: (i) the expected and returned result sets were different; or (ii) the expected keywords were not present in the student's answer (e.g., the query explicitly required the use of UNION and the student didn't use it), even if the result sets were equal.

The answer data was exported from the database to CSV files (one file per group and per round) and then cleaned in Excel. Empty and not related answers (e.g., comments like "I don't know") were removed from the answers and not considered for the analysis. Once the incorrect answers were cleaned, these were categorized into Semantic or Syntax errors. Similarly to [5], in this study, it was decided to not differentiate logical from semantic errors. Syntax errors were automatically flagged by the marking script and then manually classified, along with the semantic errors. The classification was performed by two research assistants and validated by the authors of the paper.

After the data was cleaned and classified, it was imported back into a database where the distribution of such errors across the different environments and types of queries was analyzed. DataGrip was used for running an SQL script for extracting reports and Excel was used for creating plots and calculating descriptive statistics of the data set.

4 RESULTS

4.1 Groups and Answers

In the study, 133 students participated (out of the 161 enrolled in the course) who were randomly split into groups G, C, and N (see Table 3).

The pre-test results were used for comparing the homogeneity of the groups in terms of actual SQL knowledge. The median of correct answers was 2 for all groups, while the means were 2.20, 1.92, and 2.05 for G, C, and N respectively. The standard deviations for these environments were 1.254, 1.115, and 1.234; these low and similar values show that data points are clustered around the mean. Figure 1 shows the distribution of correct answers in the pre-test for each group.

Based on the pre-test results, it can be observed that the three groups had similar knowledge levels of SQL at the beginning of the study and there were no significant differences between them.

In total, 1,009 meaningful answers (i.e., not empty and SQL related) were submitted during the core stage. After marking them, 436 were labeled as correct and 573 as incorrect (i.e., these had syntax and/or semantic errors).

The incorrect answers were analyzed in order to understand:

- The type of error (i.e., semantic vs. syntax)
- The location of the errors (i.e., where in the SQL statement the error was located)
- The relation between the complexity of the query and the errors made

The analysis of these aspects is presented in the following subsections.

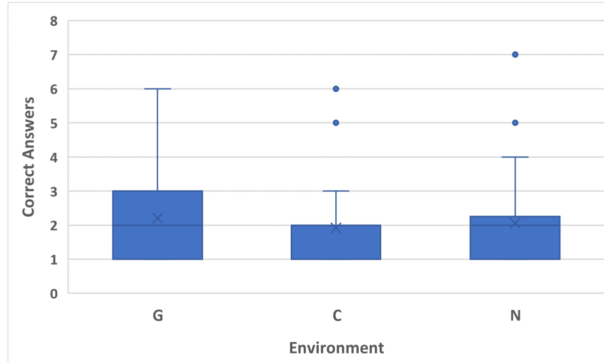

Figure 1: Pre-test correct answers by environment.

Table 3: Groups and answers overview

	G	C	N	Total
Participants	47	41	45	**133**
Meaningful answers	363	306	340	**1,009**
Correct answers	154	135	147	**436**
Incorrect answers	209	171	193	**573**
Percentage of correct answers	42.4	44.1	43.2	**43.2**
Percentage of incorrect answers	57.6	55.9	56.8	**56.8**

4.2 Analysis by Type of Error

Six hundred and sixty-one errors were found in the 573 incorrect answers. Out of these, 457 were semantic errors and 204 were syntax errors. An overview of the errors per environment and type is presented in Table 4.

Table 4: Errors overview

	G	C	N	Total
Total errors	230	194	237	**661**
Semantic errors	146	148	163	**457**
Syntax errors	84	46	74	**204**
Percentage of semantic errors	31.9	32.4	35.7	
Percentage of syntax errors	41.2	22.5	36.3	
Percentage of incorrect answers with semantic errors	69.9	86.5	84.5	
Percentage of incorrect answers with syntax errors	40.2	26.9	38.3	

Semantic Errors. The distribution of the 457 semantic errors was very similar across environments (G: 31.9%, C: 32.4%, N: 35.7%). However, a considerable difference was observed in the G environment where 69.9% of the incorrect answers had semantic errors compared with 86.5% and 84.5% from the C and N environments respectively.

The most common semantic errors were those related to missing, wrong or extra columns (200 errors), followed by those caused by missing, wrong or extra joins (121), see Table 5.

