Exploring Feedback and Gamification in a Data Modeling Learning Tool

Olav Dæhli[1], Bjørn Kristoffersen[2], Per Lauvås jr[3] and Tomas Sandnes[3]
[1]University of South-Eastern Norway, Porsgrunn, Norway
[2]University of South-Eastern Norway, Bø, Norway
[3]Kristiania University College, Oslo, Norway
olav.dehli@usn.no
bjorn.kristoffersen@usn.no
per.lauvas@kristiania.no
tomas.sandnes@kristiania.no

Abstract: Data modeling is an essential part of IT studies. Learning how to design and structure a database is important when storing data in a relational database and is common practice in the IT industry. Most students need much practice and tutoring to master the skill of data modeling and database design. When a student is in a learning process, feedback is important. As class sizes grow and teaching is no longer campus based only, providing feedback to each individual student may be difficult. Our study proposes a tool to use when introducing database modeling to students. We have developed a web-based tool named LearnER to teach basic data modeling skills, in a collaborative project between the University of South-Eastern Norway (USN) and Kristiania University College (KUC). The tool has been used in six different courses over a period of four academic years. In LearnER, the student solves modeling assignments with different levels of difficulty. When they are done, or they need help, they receive automated feedback including visual cues. To increase the motivation for solving many assignments, LearnER also includes gamifying elements. Each assignment has a maximum score. When students ask for help, points are deducted from the score. When students manage to solve many assignments with little help, they may end up at a leaderboard. This paper tries to summarize how the students use and experience LearnER. We look to see if the students find the exercises interesting, useful and of reasonable difficulty. Further, we investigate if the automated feedback is valuable, and if the gamifying elements contribute to their learning. As we have made additions and refinements to LearnER over several years, we also compare student responses on surveys and interviews during these years. In addition, we analyze usage data extracted from the application to learn more about student activity. The results are promising. We find that student activity increases in newer versions of LearnER. Most students report that the received feedback helps them to correct mistakes when solving modeling assignments. The gamifying elements are also well received. Based on LearnER usage data, we find and describe typical errors the students do and what types of assignments they prefer to solve.

Keywords: entity relationship diagrams, ERD tool, automatic formative assessment, automatic formative feedback, gamification in education

1. Introduction

More than 50 years after Codd introduced the relational model of data (Codd, 1970), relational databases are still in use today. Although we have multiple options when selecting a storage medium for our data, the relational database is still often the choice in IT projects today. In IT education, knowledge and skills regarding databases are essential. In the 2008 ACM/IEEE Curriculum Guidelines for Undergraduate Degree Programs in Information Technology (Lunt et al., 2008, p. 19), "Databases" was firmly placed as one of the five pillars within the IT discipline. In the currently newest version (Task Group on Information Technology Curricula, 2017, p. 50), "data modeling" and "database query languages" are examples of topics within the *essential* IT domain: Information management.

The order in which to teach data modeling and database query language (SQL) may vary. Our contribution context is several courses in multiple institutions, where both modeling and SQL are taught in the same course, with SQL first, then modeling. In study programs with multiple database courses, this would translate into having a course with emphasis on SQL followed by a separate course on modeling (e.g. Migler and Dekhtyar, 2020).

When designing a database, we have several modeling notations to choose from. Entity Relationship Modeling (ER) (Chen, 1976), Unified Modeling Language (UML) and Object-Role Modeling (ORM) (Halpin and Bloesch, 1999) are well-known alternatives. We also have multiple professional tools to choose from when developing a model, and later transforming the model into database schemas. But these professional tools require the user to already *have* the modeling skills required to do so.

Reference this paper: Dæhli, O., et al., 2021. Exploring Feedback and Gamification in a Data Modeling Learning Tool. *The Electronic Journal of e-Learning,* 19(6), pp. 559-574, available online at www.ejel.org

An educational tool could *help* the students *acquire* modeling skills. Intelligent Tutoring Systems (ITS) are "computer programs that use artificial intelligent techniques to enhance and personalise automation in teaching" (Alkhatlan and Kalita, 2018, p. 1). In database modeling, multiple tools have been developed over the years for that purpose. EER-Tutor (Suraweera and Mitrovic, 2004) is a mature ITS dating back to the early 2000s. Originally named KERMIT (Suraweera and Mitrovic, 2002), EER-Tutor originated from the University of Canterbury, New Zealand, and is still in use. In the last 20 years, several papers describe the tool from multiple angles (e.g., Zakharov et al., 2007; Duan et al., 2010; Mathews et al., 2012; Mitrovic and Holland, 2020). With EER-Tutor, students can select among several available modeling assignments, try to complete the task, and receive formative feedback.

Automatic grading of free-form ER diagrams is a hard problem (Jayal and Shepperd, 2009). In recent years, multiple attempts have been made to automate (or semi-automate) assessment and grading of both ER models (Batmaz and Hinde, 2006; Thomas et al., 2006; Thomas, 2013; Simanjuntak, 2015; Lino and Rocha, 2018) and UML class diagrams (Hoggarth and Lockyer, 1998; Ali et al., 2007; Soler et al., 2010; Hasker, 2011; Stikkolorum et al., 2019; Bian et al., 2019). Semi-automatic grading has the potential both to decrease the workload of teachers and achieve a fairer marking (Batmaz et al., 2010). The efficacy of the automated grading effort has also been evaluated (Bian et al., 2020). Using the grading algorithm described in Bian et al., 2019, they found that the grading strategy needs to be adapted to the level of the student and the grading style of the instructor. Further, they emphasize that the grading must consider multiple possible solutions.

The process of learning the modeling notation and the related formal rules for drawing diagrams, is normally not that demanding. Related to Bloom's taxonomy (Bloom, 1956), a way of classifying learning objectives, acquiring basic knowledge and comprehension are the least demanding part of the learning process. Nor are there so many symbols and rules to learn in this case.

After having achieved an understanding of the basic terminology, students have to learn how to use a requirement specification to analyze the need for storing data, and from this construct an appropriate data structure. This is a more complicated process, where students need to obtain an understanding at a higher level (ref. Bloom). This requires volume training, solving exercises of increasing size and complexity.

Errors made by students trying to learn data modeling have been studied and classified, and also compared with Bloom's taxonomy (Batra and Antony, 1994; Bogdanova and Snoeck, 2019; Rosenthal and Strecker, 2019).

Most students need a lot of help and feedback on their work along the way. Formative feedback should be supportive, timely, and specific (Shute, 2008; Hattie and Timperley, 2007). With a rising number of students in higher education, providing quality feedback is time demanding for teachers and it limits the students' ability to do the work when and where it suits them best.

LearnER is designed to provide students with automatically generated feedback when they are in the process of solving various modeling assignments. It is intended to be used as a tutoring tool, and not for automated grading. Professional data modeling tools should also be included in a database course, but mainly *after* the students have attained a certain level of understanding.

Whether working alone or in teams, several students need extra motivation for solving exercises. To meet this end, elements of gamification are included in LearnER, such as earning points through solving tasks and having leaderboards (high score lists) where students can compare themselves to others. Gamification is the use of game mechanics as motivation in non-game contexts, for example to increase learning in educational applications (Kapp, 2012).

Formative feedback and gamification seem to be a promising combination in educational software (Fuchs and Wolff, 2016; Menezes and Bortoli, 2016; Keuning et.al., 2018; Zainuddin et al., 2020). The latter systematic literature review concludes that "Gamification is an uprising trend that applies gaming mechanics as a driver to motivate, engage and enhance the user experience." (Zainuddin et al., 2020, p. 15).

A recent attempt at gamifying database design learning is MonstER Park (Schildgen, 2020). MonstER Park is a free online game where the learner advances through a game creating ER diagrams along the way in a step-by-step fashion.

