International Journal of Learning Analytics & Artificial Intelligence for Education

Volume No. 6

Issue No. 1

January - April 2024

ENRICHED PUBLICATIONS PVT.LTD

JE - 18,Gupta Colony, Khirki Extn,
Malviya Nagar, New Delhi - 110017.
E- Mail: info@enrichedpublication.com
Phone :- +91-8877340707

International Journal of Learning Analytics & Artificial Intelligence for Education

Focus and Scope

The journal of learning analytics and artificial intelligence for education is a peer-reviewed, open access journal that aim to disseminate highest quality research in the field. The journal aims to increase knowledge and understanding of ways in which learning analytics and artificial intelligence can support and enhance education.
The journal covers all advances in learning analytics and artificial intelligence for education, including but not limited to the following topics: tools and methodologies for learning analytics and artificial intelligence for education, applications of learning analytics and artificial intelligence for education in real-world settings, theoretical perspectives on learning analytics and artificial intelligence for education.

Fields of interest include, but are not limited to:

AI powered educational systems and tools
Analyzing Learner’s Activity, Engagement and Motivation
Modelling Learning and Teaching in Different Environments (online, blended and physical environments)
Evaluation and Development of Teaching Practices and Learning Designs
Application of Machine Learning, Educational Data Mining, and Predictive Analytics
Supporting and Developing Assessment and Feedback Practices
Supporting Personalized and Adaptive Learning
Analyzing and Improvement of Learning Organizations based on Data
Development of new Methods and Theories

International Journal of Learning Analytics & Artificial Intelligence for Education

(Volume No. 6, Issue No. 1, January - April 2024)

Contents

Mobile Learning Applications for Android und iOS for German Language Acquisition Based on Learning Analytics Measurements

Markus Friedl, Markus Ebner , Martin Ebner
Graz University of Technology, Graz, Austria

ABSTRACT

The use of digital media is increasingly being promoted in school teaching. Since this aspect changes the interaction between teachers and pupils, this research is concerned with the development of a prototype of a mobile ap plication for Android and iOS, in which different learning applications for lan guage acquisition are offered on the basis of learning analytical measurements provided by experts in the field. By logging and collecting interactions of the user, it is possible to create a variety of statistical evaluations and thus respond to the needs and weaknesses of students. For the evaluation of the application, a user experience test was carried out, whereby the child-friendly operation of the application was tested. Due to the very positive feedback, the design was found to be good and can therefore be further developed.

Keywords—Learning Analytics, Language Acquisition, Mobile Learning Ap plication, Digital Media in Teaching, Technology Enhanced Learning, Educational Data Mining

1 Introduction

In this day the use of digital media in school is steadily growing (Robin, 2008) and will play an ever-greater role in the future than it already is today. Instead of using classical learning materials, electronic aids such as computers are increasingly being used. Through the development of tablets and their widespread use, a new way of consuming digital media has been created. Due to their simple operation, most children of the youngest primary school age have already mastered the use of such devices, so that their use in the classroom is very obvious.

But s big problem is the changed communication between teacher and student. By solving tasks in digital form, they are usually corrected digitally, and the result is returned. This means that students often only see the result of their tasks without an oral explanation. It is also difficult to generate and provide individual tasks for classes or even individuals as platforms do not support this.

In order to counteract this aspect, a prototype of a mobile application for Android and iOS was developed in the course of this research. With the help of the application different exercise formats and different tasks will be offered to learn the German lan guage based on Learning Analytics Measurements.

The exercises will be used by the IDeRBlog platform (Aspalter, Edtstadler, and Martich, 2017), which will be created and maintained by experts and teachers respec tively. This platform provides an intelligent dictionary containing different data of each user, which will also be used in the app in the future.

Since all exercises and tasks are loaded by an API from an external server, the main focus of the app was a simple handling and comprehensible presentation of the contents. Particular attention was paid to child-friendly operation, as the main target group of the app are children of primary school age. In order to realize this aspect, a User Experience (UX) test with children from Austria was carried out. The task was that each child tried out all exercise formats and completed a short questionnaire after each exercise.

In addition, a discussion was held about which things might be unclear or badly formulated and whether there were any suggestions for improvement. Since the feedback from the test was very positive and the children had a lot of fun practic ing, the user-friendly design was confirmed and for this reason the prototype can be further developed in this way.

2 Related Work

2.1 Digital media in teaching

In general teaching can be explained as a process of interactions between teachers and their pupils. This process is supported by various activities by teachers with spe cific skills (Tulodziecki, Herzig, and Blömeke, 2017). The effect of digital media in teaching is influenced by different potential factors. These factors can be divided into following points (Herzig, 2014):

- The digital media and the content which is contained by the media
- Media offerings integrates different teaching processes
- The different persons which are involved in this process, in this area these are teachers and their pupils

According to (Herzig, 2014) the context of these factors there are four categories which serve to fulfil the factors. A digital media can have individual meanings, or an offer also can cause an interaction which have specific characteristics. These characteristics have more or less influence on the teaching process. Teachers usually have different areas of knowledge in which they are experts, and their expertise influences the way they teach and learn. Students also differ greatly in their knowledge, receptiveness to new knowledge and educational proximity. These factors show that the effect of digital media does not only depend on their content.

By using such media, the learning performance is to be increased in comparison to classical media, whereby an added value can be achieved. If, however, these media are too monotonous, this impairs the motivation of the learners and causes the oppo site of a worse performance. The aim is to increase the learning performance and the success with unvarying effort or have a constant success in learning with less exertion (Kerres, 2002).

2.2 Learning analytics

With LA, large amounts of data are collected and evaluated so that the results can have a positive influence on the teaching process in the future. The terms Technology Enhanced Learning (TEL) and Educational Data Mining (EDM) are often associated with LA. LA can be seen as a subdivision of TEL and has been a new trend since it was mentioned a few years ago (Johnson, Adams, and Cummins, 2012). EDM deals with the storage of data and should answer the question of which data is important and meaningful in order to gain added value from LA (Duval, 2011).