Syntax Errors. It was observed that the number of syntax errors in the G environment was greater than in the N environment and

Table 5: Most common semantic errors

	G	C	N	Total
Missing/wrong/extra columns	68	70	62	**200**
Missing/wrong/extra join	38	39	44	**121**
Missing/wrong/extra operator	14	14	17	**45**
Wrong comparison/values	14	9	15	**38**
Wrong ordering	6	8	12	**26**
Other	2	6	6	**14**
Function-related	4	2	7	**13**

almost doubled the ones found in the C environment. Out of the 204 syntax errors, 41.2% were found in the G environment (84 errors), followed by the N environment (36.3%, 74 errors) and the C environment (22.5%, 46 errors). Regarding the percentage of questions with syntax errors, the C environment has the least (26.9%) followed by the N (38.3%) and G (40.2%) environments.

The top two syntax errors were caused by incorrect field names and invalid usage of keywords (52 errors each). While the number of field name-related errors across environments was similar (G: 20, C:15, N:17), it is interesting that invalid keyword errors were less likely to happen in the C environment (8 errors) than in the G (25) and N (19) environments.

Also, schema-related errors (e.g., wrong datatype, too many columns) were considerably higher in the G environment (15 errors).

Table 6: Most common syntax errors

	G	C	N	Total
Field names	20	15	17	**52**
Invalid keyword/clause	25	8	19	**52**
Function-related	13	9	14	**36**
Join/table related	5	9	11	**25**
Schema-related	15	3	5	**23**
Comparison-related	4	2	4	**1**
Other	2	0	4	**6**

Regarding SQL error codes thrown by the DBMS (MariaDB in this case), 14 different errors were found. The top-three most common SQL errors were the 1064: "You have an error in your SQL syntax" 113 times (G:45, C:21, N:47); the 1054: "Unknown column" (G:17, C:7, N:13) 37 times; and the 1060: "Duplicate column name" (G:6, C:3, N:1) 10 times.

4.3 Analysis by Error Location

Regarding where the errors were most commonly located, the SELECT and WHERE clauses concentrated 65.8% of the errors (33.7% and 32.1% respectively); the FROM clause came third with 22.4% (see Table 7). While this makes sense since SELECT-FROM-WHERE is the basic and most used structure in SQL, the same pattern was repeated in nested queries, i.e., the location of the SELECT/FROM/WHERE clauses within the query made no difference in terms of the error distribution (33.3%, 26.0%, and 33.3% for nested queries respectively).

However, a considerable difference was observed between the semantic and syntax errors when their location was considered. 37.9% of the semantic errors were found in the SELECT clause, 30.0% in the WHERE clause, and 21.9% in the FROM clause. The syntax error location behaved differently as 36.8% of the errors were found in the WHERE clause, while 24.5% and 23.5% were in the SELECT and FROM clauses respectively.

Table 7: Location of errors within a query (in percentages)

	G	C	N	Total
SELECT	11.1	12.0	10.6	**33.7**
FROM	7.4	6.2	8.8	**22.4**
WHERE	12.3	8.2	11.6	**32.1**
GROUP BY	1.1	1.1	1.5	**3.6**
HAVING	0.6	0.6	0.3	**1.5**
ORDER BY	1.1	0.3	1.4	**2.7**
Other	1.1	1.1	1.9	**4.0**

4.4 Analysis by Query Complexity

In regard to the complexity of the queries, it was observed that semantic errors are more common than syntax errors in "easier" queries (L1, L2, and L3). This trend is consistent across the three environments where the amount of semantic/syntax errors for "easy" queries was 72/28, 79/21, and 77/23 for the G, C, and N environments respectively.

When the complexity of the queries increases, the number of syntax errors also increases. However the syntax errors percentage is always less than 50% with the exception of two cases, this happened in queries L6 of the G environment (39/61%) and L7 of the N environment (50/50%).

Table 8 presents the ratio for all the nine levels of complexity (Easy: L1-3, Medium: L4-6, and Hard: L7-9).

Table 8: Error distribution by the difficulty of the query (semantic/syntax in percentages)

	G	C	N
L1	75/25	92/8	87/13
L2	84/16	80/20	65/35
L3	65/35	72/28	77/23
L4	69/31	91/9	73/27
L5	56/44	79/21	74/26
L6	39/61	59/41	61/39
L7	62/38	56/44	50/50
L8	69/31	75/25	71/29
L9	63/37	80/20	56/44

5 DISCUSSION

5.1 SQL Errors and Gamification

The goal of this research was to answer *"how does gamification influence the errors students make while learning SQL?"*, after analyzing more than one thousand answers submitted by more than one hundred students, the following was observed:

Competition reduced the likeliness of syntax errors and improved keyword usage. C students had the least syntax errors while these were more common in the G environment. This may indicate that the competitive factor positively influenced the attention to detail of C students. Also, keyword-related errors were the least in the C environment by a considerable margin.

Gamification improved the understanding of the queries. In the G environment, students delivered higher quality answers than the rest, showing a better comprehension of what they were asked for in the queries.