In this paper, we have investigated the following research questions:

1. When and how do students use LearnER?
2. To which extent do the formative comments give adequate feedback to students?
3. Do the gamification elements contribute to learning, and if so, in what way?

We have investigated the same questions based on earlier versions of LearnER (Dæhli et al., 2018; Dæhli et al., 2020), this paper is an extended version of the latter paper. We have further developed LearnER since 2020. We present additional data on student experiences from using the tool. We expand our data set with new surveys, and we add a new data set; usage data stored in the LearnER database.

2. LearnER – combining feedback and gamification

LearnER is part of the free online resources of a Norwegian textbook written by one of the authors of this paper. The tool contains a set of predefined exercises of various difficulty, each having a model solution. Teachers may also add extra exercises. Students construct ER models by using words extracted from the model solution and may at any time check their model and receive elaborate formative feedback as well as visual cues.

A stylized version of the user interface is presented in Figure 1. The student is presented with a scenario text and can build a data model by dragging the available labels (words) onto the drawing area.

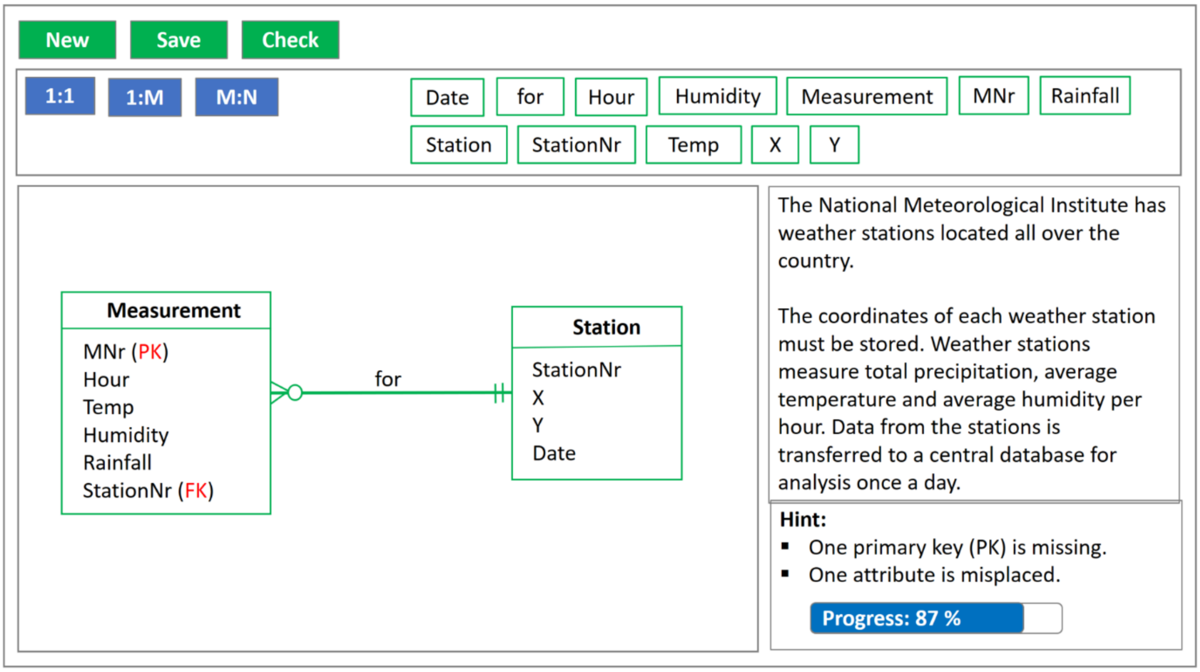

Figure 1: Stylized LearnER user interface

LearnER incorporates gamification as a motivating factor to solve exercises. Each exercise has a maximum number of points that can be achieved. When a student requests help from LearnER to solve an exercise, the maximum number of points the student can achieve is lowered. A leaderboard is kept for each exercise and for the total number of points achieved. A separate panel present feedback and progress bars, on the student's request.

LearnER supports three notations, namely UML and two variants of Crow's Foot. Furthermore, LearnER supports both high-level conceptual models and more implementation-oriented models. In addition, an SQL-script for creating an entire database, based on the student's solution, can automatically be created when students have reached the goal of making a complete model.

A new major version of LearnER has been introduced at the start of each academic year, but minor versions in-between has made it possible to roll out small changes and bug fixes for the spring semester. The most important changes during the project period have been improvements to the UI, better logging, several improvements to the feedback and scoring systems, making the tool more flexible, and adding a much larger exercise bank.

 ISSN 1479-4403

To make some of the exercises more demanding for eager students, some exercises were, in the newest version, supplemented with a few extra "inadequate" words – in the sense of being words that do not fit into the solutions. This makes it harder for students to pick the correct entity and attribute names to be used in the model.

3. Method

Different versions of LearnER have been used as a pedagogical tool in six courses, including one distance learning course, over a period of four academic years from fall 2017 to spring 2021. Qualitative and quantitative research on how students use LearnER and their experienced learning effects have been conducted in several studies. Research findings and feedback from teachers and students have been used to further develop the tool.

3.1 Surveys

Several questionnaires have been designed and distributed to all students in these courses. A core set of 9 multiple-choice questions, mostly of the Likert type, has been included in every survey from fall 2017 to spring 2020, and 5 of these were also included in a final survey conducted spring 2021. A few extra multiple-choice and free-text questions have been introduced in some of the surveys. In fall 2020 we conducted a different type of survey with mainly free-text questions.

The questionnaires have been distributed to more than 1800 students over the four-year period. Figure 2 shows an estimate of the number of active students at each campus (placing the distance learning students in their own virtual campus), which we here define as the number of students attending the final exam – according to the national Common Student System ("Felles Studentsystem", FS).

A total of 1886 active students have received one of the surveys in the period: 406 in 2017/2018, 346 in 2018/2019, 438 in 2019/2020 and 696 in 2020/2021. The number of students receiving the questionnaire were somewhat higher than this. 357 students have responded to the questionnaires, giving an overall response rate of 19% among active students.

The data for campus Oslo covers two database courses in the 2019/2020 column, and the data for the Vestfold campus covers three courses all taught in spring 2021. The response rate for campus Oslo in fall 2019 and fall 2020 were significantly lower than the rest, only 5% and 7%.

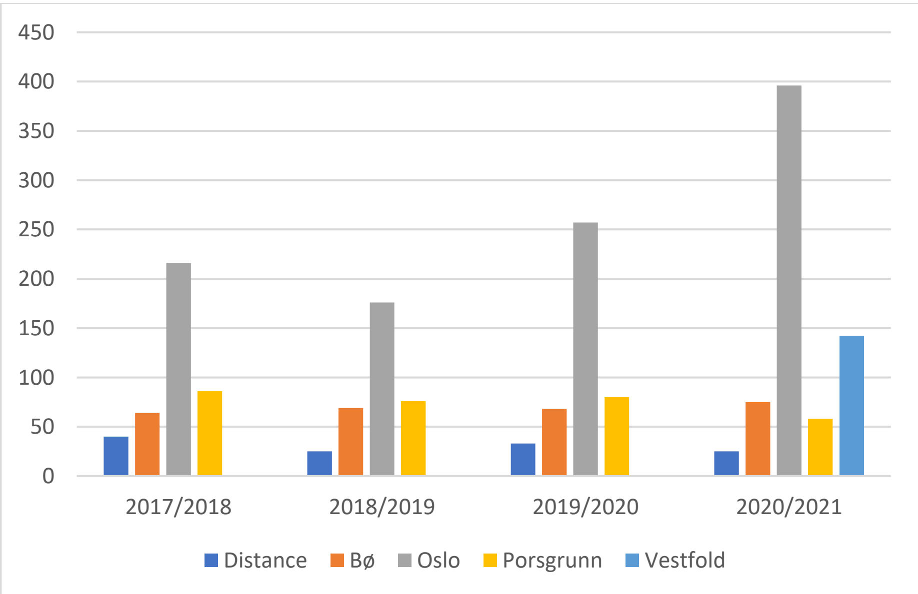

Figure 2: Number of active students at different campuses per academic year

The authors have taught all six courses combined. Some of the courses involved in this study are taught in the fall and some in the spring. Also, the data modeling part of these courses is not always taught in the same weeks within the semester, and the time interval when surveys were conducted varied slightly from campus to campus and from semester to semester. This may have affected the results somewhat, in particular in the first academic year, when there was bug fixing in between different surveys.