Learning Analytics (LA) can be seen as a process with iterative steps. These steps can be divided into the following steps (Chatti et al., 2013):

- Data collection and pre-processing: In the first step, the data is collected and con verted to a uniform format. The final result is influenced by the quality of the data.
- Analytics and action: With the aim of various techniques the collected data is analyzed and visualized. This step is aimed at increasing the effectiveness of learning.
- Post-processing: If the result obtained is not satisfactory, new data is collected in this step and the process starts again with the first step.

In addition, there are various drivers that can be used to better explain LA in the learning area. However, these drivers do not only apply to education, but are also used in many other areas and describe the complexity of LA. By the use of Virtual Learning Environment (VLE) systems large amounts of data, which is called big data, are generated. Therefore, it is important to create systems that can handle, analyze and process this amount of data within a reasonable time. It is also important that teachers receive feedback from online courses offered in order to evaluate and, where appropriate, improve the teaching process. Conversely, it is also important that students receive help from their teachers, even if they are not physically present. As a third major point, it is of crucial importance that this system is supported by politicians, as this process requires many resources and large investments.

2.3 Mobile language learning applications

Many mobile applications for language learning are offered. In relation to the developed prototypes, comparable applications were tested to get an overview and possible design suggestions. The focus was also on whether the applications could be used in schools.

What was immediately noticed during the research was that most applications offer very little content without payment. In addition, advertisements are often displayed, which makes their use in schools impossible. It should also be noted that the main target group of applications varies greatly. Many applications, which are designed for children, contain a lot of playful elements and animations, whereby the focus on learning often fades into the background and children are easily distracted. Other applications are mainly made for adults, but the design is often not very understandable for children. Furthermore, it is often difficult that the mobile applications can be used in schools, because teachers rarely have access to the data of their students and can see or this functionality is not available at all.

3 Research Design

Many applications have been developed at the Graz University of Technology (TU Graz) in the past (Ebner, 2015). Each application must be evaluated by a field test to ensure that the application meets the specified requirements. Since all exercise formats and tasks are given by professors, the focus is on testing the design of the developed application. In order to evaluate the exercise format and its design, a questionnaire is completed after an exercise has been completed. These answers do not provide representative data in order to obtain a high-quality evaluation. Furthermore, the application must be tested by several children over a longer period of time in order to get better results.

In order to start the test, each child was provided with an iPad on which the application was already installed and ready for use. The children were then asked to select an exercise format one after the other and start an exercise. In order not to lose the overview, all children were asked to start the same exercise format and wait until everyone had finished the exercise. During the exercise, notes were written down to document the different reactions or possible problems that occurred. At the end of the exercise, each child fills in a questionnaire (Fig. 1), where they are asked to answer, whether the task was comprehensible, could be solved independently, whether the exercise was fun and whether the exercise would also like to be done at home. In order to make the evaluation understandable, smileys were used as a scale, whereby the laughing smiley means very good and the sad smiley not good. At the beginning of the test it could be noticed that the children felt that their performance was being evaluated and that the notes written refer to their solutions. Therefore, it was important to mention several times that only the application should be evaluated and that the children cannot do wrong. After a short period of familiarization, the children's nervousness subsided, and the application was tested with full concentration.

In addition to the questionnaire, the children in the group were asked if anything could be improved in the design or if anything was unclear. Then the same procedure was repeated with the next exercise format until all exercise formats were evaluated.

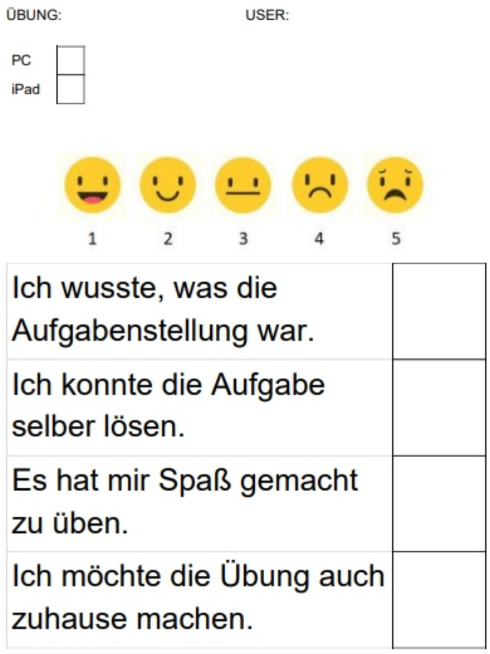

Fig. 1. UX-Feedback questionnaire

4 Prototype – IDeRBlog ii

The IDeRBlog ii application is a native developed application for Android and iOS, which is designed for the use on tablets. The application is designed to support the learning of the German language based on LA measurements. In the application one has the possibility to select different exercise formats and to start tasks with different difficulty levels after the Login by means of username and password. The user login is required, since the application can currently only be used by members of the IDeRBlog platform. In addition, this platform provides a separate intelligent dictionary for each user, which will be used in the application in the future.

The application provides five different exercise formats as shown in Fig. 2, which contain any number of exercises in different difficulty levels. The first format in the upper right corner is called “Match” and in this exercise, pupils must assign a given word to different word categories. There are two basic types of the exercise of assigning a word. The first is by drag and drop a box to the correct categories. With the second type, a displayed word must be assigned to the correct categories and typed in correctly. After checking each task, the user receives feedback as to whether the task was solved correctly or incorrectly. If everything is correct, this is indicated by text and the exercise can be continued. If a mistake has been made, the user will be shown the correct solution as shown in Fig. 3. The user receives such feedback after each task has been solved, and any incorrect information is transferred to an external server for evaluation.

Fig. 2. Overview of the exercise formats

In the second exercise format (Fig. 2) in the lower left corner, which is called "Wheel of Fortune", pupils must type an auditioned word into an input field. The word can be auditioned as often as needed before it is entered. This exercise can be made more difficult by setting a timer that limits the time of input. If the word is entered correctly, it will be shown to the Student. Otherwise, the wrong input and the correct word will be displayed one below the other as feedback.

The third format with the title "Keywords" in the middle of the screen (Fig. 2) again has two basic configurations of the exercise. In the first configuration, a word is displayed which is divided into the individual letters or graphemes and the student must mark difficult places in the word. Once the word has disappeared, it must be entered correctly. In the second configuration, a timer limits the memorization of the word before it must be re-entered. The feedback after solving a task is identical to the previously mentioned feedback.