Most semantic errors "hid" in the SELECT clause, while syntax errors are in the WHERE clause. There was observed a strong tendency to make mistakes related to the columns being projected even if the expected schema was shown as a hint in the application. Interestingly, issues related to aggregate functions were very few compared with the observed by [19].

Regarding the FROM clause, the overall impression that joining is a confusing concept for students [15], was supported by the results of this study; missing, wrong, and extra joins was the second most common semantic error.

GROUP BY and HAVING queries were not that difficult. Grouping has been reported to be a complicated issue for students [15], however, in this study GROUP BY errors were very few (out of the 478 answers that required the use of GROUP BY, less than 5% had errors related to that keyword). The use of HAVING and ORDER was not an issue for the participants in this study

The harder the query the easier it is to make mistakes. As expected, the more complex the query the more likely mistakes are made. In general, the fact that semantic errors were more common than syntax errors, may indicate a lack of understanding of the query, leading students to mismanage the complexity of the queries by writing overly elaborated queries containing unnecessary elements, overusing nesting, and incrementally building queries with persistent errors [16].

Is the effort of gamifying worth it? The N environment usually performed "in-the-middle" of the three environments, raising the question if it is worth the additional effort of gamifying a learning activity. It can be concluded that, in this particular study, gamification and competition had a positive influence in specific aspects, however, as a whole, the differences between gamified, competitive, and non-gamified environments were minimal.

Instructors may want to choose a particular environment to mitigate specific types of errors rather than tackling all the learning-SQL-related issues with a single strategy.

5.2 Threats to Validity

There are several potential threats to the validity of this study and these will be discussed in the following paragraphs.

External validity. Since the empirical study was carried out in one university, results can't be generalized, however, the results do represent a sample of the population under study.

In addition, all of the participants belonged to the same course, they were taught by the same teaching team and used the same resources, an improvement from the original study [7]. Regarding the experimental material, the queries have been validated extensively by its creators and comprise the basic SQL elements taught in introductory Database courses. As suggested in [7], to reduce the limitations of the original study and to maximize the similarity of conditions across the participants of the study, the design followed a one group — one lecturer setting.

Internal validity. Students that participated in the study did it voluntarily and their participation was independent of the course. Those who participated in the competitive environment were given the option to use nicknames so they were not easily identified. The results of the study did not have any influence on the marks the students received in the course. Furthermore, the activities were not linked to any assessment item of the course.

Regarding the previous knowledge students may have about SQL, none of the participants had prior extensive knowledge of SQL, which is supported by the pre-test results where the correct answers were, as expected, very low.

The student dropouts (mortality) were also considered but these rates were very low and similar across environments. The completion rates were over 93% in all environments (G: 95.7%, C:97.6%, and N:93.3%).

Construct validity. The correctness of the marking script used for determining if an answer was correct or not was already addressed in the original study. However, a sample of marked answers was double-checked by the authors of this paper and shared with the lecturer of the course who also validated them.

Another threat was students participating in the wrong environment. To mitigate this threat, the three instances were accessed through three different URLs and explicit checks in the AT-AT backend were implemented so that only certain users could request points and leaderboards.

Lastly, a strength of the study was that the authors did not interact with the participants disallowing potential bias or experimenter expectancies.

6 CONCLUSIONS AND FUTURE WORK

In this paper, an empirical study to observe the effects of gamification on the errors students made when learning SQL was presented. After analyzing a pool of 1,009 answers submitted by 133 students enrolled in a Database course and classifying the errors found in their answers, it was observed that students participating in a gamified environment tended to make fewer semantic errors than in competitive or non-gamified environments.

On the other hand, the competitive environment yielded better results reducing the prevalence of syntax errors. Furthermore, it was observed that SELECT and WHERE clauses were the most common locations where errors were found. Overall, it was concluded that gamification (and competitiveness) influences positively specific aspects of students' learning. However, as a whole, the differences observed between non-gamified and gamified environments were minimal.

In future work, an in-depth analysis should be done to understand the root causes of the semantic errors since it is not possible to know if these were caused because of a lack of understanding of the query or the "rush" to submit an answer. In addition, replicating the study and refining the error classification would contribute to identifying and developing strategies that improve students' efficiency when learning SQL.

REFERENCES

[1] [n. d.]. Computer sciencewithout a computer. https://www.csunplugged.org/en/ Online; accessed 22 January 2023.

[2] [n. d.]. EDU – Code Avengers. https://www.codeavengers.com/edu. Online; accessed 22 January 2023.

[3] [n. d.]. The Learn to Read Process – How Reading Eggs works. https://readingeggs.co.nz/about/why-it-works/. Online; accessed 22 January 2023.

[4] [n. d.]. The schemaverse. https://schemaverse.com/ Online; accessed 22 January 2023.