Still, all courses within a single academic year used similar versions of the tool. When comparing results, we therefore regard all responses for courses held in one academic year as one dataset and compare it to the datasets from the other academic years.

3.2 Interviews and observations

In 2017/2018 we did semi-structured individual interviews with 19 students from three campuses, including some distance learning students, about their opinion on LearnER (Dæhli et al., 2018). 8 more students were interviewed fall 2018, following the same procedure, but now based on a new version of LearnER.

In spring 2019, 3 additional students were interviewed, again based on the same interview guide, but now the students solved a data model exercise in LearnER immediately before the interview. The students were told to think aloud during problem solving and one of the authors was passively observing the process.

We have not interviewed students based on the latest version of LearnER. However, all four paper authors also teach the courses being investigated, and have all been involved in LearnER lab exercises, giving rich possibilities for informal observations and oral feedback from students during their work with LearnER. Also, in 2020/2021 surveys, more free-text questions have been included to obtain more qualitative data.

Permission to retrieve non-anonymous data was obtained from Norwegian Centre for Research Data (NSD) and the participating students. Norwegian quotes were translated into English by the authors of the paper. The translated quotes are listed in quotation marks, although they are not language direct quotes.

3.3 LearnER usage data

LearnER is a freely available web tool, but most users are probably students taking our courses. The number of user accounts have increased from a little over 100 in the first year to over 400 per year.

A valid email is needed for user registration but is not permanently stored. Students choose a nickname and a password during registration and can then use the system anonymously. They are warned that nicknames may appear on leaderboards and are also asked not to use passwords they use in other (more critical) applications. All student data models are kept in the LearnER database, together with nicknames and time stamps showing when the student started and finished working on the exercise. Also, each time the student asks for help, the current state of the data model is persisted to the database.

The usage data in the LearnER database is archived at the end of each semester, and then all users and the data models they have built are deleted from the online system, which means that all high score lists are empty at the start of a new semester.

4. Results

All courses are introductory with no prerequisites. Most of the 2017/2018 students reported they had no knowledge about data modeling concepts before taking the course (Dæhli et al., 2018).

4.1 Increased student activity

Students reported that they did more exercises in newer versions of LearnER, see Figure 3. We observe that only 11% of the students said they did more than 10 exercises in 2017/2018. In the following three years, the corresponding numbers are 24%, 34% and 28%.

The apparent increase in student activity based on these self-reporting numbers can be further supported by inspecting LearnER usage data, where the average number of completed exercises goes up from 3.0 in 2017/2018 to 6.5 in 2018/2019, 7.9 in 2019/2020 and 11.9 in fall 2020. The number of exercises for 2019/2020 deviates somewhat from what we reported in (Dæhli et al., 2020), due to improved data cleansing.

Survey data and LearnER usage data is not directly “linked”. Some LearnER users may not be students at our courses. Also, LearnER usage data for spring 2021 is not included, whereas the 2020/2021 data in Figure 3 is based *only* on spring 2021. What we can say, however, is that both survey data and LearnER usage data indicate that students solve more exercises in newer versions of the tool.

.

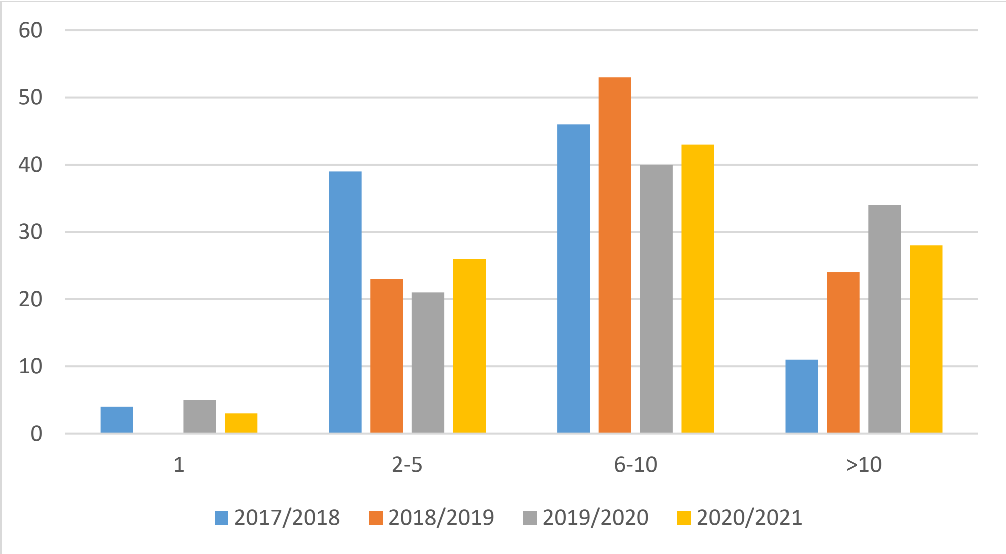

Figure 3: How many exercises did you attempt in LearnER?

We can think of several possible explanations for the increased student activity:

- More exercises have been added to the system before each academic year. In 2017/2018, LearnER had less than 20 exercises, now it contains approximately 50 exercises. This should make LearnER more interesting.
- In early versions of the tool, the teacher labeled exercises “easy”, “medium” or “difficult”. From 2018/2019 onwards, exercises were reassigned a difficulty level from 1 to 10, again discretionary by the teacher. This fine-grained categorization means that a student needs to solve more exercises to reach the “top”.
- The increased activity might also be caused by general improvements to the system. A better tool is more fun and rewarding to work with. As reported in (Dæhli et al., 2018), several 2017/2018 students mentioned the tool was “prototypical” and “buggy”. There were fewer comments about bugs in the 2018/2019 survey, and even fewer again in later surveys.
- Changes in teaching arrangements could also affect the activity, of course, but we have been using LearnER in much the same way in all six courses over the period.

4.2 Collaboration

During the first three years, a relatively large number of students were using LearnER together with others to a large degree – 47%, 32% and 42%, respectively. See figure 4. In spring 2021, this dropped to only 18%. We speculate that this is a coronavirus effect.

Spring 2020 was also affected by the coronavirus, but not until mid-term. The students had already gotten to know each other and had formed relations and study-groups by that time, which could be the reason we did not find a decrease in collaboration for the 2019/2020 academic year.

We observe differences in reported collaboration between the campuses/courses also before the pandemic. The course in Porsgrunn has a high degree of student collaboration, while the distance learning students prefer to work alone. Some of these differences can be explained by how learning activities are organized for various courses, e.g. if LearnER exercises are given as group assignments.