In the upper right corner (Fig. 2) is the "Insert" exercise format, where sentences with text gaps are displayed. The task is to fill in these gaps correctly so that the sentence makes sense. A description text can be displayed as an aid. When checking a sentence, correctly filled gaps are displayed in green and incorrect ones in red. The pupil cannot continue with a new sentence until all the gaps have been filled in correctly.

The last format in the lower right corner (Fig. 2) is called "Error locations". In this exercise, a sentence is displayed again, but this time it contains incorrect words. The pupil is asked to find and mark these wrong words. If an incorrect word is selected, an additional window will appear in which the reasons must be given by multiple choice selection for the correct spelling.

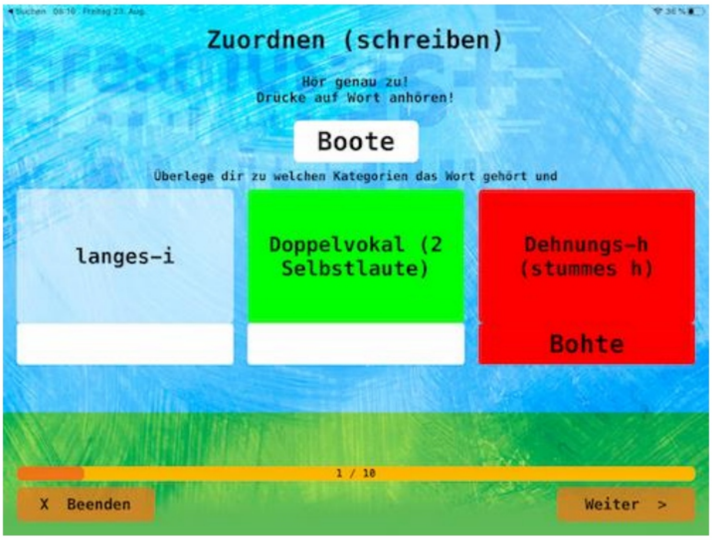

Fig. 3. Right answer in green with the wrong answer in red as feedback after evaluating a task.

By correctly solving tasks, the pupil has the opportunity to collect points, which are displayed in the overview (Fig. 2). These points can be used to unlock tasks within different exercise formats.
The same database is currently used as the exercise content for all users, whichmeans there is currently no possibility to create individual exercises for a pupil.
As already mentioned, in the future the intelligent dictionary of the IDeRBlog plat form will be integrated, so that every user can practice all individual words and sentences. This ensures that specific weaknesses can be practiced and eliminated.

5 Field Study

As a test group for the evaluation of the application, children from a third-grade el ementary school in Austria should serve, which unfortunately did not come about due to technical problems or a lack of equipment. Therefore, a smaller group with five children aged between eight and eleven years was chosen as a replacement. The five children included two girls and three boys. Three iPads were used as test devices, on which the application was installed and ready for operation. Since the application can also be used offline, i.e. without an Internet connection, there is no need for an active Internet connection. Only the data sent to an external server and stored in a database could not be uploaded. Since not every child had an iPad, the children were divided into a group of three and a group of two children.
Before the test was started, the reason and the procedure for this evaluation were briefly explained. It was also mentioned several times that not the performance of the children, but the application should be evaluated, since the question often arose whether something happened if an exercise was not solved correctly.
After all the questions had been answered, the first group was asked to start the first exercise format.
It was immediately obvious that all children knew how to operate tablets from home or school and that the application was quickly understood.
After one exercise was completed, each child was asked to complete the question naire in Figure 1. In addition, questions were asked about the design and possible suggestions for improvement before the next exercise format was started.

The following questions were prepared:

I had understood the assignment definition.
I was able to solve the problem myself.
I had fun to practice.
I would also like to do the exercise at home.

The grading of the individual questions was done with the help of numerical values and smileys, where 1 or a laughing smiley means very good and 5 or a sad smiley does not mean good.
During the exercise, notes were written down again and again to document the be havior or possible problems which were taken into account in the further development of the application.
After about 40 minutes the first group was finished with all exercise formats and then the second group started the same procedure, which lasted about 40 minutes.
At the end a joint discussion was started again, which resulted in only positive feedback from the children. The question was even asked when the application would be available for everyone so that the children could continue to use it in the future. A summary of all questions and the calculated mean value is shown in Table 1.

Table 1. Result of the evaluation questionnaire

	I had understood the assignment definition.	I was able to solve the problem myself.	I had fun to practice.	I would also like to do the exercise at home.
Error locations	1	1	1.2	1.2
Insert	1.2	1	1.2	1.2
Keywords - Easy	1	1	1	1
Keywords - Hard	1	1	1	1
Match – Drag & Drop	1	1	1	1.2
Match - Write	1	1	1	1.5
Wheel of Fortune	1	1	1	1
Average	1.03	1	1.06	1.16

6 Conclusion

This research deals with the use of a mobile application for Android and iOS to learn the German language based on LA measurement and to address primary school children in Austria as the main target group. A prototype of the application was de veloped, which was designed to be easy to use and child friendly. The design was evaluated with the help of a UX test of five children, which provided valuable feedback and was taken into account in the further development of the application. All exercise formats were specified by professors from the University College of Teacher Education Styria and together we developed the design and discussed how each exercise format could be implemented. The chosen design of the application and the user interface had to be as simply and understandable as possible, so children have no problems with the use of the application. The exercise data will be provided by the IDeRBlog platform, and the intelligent dictionary of this platform will also be used in the future. By this aspect it is then possible to make individual tasks available for each user together and these in the application, which are to go into the weaknesses of each individual purposefully and improve these.
Based on the very positive feedback, the decision can be made that the design con cept of the first prototype can be described as well-chosen and it can be built on. However, it should be mentioned that

the number of test users should be larger in order to get more representative data by the UX test.