[5] Alireza Ahadi, Julia Prior, Vahid Behbood, and Raymond Lister. 2016. Students' semantic mistakes in writing seven different types of SQL queries. In *Proceedings of the 2016 ACM Conference on Innovation and Technology in Computer Science Education*. 272–277.

[6] Manal M. Alhammad and Ana M. Moreno. 2018. Gamification in software engineering education: A systematic mapping. *Journal of Systems and Software* 141 (2018), 131–150. https://doi.org/10.1016/j.jss.2018.03.065

[7] Anonymous authors. 0000. Anonymous title. 0 (Anonymous 0000), 0–0. Issue 0. https://doi.org/0000000

[8] Anonymous authors. 0000. Anonymous title. *Anonymous conference* (0000), 0–0.

[9] Jackson Best. 2019. The Value of Games and Gamification with Mathematics. https://www.mathletics.com/blog/educators/value-of-games-in-mathematics/. (Accessed on 26/03/2022).

[10] Stefan Brass and Christian Goldberg. 2006. Semantic errors in SQL queries: A quite complete list. *Journal of Systems and Software* 79, 5 (2006), 630–644. https://doi.org/10.1016/j.jss.2005.06.028 Quality Software.

[11] Gretchen I. Casterella and Leo Vijayasarathy. 2019. Query structure and data model mapping errors in information retrieval tasks. *Journal of Information Systems Education* 30, 3 (2019), 178 – 190. https://www.scopus.com/inward/record.uri?eid=2-s2.0-85081174751&partnerID=40&md5=43b2bf018d14881e91b93c1a12522187 Cited by: 5.

[12] Christo Dichev and Darina Dicheva. 2017. Gamifying education: what is known, what is believed and what remains uncertain: a critical review. *International Journal of Educational Technology in Higher Education* 14, 1 (2017). https://doi.org/10.1186/s41239-017-0042-5 Cited by: 373; All Open Access, Gold Open Access.

[13] Juho Hamari, Jonna Koivisto, and Harri Sarsa. 2014. Does Gamification Work? – A Literature Review of Empirical Studies on Gamification. In *2014 47th Hawaii International Conference on System Sciences*. 3025–3034. https://doi.org/10.1109/HICSS.2014.377

[14] Theresia Devi Indriasari, Andrew Luxton-Reilly, and Paul Denny. 2021. Improving Student Peer Code Review Using Gamification. In *Australasian Computing Education Conference* (Virtual, SA, Australia) *(ACE '21)*. Association for Computing Machinery, New York, NY, USA, 80–87. https://doi.org/10.1145/3441636.3442308

[15] Daphne Miedema, Efthimia Aivaloglou, and George Fletcher. 2022. Identifying SQL Misconceptions of Novices: Findings from a Think-Aloud Study. *ACM Inroads* 13, 1 (feb 2022), 52–65. https://doi.org/10.1145/3514214

[16] Daphne Miedema, George Fletcher, and Efthimia Aivaloglou. 2022. So Many Brackets! An Analysis of How SQL Learners (Mis)Manage Complexity during Query Formulation. In *Proceedings of the 30th IEEE/ACM International Conference on Program Comprehension* (Virtual Event) *(ICPC '22)*. Association for Computing Machinery, New York, NY, USA, 122–132. https://doi.org/10.1145/3524610.3529158

[17] Seth Poulsen, Liia Butler, Abdussalam Alawini, and Geoffrey L. Herman. 2020. Insights from Student Solutions to SQL Homework Problems. In *Proceedings of the 2020 ACM Conference on Innovation and Technology in Computer Science Education* (Trondheim, Norway) *(ITiCSE '20)*. Association for Computing Machinery, New York, NY, USA, 404–410. https://doi.org/10.1145/3341525.3387391

[18] Toni Taipalus. 2020. Explaining Causes Behind SQL Query Formulation Errors. In *2020 IEEE Frontiers in Education Conference (FIE)*. 1–9. https://doi.org/10.1109/FIE44824.2020.9274114

[19] Toni Taipalus, Mikko Siponen, and Tero Vartiainen. 2018. Errors and Complications in SQL Query Formulation. *ACM Trans. Comput. Educ.* 18, 3, Article 15 (aug 2018), 29 pages. https://doi.org/10.1145/3231712

[20] Dr. Zachary Fitz Walter. 2022. What is Gamification? https://www.gamify.com/what-is-gamification. (Accessed on 26/03/2022).

[21] Marta Elena Zorrilla Pantaleón, Diego García-Saiz, and Alfonso de la Vega. 2021. Fostering study time outside class using gamification strategies: An experimental study at tertiary-level database courses. *Computer Applications in Engineering Education* 29, 5 (2021), 1340–1357. https://doi.org/10.1002/cae.22389 arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1002/cae.22389