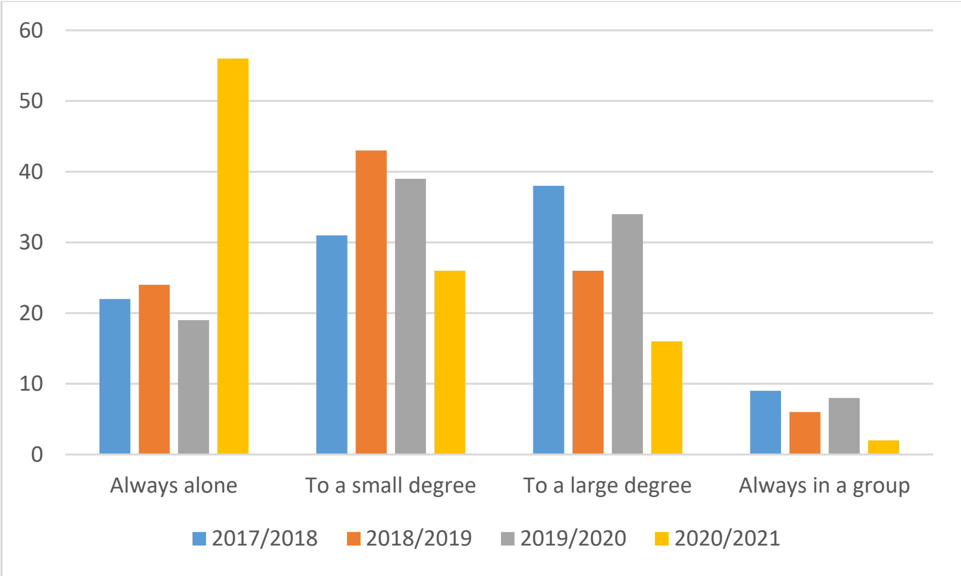

Figure 4: To what extent did you use LearnER together with others?

4.3 Formative feedback and visual cues

The feedback system was redesigned in the summer of 2018. The new system gave more elaborative feedback, first at a general level and then more detailed on the students' request. Students could also get visual cues that pinpointed errors in their models.

In 2017/2018, students said they wanted more detailed feedback (Dæhli et al., 2018), but the surveys indicate that even more students read the feedback carefully in the early versions. 55% of the students strongly agreed with the statement "I read the feedback from LearnER carefully", compared to 30% in 2018/2019 and 37% in 2019/2020. The percentage of students who either agreed or strongly agreed was, however, about the same over the three years. We did not include this question in the 2020/2021 survey.

The amount of feedback text is substantially *higher* in newer version, possibly explaining why fewer students *strongly* agreed with the statement. Clearly, improvements can be made in feedback design, several students find the feedback texts too verbose and perhaps too general or theoretical. A student commented: "LearnER is a good program, but the feedback should be more focused on each exercise. It should provide a little more information about the exercise itself, and not generalize it. Otherwise very happy :)".

A similar argument can be made regarding visual cues that are added to the newer versions, "showing exactly where it is wrong" as one of the 2019 students puts it. Students now have alternative means for correcting their data models and may not *need* to read all the feedback.

It seems more students are now able to *correct* their data models based on feedback from the system. In 2017/2018, 54% of the respondents agreed or strongly agreed that they were able to correct their data model based on feedback from the system. This percentage was 80% in 2018/2019 and 71% in 2019/2020. We did not include this question in the 2020/2021 survey.

We found the same pattern when asking students if feedback from LearnER is helpful in learning data modeling, see Figure 5. 52% agreed or strongly agreed with this in 2017/2018, rising to 88% in 2018/2019, but then a little down to 71% and 75% the last two years.

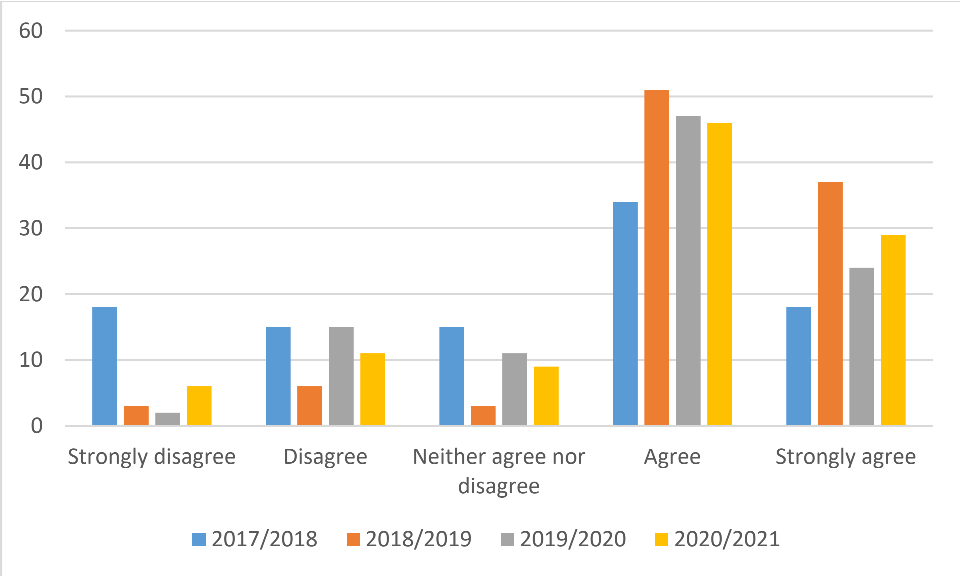

Figure 5: Feedback from LearnER helped me to learn data modeling

The 2019/2020 version of LearnER was made more flexible by allowing several related and equally "good" answers, by introducing what might be called "don't care" constraints. Technically speaking, this means that both "0 or many" and "1 or many" could be defined as correct for a given relationship. In early versions of LearnER, the person adding the model solution, had to somewhat arbitrary choose between one or the other. Students then had to guess the correct solution, but sometimes they got stuck – maybe (correctly) thinking that "This can't be wrong!".

Both visual cues and flexible solutions made it easier for the students to find the correct solution in the 2019/2020 version, and it also possibly made them less dependent on explanations from textual feedback.

4.4 Gamification and motivation

Some elements of gamification are included to stimulate student activity, such as difficulty levels, score points and leaderboards. Results from both interviews and questionnaires indicate that most students find this motivating and solve more exercises because of it, even though a few find it irrelevant. "To me, it's motivating. Absolutely. For me who likes games, I look at it as a challenge. It's an exciting part of the challenge."

Students who like the gamification often say they have a competitive instinct. It triggers them to try to get on the high score list, and to compete against classmates or team members. "Extremely good concept. For people with competitive instinct, it is always fun to get points and be measured against others."

After redesigning the algorithm for computing scores before the second year (2018/2019), more students found the game mechanics motivating, see Figure 6. 59% of the students in 2017/2018 agreed or strongly agreed that earning points are motivating, rising to 89% in 2018/2019, but (again) down, to 79% in 2019/2020 and 78% in 2020/2021. Very few students disagreed in the newer versions.

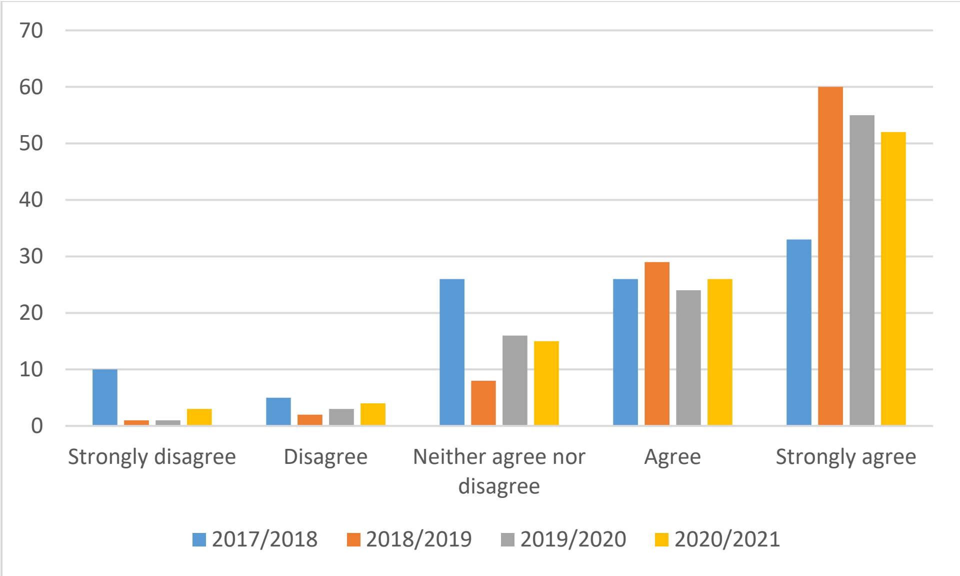

Figure 6: I find it motivating to get points for each exercise

The way scores are calculated is quite simple, and should perhaps be made more sophisticated, as several students point out. A student said it like this: "So one can get a high score by taking many, many exercises, and one can get a high score by just taking a few and being good at it. So maybe the number of times you've tried should count." Even if students "see through it", some are still able to fool themselves into being motivated: "Looked through it a bit, but other than that it was fun, it was motivating."