7 References

[1] Aspalter, Christian, Edtstadler, Konstanze and Martich, Susanne (2017). "Recht-/schreibdidaktische Uberlegungen fur̈ den Einsatz der Internetplattform IDeRBlog" In: Open Online Journal for Research and Education (cit. on p. 1-7)

[2] Chatti, Mohamed Amine et al. (2013). "A reference model for learning

[3] Analytics." In: International Journal of Technology Enhanced Learning 4.5-6,

[4] pp. 318–331 (cit. on pp. 9–14)

[5] Duval, Erik (2011). "Attention please!: learning analytics for visualization and recommendation." In: LAK 11, pp. 9–17 (cit. on p. 7). https://doi.org/10.1145/2090116.2090118

[6] Ebner, M. (2015). Mobile Learning and Mathematics. Foundations, Design, and Case Studies. Crompton, H., Traxler, J. (ed.). Routledge. New York and London. pp. 20-32

[7] Herzig, Bardo (2014). "Wie wirksam sind digitale Medien im Unterricht." In: Gütersloh: Bertelsmann Stiftung (cit. on p. 4)

[8] Kerres, Michael (2002). "Bunter, besser, billiger? Zum Mehrwert digitaler Medien in der Bildung (Why? On the Benefit of Digital Media in Educa- tion)." In: it-Information Tech nology 44.4, pp. 187–192 (cit. on p. 6). https://doi.org/10.1524/itit.2002.44.4.187

[9] Robin, Bernard R. (2008). "Digital Storytelling: A Powerful Technology Tool for the 21st Century Classroom." In: Theory into Practice 47.3, pp. 220–228. (cit. on p. 3). https://doi.org/10.1080/00405840802153916

[10] Tulodziecki, Gerhard, Bardo Herzig, and Sigrid Blömeke (2017). "Gestaltung von Unterricht: eine Einführung in die Didaktik." In: 3311 (cit. on p.3)

8 Authors

Markus Friedl studied Software Engineering and Management at Graz University of Technology, Graz, Austria.

Markus Ebner is currently working as a Researcher in the Department Educational Technology at Graz University of Technology. He deals with e-learning, mobile learning, technology enhanced learning and Open Educational Resources. His focus is on Learning Analytics at K-12 level. In addition, several publications in the area of Learning Analytics were published and workshops on the topic were held.

Martin Ebner is with the Department Educational Technology at Graz University of Technology, Graz, Austria. (E-mail: martin.ebner@tugraz.at). As head of the De partment, he is responsible for all university wide e-learning activities. He holds an Assoc. Prof. on media informatics and works at the Institute of Interactive Systems and Data Science as senior researcher. For publications as well as further research activities, please visit: http://martinebner.at. Email: martin.ebner@tugraz.at

Fearing the Robot Apocalypse Correlates of AI Anxiety

David John Lemay, Ram B Basnet, Tenzin Doleck
McGill University, Montreal, Quebec
Colorado Mesa University, Colorado, USA
University of Southern California, California, USA

A B S T R A C T

This study examines the relationship between individuals' be liefs about AI (Artificial Intelligence) and levels of anxiety with respect to their technology readiness level. In this cross-sectional study, we surveyed 65 stu dents at a southwestern US college. Using partial least squares analysis, we found that technology readiness contributors were significantly and positively related to only one AI anxiety factor: socio-technical illiteracy. In contrast, all four links between technology readiness inhibitors and AI anxiety factors were significant with medium effect sizes. Technology readiness inhibitors are positively related to learning, fears of job replacement, socio-technical illiteracy, and particular AI configurations. Thus, we conclude that AI anxiety runs through a spectrum. It is influenced by real, practical consequences of immedi ate effects of increased automatization but also by popular representations and discussions of the negative consequences of artificial general intelligence and killer robots and addressing technology readiness is unlikely to mitigate effects of AI anxiety.

Keywords—Artificial intelligence, anxiety, technology readiness contributors,technology readiness inhibitors, technology dispositions.

1 Introduction

A quick search shows that people are generally ambivalent about the advent of arti ficial intelligence (AI). On the one hand, individuals perceive advantages afforded by AI applications such as recent advances in image and voice recognition. On the other, they are cognizant of attendant negative consequences of increased automation suchas job displacement and erosion of human rights to privacy, liberty, and agency.

Popular representations of killer robots and the enslavement of humanity to technology contribute to the general distrust. Present controversies over social media technol ogy's stewardship of privacy and public discourse and government supervision of its citizens render notions of technological utopianism naive at best and harmful at worst.

AI anxiety is an increasingly recognized phenomenon as individuals grapple with a changing present and an uncertain future [1]. To address AI anxiety, it is important to examine the correlates of AI anxiety to identify mechanisms that can mitigate distress and better manage individuals' emotions and perceptions.

Johnson and Verdicchio [1] have argued that much of AI anxiety is overblown and can be attributed to three factors: an exclusive focus on AI programs that leaves humans out of the picture, confusion about autonomy in computational entities and in humans, and an inaccurate conception of technological development. They conclude there are good reasons to worry about AI but not for the reasons advanced by AI alarmists.

In an Anthropology of Robots and AI: Annihilation Anxiety and Machines, Richardson [2] discussed the

existential fears ("annihilation anxieties") that another intelligence poses to humans as represented by popular fiction. The author argued that there is a resonant relationship between our fiction of robots and our "lived realities of robotic practices" that feedback into each other influencing our experiences of both.
As our fears often reflect ourselves more than anything else, perhaps our AI anxiety stems from our primal nature, motivated by dominance and our own history of genocide, subjugation, and intolerance of others. But also, the fear of being made redundant, replaced, or erased. As Richardson [2] writes about R.U.R. (Rossum's Universal Robots) by Karel Capek, the first play to feature a robot: "The robot [...] is a device to explore the fears of terminus in human existence brought about by mechanization, political ideologies and high modernism, and it speaks to the theme of humanity's end" (p. 2). Discussions of AI and robots inevitably are faced with the dehumanizing effects of technology, as has been a prevailing theme of fiction in the modern, industrial era.