4.5 Exercises organized by difficulty levels

Figure 7 shows the number of exercises solved at each difficulty level, for the last three semesters, based on LearnER usage data. We observe that many students attempt to do exercises up to level 5. The number of exercises solved at difficulty level 1 to 5 make up for 88 %, 86 % and 84 % of the total number of exercises solved for the three semesters.

It seems that almost every student starts with a demo exercise at level 1, then works their way through the exercises of increasing difficulty levels. Looking at the numbers for attempts per exercise (disregarding a couple of exercises we encourage students to start with), these are evenly distributed within each difficulty level, suggesting that the students pick their exercises within a difficulty level at random, or possibly does all exercises within a level before they move on to the next difficulty level.

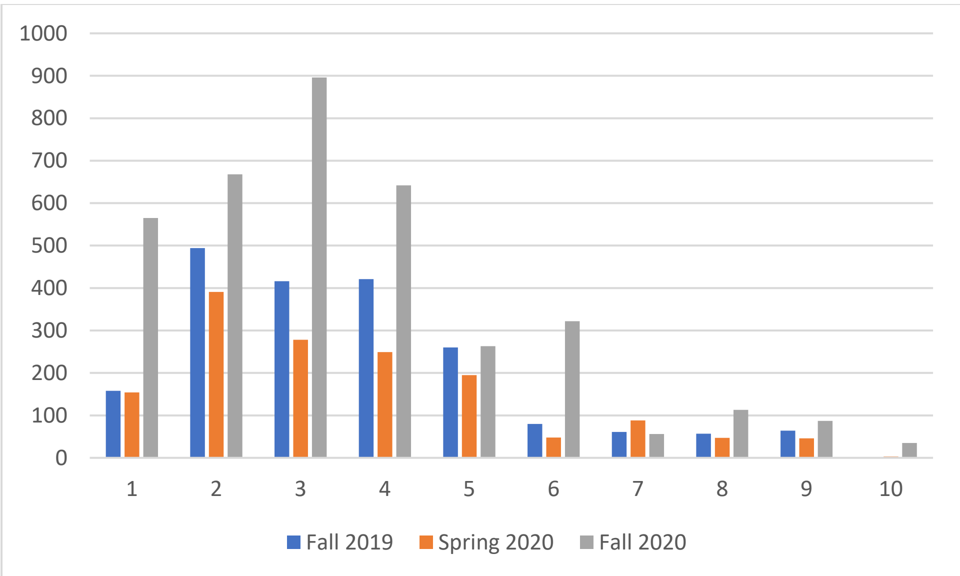

Figure 7: Exercises solved for each difficulty level

Figure 8 shows the average amount of time spent to solve exercises at each difficulty level. Difficulty levels are set at the teacher's discretion, by an overall assessment based on both size and complexity of the proposed solution as well as the exercise text itself. According to Figure 8, it seems fair to say that the difficulty levels have been set reasonably "correct", possibly with an exception for some of the level 6–7 exercises.

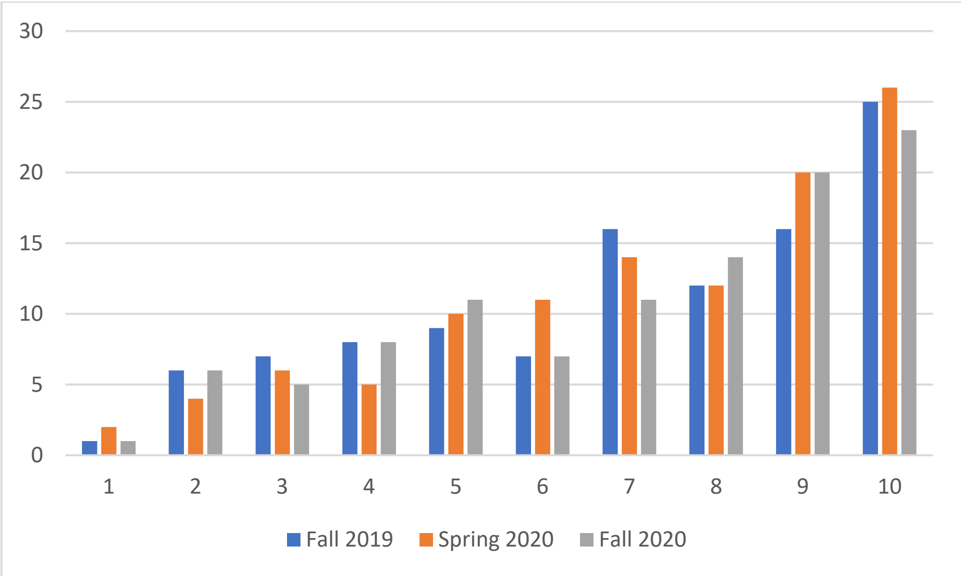

Figure 8: Average amount of time spent to solve exercises at each difficulty level

4.6 The useful, but possibly dangerous Check button

When students solve exercises in LearnER, the idea is that it should happen in a similar way as when working on assignments where they can get guidance from a supervisor. In such a situation, students will be able to request assistance if they are unable to move forward on their own. A clever supervisor will not give answers to students straight away, but will rather guide them in the right direction, with the aim of enabling them to complete the tasks themselves. This is also how we think about LearnER and how it should be used.

To achieve something in this direction, a Check button is made available. It is not possible to view the solution model in LearnER, you must solve it yourself, but the Check button gives the student an opportunity to ask for help. The response will not be a definitive solution to the problem, but rather some hints about parts of the solution that needs to be corrected.

The students use this Check button a bit more than we expected, especially for difficult and large exercises. See Figure 9. One explanation for this is that even though the Check button in LearnER shows explanations for all errors in the model, students often fix only one or two errors, and then click the Check button again.