Faced with contemporary realities of job loss and economic insecurity brought upon workers worldwide from increasingly automated, globe-spanning production chains, it is not surprising that many might be skeptical about the benefits of AI and question to whom those benefits might accrue. Contemporary debates about technolo gy and high-profile critics [3, 4] stoking fears of an inevitable robot apocalypse—if artificial general intelligence were ever created—contribute to the distrust and the malaise. Privately funded AI research labs with little governmental oversight and publicly funded surveillance do not instill trust either, especially when new research continues to make huge progress on tasks that were once believed to be the exclusive remit of humans and are now mastered by machines. However, the end of humanity could be much more mundane, as reported by Achenbach [3] in the The Washington Post:

"The world's spookiest philosopher is Nick Bostrom […] in his mind, human extinction could be just the beginning. Bostrom's favorite apocalyptic hypothetical involves a machine that has been programmed to make paper clips This machine keeps getting smarter and more powerful, but never de velops human values. It achieves "superintelligence." It begins to convert all kinds of ordinary materials into paper clips. Eventually it decides to turn eve rything on Earth — including the human race (!!!) — into paper clips. Then it goes interstellar."

Despite these prognostications of doom, a 2018 Workforce Institute survey of 3,000 individuals found that four out of five employees (82%) saw AI as an opportunity to improve their jobs, while only a third (34%) worried that AI might replace them at work [5]. Whereas a Canadian government policy paper [6] found a correlation between Canadians' fear of losing jobs due to automation and populist and nativist views, but that Canadians supported traditional government policy solutions such as workforce retraining more than limiting labor mobility.

1.1 Research question

The diversity of reactions to AI and general confusion about the state of AI technological development compelled us to examine the range of antecedent factors that might help mitigate current popular confusion about the state of AI and individual AI anxiety. Thus, we sought to understand the relationship between technology readiness and AI anxiety. We asked two closely related questions:

a) What are the relationships between technology readiness contributors and AI anxiety?

b) What are the relationships between technology readiness inhibitors and AI anxiety?

2 Background

Given current debates on AI impact on society for good and for bad, Wang and Wang [7] developed an AI anxiety scale to assess to what degree these uncertainties provoke an existential malaise and anxiety for the future and how to mitigate negative consequences for individuals. Wang and Wang [7] situate AI anxiety with respect to technophobia which they define as an irrational fear of technology characterized by negative attitudes toward technology, anxiety about the future impacts of advancing technology, and self-admonishing beliefs about their ability. They divide AI anxiety in two aspects, computer anxiety [1] and robot anxiety [8]. They term AI anxiety as a distinct and independent variable. They define AI anxiety as "an overall, affective response of anxiety or fear that inhibits an individual from interacting with AI. Thus, AIA may be operationally considered as a general perception or belief with multiple dimensions" (p. 3). Informed by the theory of reasoned action [9], Wang and Wang [7] argue that it is necessary to address negative affect as it is known to negatively affect future performance. The scale is established to measure the degree of motivated learning behavioral intention as anxiety can be a facilitative, motivational factor for proactively addressing anxiety by becoming better informed. Although it can also be an obstacle to action. Hence, we expect a different behavioral profile for divergent reactions to anxiety.

As Haring et al. [8] noted, the research to date has demonstrated the cultural variability of reactions to technology, however, it is precisely the fact of the variability of social groups' reactions to AI that merits study to aid in determining HR policies for managing workforces worldwide. However, cultural variability alone cannot explain all the remaining variance as Wang and Wang [7] only found low correlation (r=0.19) between the AI construct and behavioral intentions. However, such low correlations are very common in the attitude-intention literature [9]. It is likely the case that there are other sources of unaccounted variability that influence individuals' affect and behavior. Such individual differences and social determinants can be modelled and their influences accounted for [10-12].

Wang and Wang [7] found that AI anxiety was facilitative to some extent as it appeared to influence motivated learning behaviors. We wished to determine how enabling and inhibiting determinants might interact to inform a range of behavioral profiles and responses to AI anxiety. Following Khatri, Samuel, and Dennis [13], we sought to determine the extent that an individual's technology-based predispositions (technology readiness contributors and technology readiness inhibitors) might influence their AI anxiety.

3 Method

3.1 Research design

The present exploratory study employs a cross-sectional survey design and partial least squares modeling to assess the influence of technology-based predispositions (technology readiness contributors and technology readiness inhibitors) on AI anxiety.

3.2 Participants and procedure

We drew on data from computer science students enrolled in a southwestern college in the US. In total, 65 students participated in this study. The sample comprised 8 females and 57 males with an average age of 23.86 years (SD=6.09). Students were asked to voluntarily participate in the study. No compensation was provided for participation in his study. Students participating in the study were emailed a link to a self-report questionnaire.

3.3 Measure

Along with demographic information, participants responded to statements related to the study measures: AI anxiety scale and Technology Readiness Contributors and Technology Readiness Inhibitors scales.

Scales were gathered from previous literature. AI Anxiety scale [7] was used to evaluate an individual's AI anxiety. The AI anxiety scale contains 21 items measuring the following four factors: Learning; AI Configuration; Job Replacement; and, Sociotechnical Blindness, however we prefer the more neutral term Social-Technical Illiteracy, which we also feel is the more descriptive expression. The items were measured on a 7-point Likert scale (1=Strongly Disagree; 7=Strongly Agree). To measure individual's enabling and inhibiting predispositions, we adapted the Technology Readiness Contributors (TRC) and Technology Readiness Inhibitors (TRI) scales [13]; for each of the items, students responded on a 7-point Likert scale (1=StronglyDisagree; 7=Strongly Agree).

4 Analysis and Results

4.1 Analytic plan

The variance based partial least squares structural equation modeling (PLS-SEM) [14] was used to model and analyze the links between the constructs. PLS-SEM can simultaneously estimate both the measurement model and the structural model, and is a widely used estimation method in educational technology research [15-16]. The analysis was conducted via a two-step procedure: first assessing the measurement model and then the structural model. WarpPLS software [17] was used for the analysis of the measurement and structural model.

4.2 Measurement model

Prior to the structural model analysis, we assessed the psychometric properties of the measurement model. We followed the measurement model evaluation guidelines suggested in the literature [14, 18]. In Table 1, we find that the model-fit indices meet the suggested acceptance levels [18].

Table 1. Model Fit Statistics

Measure	Values	Recommended Criterion
Average path coefficient (APC)	0.273, *P*=0.005	Acceptable if *P*<0.05
Average R-squared (ARS)	0.192, *P*=0.026	Acceptable if *P*<0.05
Average adjusted R-squared (AARS)	0.166, *P*=0.041	Acceptable if *P*<0.05
Average block VIF (AVIF)	1.011	Acceptable if <=5
Average full collinearity VIF (AFVIF)	1.748	Acceptable if <=5

The evaluation of the measurement model entailed examining reliability, convergent validity, and discriminant validity [14, 18]. Loadings were greater than or equal to 0.5 (along with p-values less than or equal to 0.5). Internal consistency reliability was established (composite reliability coefficients of the measures were greater than the threshold value of 0.70). All average variance extracted (AVE) values exceeded the recommended threshold value of 0.50. Discriminant validity was also assessed, using the Fornell-Larcker criterion [19]. Table 2 presents the correlation matrix for the Fornell-Larcker criterion. We find that all the diagonal values are greater than the off-diagonal numbers in the corresponding rows and columns, thus, the requirements of the Fornell-Larcker criterion were met and discriminant validity was confirmed. In sum, the constructs were empirically established to be both reliable and valid.

Table 2. Discriminant Validity Test

	TRC	TRI	Learning	Job Replacement	Socio-Technical Illiteracy	AI Configuration
TRC	(0.732)	-0.006	-0.125	0.076	0.235	-0.089
TRI	-0.006	(0.743)	0.426	0.413	0.322	0.265
Learning	-0.125	0.426	(0.856)	0.454	0.482	0.552
Job Replacement	0.076	0.413	0.454	(0.783)	0.589	0.630
Socio-Technical Illiteracy	0.235	0.322	0.482	0.589	(0.747)	0.546
AI Configuration	-0.089	0.265	0.552	0.630	0.546	(0.925)

4.3 Structural model

In the second stage, the relationships between the constructs in the research model were ascertained by evaluating the structural model. Variance Inflation Factors (VIF) values were inspected to check for potential multicollinearity problems. All VIF values were below the suggested threshold of 5, thus, there was no indication of multicollinearity. At the same time, since Q2 coefficient values were greater than zero, there was an acceptable level of predictive relevance [18].

The structural model was assessed through (see Table 3): path coefficients (β), path coefficients' significance levels (p-value), and effect sizes (f^2). Note that for assessment of f^2, values of 0.35, 0.15, and 0.02 indicate large, medium, and small effect sizes, respectively [20].

The path coefficients were assessed to determine the significance of the relationships between the constructs. As indicated in Table 3, for the links between TRC and AI anxiety factors, the results show that only TRC was significantly (and positively) related to Socio-Technical Illiteracy ($\beta=0.256$, $p=0.014$), with a small effect size. In contrast, all four links between TRI and AI anxiety factors were significant with medium effects: TRI was significantly (and positively) related to Learning ($\beta=0.474$, $p<0.001$); TRI was significantly (and positively) related to Job Replacement ($\beta=0.404$, $p<0.001$); TRI was significantly (and positively) related to Socio-Technical Illiteracy ($\beta=0.330$, $p=0.002$); and, TRI was significantly (and positively) related to AI Configuration ($\beta=0.345$, $p=0.001$).

Table 3. Path Testing Results

Path	Path coefficient (*β*)	*P value*	Effect size (f^2)	Result
TRC→Learning	-0.184	*P*=0.059	0.035	Not Significant
TRC→Job Replacement	0.059	*P*=0.314	0.008	Not Significant
TRC→Socio-Technical Illiteracy	0.256	*P*=0.014	0.073	Significant
TRC→AI Configuration	-0.130	*P*=0.139	0.020	Not Significant
TRI→Learning	0.474	*P*<0.001	0.225	Significant
TRI→Job Replacement	0.404	*P*<0.001	0.167	Significant
TRI→Socio-Technical Illiteracy	0.330	*P*=0.002	0.116	Significant
TRI→AI Configuration	0.345	*P*=0.001	0.122	Significant

5 Discussion

This study examined the relationship between individuals' beliefs about AI (Artificial Intelligence) and levels of anxiety with respect to their technology-based predispositions (technology readiness contributors and technology readiness inhibitors). We surveyed 65 students at a southwestern US college and found that TRC were significantly and positively related to socio-technical illiteracy. In contrast, all four links between TRI and AI anxiety factors were significant with medium effect sizes:

TRI was positively related to learning, fears of job replacement, socio-technical illiteracy, and specific AI configurations. Thus, we conclude that AI anxiety runs through a spectrum and is influenced by real, practical consequences of immediate effects of increased automatization but also influenced by popular representations and discussions of the negative consequences of artificial general intelligence and killer robots.

Both TRC and TRI influenced socio-technical illiteracy. This could potentially be explained by a mediating influence; however, we did not include mediators in the present study. Learning appeared to mitigate TRI but not TRC. This is not surprising since assessing the limits of technology is expected to lessen AI anxiety whereas discussing its potential can actually increase fears. In fact, our results tend to indicate however counterintuitively that discussions of technology readiness contributors and inhibitors can actually increase AI anxiety and fears of job replacement and the actual diversity of AI configurations. Thus, our results are contrary to Johnson and Verdic chio's [1] contention that clarification of the true status of artificial intelligence would allay fears and AI anxiety would recede. It would appear that the actual state of AI and the wide diversity of current AI applications—that AI is actually 'eating the world'—is concerning for individuals in general. Our results tend to support Wang and Wang's [7] findings that AI anxiety can have a facilitative effect and support motivated learning. However, the results do not show the expected beneficial effects from motivated learning, as socio-technical illiteracy is related to both TRC and TRI.

Whereas Khatri, Samuel, and Dennis [13] have argued in favor of a two-system behavioral model in technology acceptance research, where individuals are both in conscious, deliberative (system 2) thinking and unconscious, automatic (system 2) behavior. They argue that our default mode is unconscious, or automatic (system 1) and is influenced by past experience and individual preferences and beliefs; system 2 can influence system 1 through effortful practice. Coming from a more social per spective, automatic behavior that characterizes system 1 behavior is not simply unreflective autonomic behavior (we are not ready to countenance zombie computer users devoid of consciousness or agency/will); c.f. [21]. Rather, an important proportion of system 1 behavior can be seen as socially learned, influenced by customs, values, and other social imperatives that dictate appropriate and expected reactions. A more social perspective would term such behavior as "transparent" to conscious [22-23], cognitive processes, because they are not individually determined but socially learned through reflexive activity [24]. Thus, it ought not appear to be a matter of preference alone. The argument inherent in addressing AI anxiety through motivated learning is that an individual's predispositions can be changed by increasing awareness and rational deliberation. Such behavior and perceptions are not automatic in the mode of autonomic physiological processes; perceptions can be apprehended and responses adapted. Their constructed nature suggests that inhibitors and enablers might reveal hidden dimensionality beyond individual self-perceptions that are the focus of Khatri, Samuel, and Dennis's [13] two-system approach.