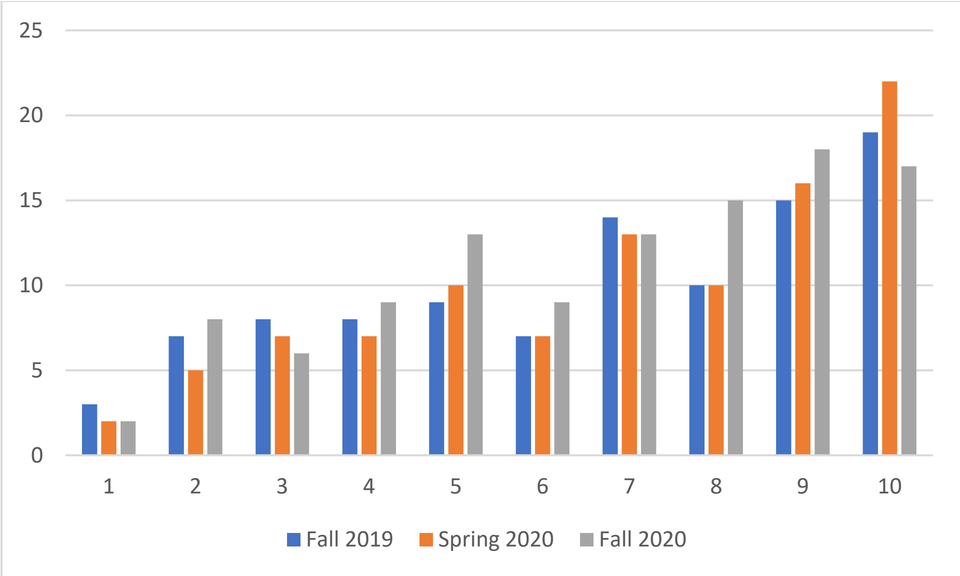

Figure 9: Average number of checks for each difficulty level

In the follow-up interviews, students said that some of the feedback texts were cryptic or too general, especially those concerning relationships: "Could have been even better feedback explaining why the answer is what it is. Sometimes it's just trial and error until you get it right, but you don't always understand why it's right." Other students found it difficult to understand the scenario text itself. That is not necessarily something that should be fixed. After all, learning to translate an informal text into a data model is what this is all about.

Of the three 2019 students that accepted to be observed and interviewed, there were one campus student and two distance learning students. All three used between 12 and 14 minutes to complete a given (and new) exercise at level 7. They used similar strategies, starting with entities and attributes before adding relationships. They all used both textual feedback and visual cues combined with some trial and error to get all the relationships correct. But the number of clicks on the Check button was significantly lower than average. With only three students, this could be a coincidence, but there could also be other explanations: the three students were all high achievers, the exercise was new and maybe assigned a "wrong" difficulty level, or the fact that the teacher was observing them could make the students read the feedback more thoroughly.

Many students are aware that they lose points by clicking on the Check button, and some try to avoid it: "I used it a few times (i.e., the Check button). I tried not to use it, because then I lost too many points." Others are determined on solving problems on their own. One of the 2019 students, being asked about the new two-level feedback system and more possibilities for getting help, answered: "Yes, that's a good feature. But preferably you don't want to press help at all, (pause), at least I want to figure it out myself. It's a bit like, ah, if you have to press for help, it feels a bit bad."

4.7 Typical data modeling errors

The system detects the following errors and omissions:
Entity: Incorrect entity name, i.e., an attribute or relationship name is used as the name of an entity.

Attribute: Either an attribute is placed in the wrong entity, or a word selected as the name of an attribute is not an attribute.

Keys: An attribute is incorrectly marked as primary key or foreign key.

Relationship: Errors in cardinality or requirements for participation in relationships, e.g., that it is set to "0 or many" where it should have been "exactly 1" or maybe "1 or many". Relationships created between wrong entities or wrong relationship name. Identifying relationships defined as non-identifying – or vice versa.

Missing elements: Entity, attribute or relationship is missing. Primary key or foreign key is missing.

Inspection of student models for simple exercises shows that most of the entities and attributes are in place when students click Check for the first time. For difficulty levels 5–10, most students choose a more step-by-step procedure with checking along the way.

Figure 10 shows the distribution of these different types of errors for different difficulty levels. Relationship errors is by far the most common type of error, and many of these are cardinality errors. Level 1 exercises include only one entity and are therefore without relationship errors. Misplaced attributes are overall the second-most common error.

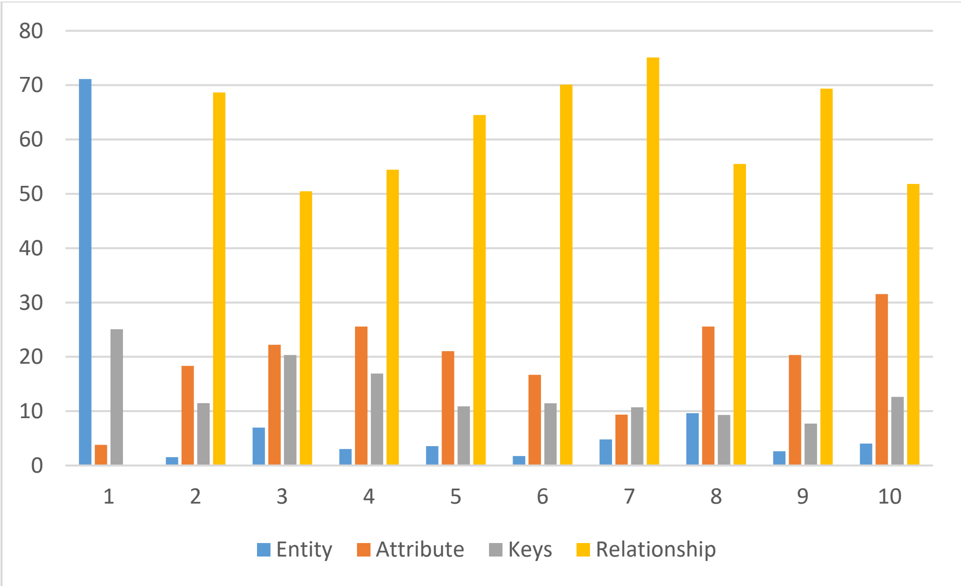

Figure 10: Types of errors for each difficulty level (fall 2020)

4.8 Extra, inadequate words and working without help

To make an exercise more challenging, extra, inadequate words may be added to an exercise's dictionary. These extra words are listed among the relevant words but are not part of the solution. Exercises facilitating extra words state that they use this feature at the start of their descriptive text.

The option to include extra words in exercises was added as a feature from autumn 2020 onwards. The student surveys for autumn 2020 and spring 2021 both have questions regarding the usage of extra, inadequate words.

Many of the extra words were values that are natural to store in the database, but which were not suitable as part of the structure (table or column names), e.g., specific genres “thriller” and “western” in a movie database, or “red”, “white” and “blue” in a clothing store database. Other words were imprecise or too general words such as “overview”, “percentage” or “database”.

Inspection of student models shows that the extra words are rarely chosen. But students seem to appreciate the opportunity for more challenging work. One student writes: “It gives us the opportunity to think more, and actually ponder a bit.”

Several students also observe that extra words is a step towards doing data modeling in a professional tool, with no assistance, as this student puts it: “A good middle ground between finding all the words from the assignment text alone and getting only the words you need.”

LearnER also includes the possibility of solving exercises without assistance. If so, students choose entity and attribute names freely. LearnER does not offer any assistance or feedback in this case, so this is the same as solving exercises in a standard ER modeling tool or by free hand on paper.

Some students have tried this, and naturally also want to get help in this situation: "It would be nice with a middle ground [...] where you could choose your own names for the entities and attributes but get some help along the way and some tips."

One of the 2019 students talked about this in the interview, explaining how he had developed an interesting learning strategy: "It's a really cool tool and I feel like I'm learning from it. But the best strategy is to solve the exercises on paper first. Find the entities based on the scenario text, and then go in and check that, yes, [my data model] matches this and that."

4.9 Experienced learning effects

In a final multiple-choice question, we asked students how they felt LearnER contributed to their learning of data modeling, see Figure 11. In 2017/2018, 31% of the respondents reported that LearnER contributed to a high degree or to a very high degree of their data modeling learning. This increased to 56%, 50% and 57% in the following three academic years, respectively.

Students think LearnER is a useful tool to get started with data modeling: "Easy way to learn modeling without knowing much in advance." One student mentioned the effect of being aware of proper naming (which is important): "It showed how to name tables and columns."

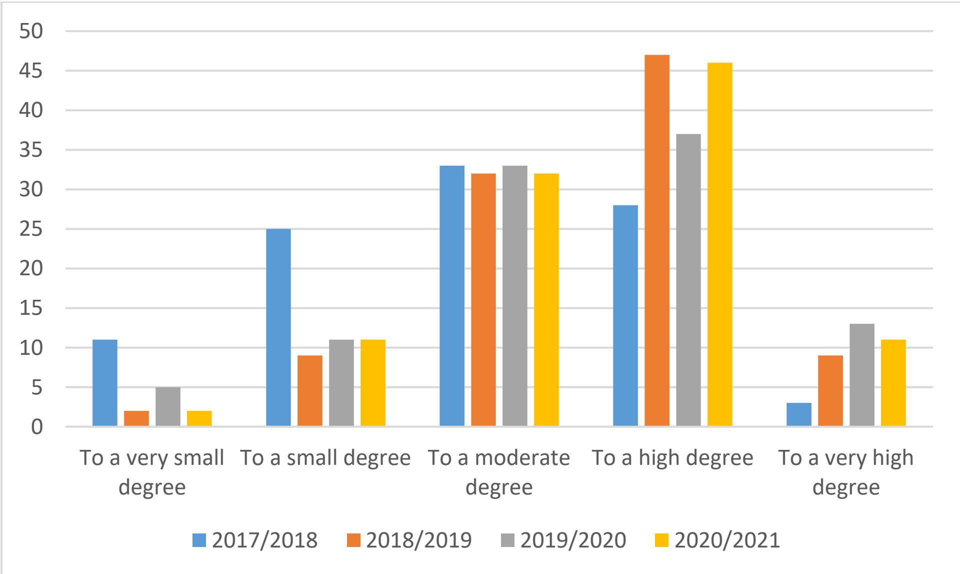

Figure 11: To what degree did LearnER contribute to your learning of data modeling?

5. Discussion

We will now discuss findings related to each research question stated in the introduction.

5.1 When and how do students use LearnER?

As stated in the introduction, formative feedback should be supportive, timely, and specific (Shute, 2008). LearnER meets these requirements to a large extent. It enables students to solve problems when and where it suits them, while still having the opportunity to receive formative feedback aimed specifically at the task they are working on.

The pandemic situation worldwide from march 2020, due to Covid 19, has made it even more important to support the learning process with tools that enable students to work actively on their own, without having a supervisor physically available to support the learning process. LearnER provides precisely this opportunity.

LearnER was designed primarily for individual work but seems to be well suited also for collaboration. In classes where students have been physically divided into teams, feedback from supervisors tells us that students often collaborate when they solve exercises in LearnER. Collaboration takes place through discussing solutions together, helping each other when someone is not moving forward, as well as competing to be the first to complete a task.

We observe that students have been able to solve more tasks by themselves using feedback from the tool, or by collaborating with others. Our experience is that LearnER facilitates active learning, allowing teachers to focus on non-trivial issues, which is beneficial for both teachers and students.

It can be hard to engage students in voluntary use of learning tools. LearnER usage data and survey results show that students solve more exercises in newer versions of the tool. We find this promising since LearnER exercises are not mandatory assignments in our courses. It would be interesting to get more exact data for different courses, like the results reported for EER-Tutor (Mitrovic and Holland, 2020).

It is important that the use of such applications is put into a pedagogical context. In courses where LearnER is used, students are first introduced to the basic concepts of modeling, then they are working with LearnER to construct their own understanding, as well as gaining practice by solving many exercises. Finally, after having achieved sufficient skills, they work out "real world" solutions on their own, by using professional modeling tools.

5.2 To which extent do the formative comments give adequate feedback to students?

Even if students have access to the entity and attribute names in the model solution, they sometimes "get lost". Still, it seems that this happens less often with newer versions of LearnER. Visual cues and flexibility in relationship cardinalities are probably the most important new measures for guiding students towards the solution. Adding several different solutions to each exercise is an alternative approach (Bian et al., 2020).

More elaborate feedback is useful and was requested by many students in the early versions. It seems that some explanations, in particular concerning relationships, are still considered to be too verbose and general. LearnER usage data shows that relationships errors are the most common, and this is also reported in (Rosenthal and Strecker, 2019).

Instead of trying to understand the feedback, some students use the Check button in a trial-and-error manner, when working with difficult or large exercises. By experience, we know that many students want to have a solution available, while they are working on exercises. But that makes it easy for students to "trick themselves" into thinking they have solved the tasks themselves, while they mainly have recreated another's solution. It is by purpose we don't give away complete solutions in LearnER. We want the students to actively create solutions by themselves.

5.3 Do the gamification elements contribute to learning, and if so, in what way?

Gamification has been shown to motivate and engage students in learning (Zainuddin et al., 2020). In the introduction we referred to MonstER Park (Schildgen, 2020). There, the entire application is developed as a game, guiding the students step-by-step through the phases of modeling.

We have taken a different approach, with a user interface looking more like a traditional modeling tool, but with several game mechanics added, such as progress bars and high score lists, where students earn points by solving tasks and loose points when they ask for help.

As reported in (Dæhli et al., 2018), students found LearnER to be a useful tool for learning basic data modeling skills. Many students reported that gamification, even though quite simple, were something that motivated them to do more exercises, in particular earning points and viewing and comparing their own and other student's results on high score lists. But they wanted clearer information about what led to points, such as solving many tasks, solving tasks quickly, etc. We also found that some students would be further motivated by more advanced game mechanics.

Students freely choose exercises marked with a level of difficulty. LearnER usage data shows that students work their way up from simple to more difficult tasks, as we expected and hoped for. The introduction of extra inadequate words in some exercises was appreciated.

6. Future work

The combination of gamification and formative feedback seems promising in learning tools (Fuchs and Wolff, 2016) and is not yet fully explored. We are looking for even more specific and context dependent ways of providing feedback to students. The game mechanics are simple and can be enriched along several dimensions (Toda, et al., 2019), e.g. lead the players from level to level based on their achievements, award them with badges, or adding more advanced forms of cooperation and competition.

We are also looking into ways to let students solve problems more freely, which means that they can choose names of entities and attributes, but still get feedback and help. The solution must then be extended with a matching algorithm, and one must solve problems related to typos, synonyms, word contractions and so on (Bian, 2019).

LearnER is a tailer-made tool for IT students, but we think this way of stimulating active learning can be applied also in other subjects. Within the field of ITS, data modeling can be considered an ill-defined domain (Fournier-Viger et. al., 2010), in the sense that modeling problems can have several valid solutions. We believe that the approach taken for developing LearnER, based on exercises having a single (but flexible) solution, combining gamification and formative feedback, can be used to build similar tools for other diagram types, such as flowcharts and different UML diagrams within the IT field, or even mind maps used as learning tools in various subjects.

Acknowledgments

The project has received financial support from Norwegian Agency for International Cooperation and Quality Enhancement in Higher Education (Diku) for 2018–2019 (grant 8600036). University of South-Eastern Norway and Kristiania University College have facilitated the project. We also wish to thank our developer Håvard Myrbakken for quality work.

References

Ali, N.H., Shukur, Z. and Idris, S., 2007. Assessment system for UML class diagram using notations extraction. *International Journal on Computer Science Network Security*, *7*, pp.181-187.

Alkhatlan, A. and Kalita, J., 2018. Intelligent tutoring systems: A comprehensive historical survey with recent developments. *arXiv preprint arXiv:1812.09628*.

Batmaz, F. and Hinde, C.J., 2006. A diagram drawing tool for semi–automatic assessment of conceptual database diagrams.

Batmaz, F., Stone, R. and Hinde, C., 2010. Personalized feedback with semi-automatic assessment tool for conceptual database model, in *Teaching and Learning in Information and Computer Sciences*, 9(1), pp 105-109.

Batra, D. and Antony, S.R., 1994. Novice errors in conceptual database design. *European Journal of Information Systems*, *3*(1), 57-69.