5.1 Limitations

This study is limited by design. As a cross-sectional study, we are limited to exam ining correlational relationships, no causal relationships can be inferred from these findings. The study is also limited in the number of explanatory variables. A socially oriented study of technology perceptions and beliefs [15] may suggest other enabling and inhibiting variables that could explain the variance in beliefs of AI anxiety.

5.2 Future directions

Future studies may employ experimental designs to infer causality, and may employ different statistical learning approaches to build models that account for the hid den dimensionality in enabling and discouraging variables influencing technology readiness and AI anxiety.

6 References

[1] D. Johnson and M. Verdicchio, "AI Anxiety", Journal of the Association for Information Science and Technology, vol. 68, no. 9, pp. 2267-2270, 2017. Available: 10.1002/asi.23867.

[2] K. Richardson, An anthropology of robots and AI. Routledge, 2015.

[3] J. Achenbach, "The A.I. Anxiety: Big-name scientists worry that runaway artificial intelli gence could pose a threat to humanity. Beyond the speculation is a simple question: Are we fully in control of our technology?", Washington Post, 2015.

[4] M. Tegmark, Life 3.0: Being Human in the Age of Artificial Intelligence. New York (New York): Alfred A. Knopf, 2017.

[5] Workforce Institute, "Can Artificial Intelligence Make Work Better?", 2018.

[6] P. Loewen and B. Stevens, "Automation, AI, and Anxiety: Policy Preferred, Populism Pos sible", Public Policy Forum, Ottawa, Ontario, 2019.

[7] Y. Wang and Y. Wang, "Development and validation of an artificial intelligence anxiety scale: an initial application in predicting motivated learning behavior", Interactive Learning Environments, pp. 1-16, 2019. Available: https://doi.org/10.1080/10494820.2019.167 4887

[8] K. Haring, C. Mougenot, F. Ono and K. Watanabe, "Cultural Differences in Perception and Attitude towards Robots", International Journal of Affective Engineering, vol. 13, no. 3, pp. 149-157, 2014. Available: https://doi.org/10.5057/ijae.13.149

[9] M. Fishbein, "Predicting and Changing Behavior", 2011. Available: 10.4324/9780203838 020.

[10] D. Lemay, T. Doleck and P. Bazelais, "Do instrumental goal pursuits mediate feelings of envy on Facebook and Happiness or subjective well-being?", Computers in Human Behavior, vol. 91, pp. 186-191, 2019. Available: https://doi.org/10.1016/j.chb.2018.09.043

[11] T. Doleck, P. Bazelais and D. Lemay, "Examining CEGEP students' acceptance of computer-based learning environments: A test of two models", Education and Information Technologies, vol. 22, no. 5, pp. 2523-2543, 2016. Available: https://doi.org/10.1007/ s10639-016-9559-9

[12] T. Doleck, P. Bazelais and D. Lemay, "Examining the antecedents of social networking sites use among CEGEP students", Education and Information Technologies, vol. 22, no. 5, pp. 2103-2123, 2016. Available: https://doi.org/10.1007/s10639-016-9535-4

[13] V. Khatri, B. Samuel and A. Dennis, "System 1 and System 2 cognition in the decision to adopt and use a new technology", Information & Management, vol. 55, no. 6, pp. 709-724, 2018. Available: https://doi.org/10.1016/j.im.2018.03.002

[14] J. Henseler, G. Hubona and P. Ray, "Using PLS path modeling in new technology re search: updated guidelines", Industrial Management & Data Systems, vol. 116, no. 1, pp. 2-20, 2016. Available: https://doi.org/10.1108/imds-09-2015-0382

[15] D. Lemay, T. Doleck and P. Bazelais, "Context and technology use: Opportunities and challenges of the situated perspective in technology acceptance research", British Journal of Educational Technology, vol. 50, no. 5, pp. 2450-2465, 2019. Available: https://doi.org/ 10.1111/bjet.12859

[16] T. Doleck, P. Bazelais and D. Lemay, "The role of behavioral expectation in technology acceptance: a CEGEP case study", Journal of Computing in Higher Education, vol. 30, no. 3, pp. 407-425, 2017. Available: https://doi.org/10.1007/s12528-017-9158-9

[17] N. Kock, WarpPLS. http://www.warppls.com, 2019.

[18] N. Kock, "WarpPLS 7.0 user manual", 2020.
[19] C. Fornell and D. Larcker, "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error", Journal of Marketing Research, vol. 18, no. 1, pp. 39-50, 1981. Available: https://doi.org/10.1177/002224378101800104
[20] J. Cohen, Statistical Power Analysis for the Behavioral Sciences. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
[21] R. Kirk, "Zombies", in The Stanford Encyclopedia of Philosophy (Spring 2019 Edition), E.Zalta, Ed. 2019.
[22] N. Roughley and K. Bayertz, The Normative Animal? On the Anthropological Significance of Social, Moral, and Linguistic Norms. Oxford University Press, 2019. https://doi.org/10.1093/oso/9780190846466.001.0001
[23] J. Sidnell and N. Enfield, "Deixis and the interactional foundations of reference", in Oxford Handbook of Pragmatics, Y. Huang, Ed. Oxford University Press, 2019, pp. 217–239. https://doi.org/10.1093/oxfordhb/9780199697960.013.27
[24] J. Gee and J. Green, "Chapter 4: Discourse Analysis, Learning, and Social Practice: A Methodological Study", Review of Research in Education, vol. 23, no. 1, pp. 119-169, 1998. Available: https://doi.org/10.3102/0091732x023001119

7 Authors

David John Lemay is a postdoctoral fellow at McGill University, Montreal, Quebec.