Bian, W., Alam, O. and Kienzle, J., 2019. Automated grading of class diagrams. In *2019 ACM/IEEE 22nd International Conference on Model Driven Engineering Languages and Systems Companion (MODELS-C)*, pp. 700-709.

Bian, W., Alam, O. and Kienzle, J., 2020. Is automated grading of models effective? assessing automated grading of class diagrams. In *Proceedings of the 23rd ACM/IEEE International Conference on Model Driven Engineering Languages and Systems*, pp. 365-376.

Bloom, B.S., 1956. Taxonomy of educational objectives: The classification of educational goals. *Cognitive domain*.

Bogdanova, D. and Snoeck, M., 2019. CaMeLOT: An educational framework for conceptual data modelling. Information and Software Technology, 110, 92-107.

Chen, P.P.S., 1976. The entity-relationship model—toward a unified view of data. *ACM transactions on database systems (TODS)*, *1*(1), pp. 9-36.

Codd, E.F., 1970. A relational model of data for large, shared data banks. *Comm. ACM 13, 6*, pp 377-387.

Duan, D., Mitrovic, A., and Churcher, N., 2010. Evaluating the effectiveness of multiple open student models in EER-tutor. In S. L. Wong et al. (Eds.), *International Conference on Computers in Education*, Putrajaya, Malaysia, Asia-Pacific Society for Computers in Education, pp. 86–88.

Dæhli, O., Kristoffersen, B., Lauvås Jr, P. and Myrbakken, H., 2018. A Supportive Web-Based Tool for Learning Basic Data Modeling Skills. In *ECEL 2018 17th European Conference on e-Learning* (p. 116). Academic Conferences and Publishing Limited.

Dæhli, O., Kristoffersen, B. and Sandnes, T., 2020. Lessons Learned from Developing and Evaluating an Educational Database Modeling Tool. In *European Conference on e-Learning* (pp. 129-XVI). Academic Conferences International Limited.

Fournier-Viger, P., Nkambou, R., and Nguifo, E. M. (2010). Building intelligent tutoring systems for ill-defined domains. In *Advances in intelligent tutoring systems* (pp. 81-101). Springer, Berlin, Heidelberg.

Fuchs, M. and Wolff, C., 2016. Improving programming education through gameful, formative feedback, in *2016 IEEE Global Engineering Education Conference (EDUCON)*, pp. 860-867. IEEE.

Halpin, T. and Bloesch, A., 1999. Data Modeling in UML and ORM: A Comparison. *Journal of Database Management (JDM), 10*(4), 4-13. doi:10.4018/jdm.1999100101

Hasker, R.W., 2011. UMLGrader: an automated class diagram grader. *Journal of Computing Sciences in Colleges*, *27*(1), pp.47-54.

Hattie, J. and Timperley, H., 2007. The power of feedback, in *Review of educational research*, Vol. 77, No. 1, pp 81–112.

Hoggarth, G. and Lockyer, M., 1998. An automated student diagram assessment system. *ACM SIGCSE Bulletin*, *30*(3), pp.122-124.

Jayal, A. and Shepperd, M.J., 2009. The problem of labels in E-assessment of diagrams, in *Journal on Educational Resources in Computing* (JERIC), 8(4), pp 12:1-12:13.

Kapp, K.M., 2012. *The gamification of learning and instruction: game-based methods and strategies for training and education*, John Wiley & Sons, San Francisco.

Keuning, H., Jeuring, J. and Heeren, B., 2018. A systematic literature review of automated feedback generation for programming exercises. *ACM Transactions on Computing Education* (TOCE), 19(1), 1-43.

Lino, A.D.P. and Rocha, A., 2018. Automatic evaluation of ERD in e-learning environments, in *2018 13th Iberian Conference on Information Systems and Technologies (CISTI)*, pp. 1-5. IEEE.

Lunt, B.M., Ekstrom, J.J., Gorka, S., Hislop, G., Kamali, R., Lawson, E., LeBlanc, R., Miller, J. and Reichgelt, H., 2008. *Curriculum guidelines for undergraduate degree programs in information technology*. Association for Computing Machinery.

Mathews, M., Mitrovic, A., Lin, B., Holland, J. and Churcher, N., 2012. Do your eyes give it away? Using eye tracking data to understand students' attitudes towards open student model representations. In *International Conference on Intelligent Tutoring Systems*, pp. 422-427. Springer, Berlin, Heidelberg.

Migler, A. and Dekhtyar, A., 2020. Mapping the SQL Learning Process in Introductory Database Courses. In *Proceedings of the 51st ACM Technical Symposium on Computer Science Education*, pp. 619-625.

Mitrovic A. and Holland J., 2020. Effect of Non-mandatory Use of an Intelligent Tutoring System on Students' Learning. In: Bittencourt I., Cukurova M., Muldner K., Luckin R., Millán E. (eds) Artificial Intelligence in Education. AIED 2020. *Lecture Notes in Computer Science*, vol 12163. Springer, Cham. https://doi.org/10.1007/978-3-030-52237-7_31

Menezes, C.C.N. and Bortoli, R.D., 2016. Potential of Gamification as Assessment Tool. *Creative Education*, 7(4), pp. 561-566.

Rosenthal, K. and Strecker, S., 2019. Toward a taxonomy of modeling difficulties: a multi-modal study on individual modeling processes.

Schildgen, J., 2020. MonstER Park-The Entity-Relationship-Diagram Learning Game.

Shute, V.J., 2008. Focus on formative feedback, in *Review of educational research,* Vol. 78(1), pp 153–189.

Simanjuntak, H. (2015) Proposed framework for automatic grading system of ER diagram, in *7th International Conference on* Information Technology and Electrical Engineering (ICITEE), pp 141–146.

Soler, J., Boada, I., Prados, F., Poch, J. and Fabregat, R., 2010. A web-based e-learning tool for UML class diagrams. In *IEEE EDUCON 2010 Conference*, pp. 973-979. IEEE.

Stikkolorum, D.R., Putten, P.V., Sperandio, C., and Chaudron, M., 2019. Towards Automated Grading of UML Class Diagrams with Machine Learning. *BNAIC/BENELEARN*.

Suraweera, P. and Mitrovic, A., 2002. KERMIT: A constraint-based tutor for database modeling. In *International Conference on Intelligent Tutoring Systems*, pp. 377-387. Springer, Berlin, Heidelberg.

Suraweera, P. and Mitrovic, A., 2004. An intelligent tutoring system for entity relationship modeling, in *International Journal of Artificial Intelligence in Education*, Vol. 14(3, 4), pp 375–417.

Task Group on Information Technology Curricula, 2017. Information technology curricula 2017: Curriculum guidelines for baccalaureate degree programs, ACM, New York, NY, USA.

Thomas, P., 2013. Online automatic marking of diagrams, in *Systemic Practice and Action Research*, Vol. 26(4), pp 349–359.

Thomas, P., Waugh, K. and Smith, N., 2006. Using patterns in the automatic marking of ER-diagrams. In *Proceedings of the 11th annual SIGCSE conference on Innovation and technology in computer science education*, pp. 83-87.

Toda, A. M., Klock, A. C., Oliveira, W., Palomino, P. T., Rodrigues, L., Shi, L., ... & Cristea, A. I. (2019). Analysing gamification elements in educational environments using an existing Gamification taxonomy. *Smart Learning Environments*, 6(1), 1-14.

Zakharov, K., Mitrovic, A. and Johnston, L., 2007. Pedagogical agents trying on a caring mentor role. *Frontiers in Artificial Intelligence and Applications*, *158*, p.59.

Zainuddin, Z., Chu, S.K.W., Shujahat, M. and Perera, C.J., 2020. The impact of gamification on learning and instruction: A systematic review of empirical evidence. *Educational Research Review*, *30*, p.100326.