Ram B. Basnet is an associate professor of Computer Science at Colorado Mesa University, Grand Junction, Colorado, USA.

Tenzin Doleck is a postdoctoral fellow at the University of Southern California, Los Angeles, California, USA.

Potentials of Chatbots for Spell Check among Youngsters

Jeton Arifi, Markus Ebner, Martin Ebner (*)
Graz University of Technology, Graz, Austria

ABSTRACT

Chatbots are already being used successfully in many areas. This publication deals with the development and programming of a chatbot prototype to support learning processes. This Chatbot prototype is designed to help pupils in order to correct their spelling mistakes by providing correction proposals to them. Especially orthographic spelling mistake should be recognized by the chatbot and should be replaced by correction suggestions stored in test data.

Keywords—Chatbot, spelling, misspelling correction, learning platform

1 Introduction

This publication is about developing and programming a Chatbot prototype for the IDeRBlog learning platform. IDeRBlog is an internet platform that includes a practice database of numerous free exercises licensed under a Creative Commons license. The user of the IDeRBlog learning platform are pupils in childhood who are learning the German language. The current version of the IDeRBlog learning platform helps chil dren to correct spelling errors in texts. The blog entries or essays that are written online by the students are checked for their spelling by an intelligent dictionary and afterwards suggestions for correction are offered. Based on these spelling errors, spelling corrections are suggested to the user.

Teachers can access this spelling error analysis for further use in lessons. The de veloped IDeRBlog Chatbot prototype is intended to provide better correction suggestions for mistakes, while allowing easy interaction with the users, which also encour ages them to learn and write. The interaction, input and operation of the chatbot should be very simple for the users. Pupils should also be given the feeling of being able to communicate with a supervisor or tutor who supports them in their learning processes in the best possible way. Therefore, an enhancement of the platform by a chatbot is being considered in order to offer even more motivation and an improved error analysis to the users. In this research work we address the research question: How can a chatbot for children be implemented to assist during their learning experiences in a language learning platform?

2 Introduction of Chatbot

The term "chatbot" consists of the English word "Chat" (interview) and "bot" (abbreviation of "robot"). Therefore, a chatbot is a computer system that allows humans to interact with the computer through natural language-based keyboard or voice recognition. A chatbot works always on repetitive patterns without human interven tion and simulates human beings. In doing so, chatbots access a stored knowledge base in which they select actions and answers by finding matches of asked questions with the existing questions or answers previously created by programmers and output them as answers to the questioner.

Chatbots are partially personified and are sometimes used in conjunction with an avatar (human, animal, or mythical creature). With the help of this "living organism" and its naturalness, the use of the computer should be entertaining and effortless [1], [2]. Likewise, chatbots can provide useful services, refer important information and present websites of interest to users in addition to directly answering the question. Thus, the users can find the information they are looking for, with little effort, or the conversation can be guided by the chatbot in a certain direction [3].

Chatbots also perform other tasks besides communication such as customer service in online shops, the management of forums, Internet auctions, information search on the World Wide Web or represent artificial characters in MUD's (Multiuser Dialog). Corporate websites are increasingly being complemented with a chatbot that acts as the virtual host, answering questions about products and services and navigating through the site [4]. The companies hope for stronger customer loyalty and better marketing information through call log analysis. There are also several noncommercial chatbots. These are developed primarily for the fun of experimenting, not least in the hope of creating a truly intelligent program. Every year, the developers of these non-commercial chatbots mainly meet to compete for the Loebner Prize, where the most human-like program is determined in a modified Turing test [5].

The British mathematician Alan M. Turing introduced his essay "Computing Ma chinery and Intelligence" published in the journal "Mind" in 1950 with the question "Can Machines Think?" [5], [6]. Therefore, the history of the chatbots can be determined back to 1950. In his essay, Turing described the basics of Artificial Intelligence(AI) and proposed to use the Turing test as the benchmark for measuring the intelligence of a computer system [6]. The question „Can Machines Think?" was finally replaced by the question of the communicative ability of a computer program by Tu ring. If a computer program succeeds in imitating a person's ability to communicate in natural language the machine should be considered as intelligent. The requirement for this is, that the computer succeeds in independent attempts and permanently deceives.Even today, 50 years after Turing, there is no agreement in AI research on the significance of the Turing Test, but it seems to serve the Chatbot developers as a relevant orientation. It is their goal to develop a system that passes the Turing Test. In the "conversation" the chatbot should no longer be distinguishable from humans and he must perfectly imitate a human conversation partner. In its original form of Turing, the Turing test was never realized. It was only towards the end of the eighties that a sufficiently powerful computer was developed that could seriously start practical experiments.

Since 1990, when the American Hugh Loebner launched the "Loebner Prize Com petition in Artificial Intelligence," a modified Turing test is conducted each year, with various chatbot programs. The first competition took place at the Boston Computer Museum in 1991, and consequently the Turing Test was also conducted for the first time [5]. Various prize money was defined for the competition. The first untraceable program in the Turing Test will be awarded a gold medal and a prize of $ 100,000. In order to win this prize, the program must be able to deal with audiovisual input (recognition of spoken language, facial expressions, gestures, emotions, etc.) and to convince the jury members [3].

Only bronze medals have been awarded till now, although the performance of the participating programs has improved significantly since 1991 [5]. Every year, several programs take part in the competition. Over the last few years, the ALICE project with Richard Wallace has repeatedly appeared and performed best in the years 2000, 2001 and 2004. The Turing test is just as controversial as the Loebner Prize. Loebner himself sees the promotion of AI research as the goal of his competition, but the majority of scientists are critical of him and his test [3]

2.1 How chatbots work

Today's chatbots are modular stimulus-response systems that compare the linguis tic input with an internal sample database and then output appropriate answers. The so-called knowledge database is also

today the heart of the chat bot, in which recognition patterns, keywords and answers are stored.
The conversation flow is controlled by the actual program and coordinates the in put and output, the activation of the knowledge database and possibly other models such as the output of spoken language. Commercial systems provide an editor that makes it easier for the user to set up and maintain the knowledge database. In the log function, which each Chatbot has, all dialogues are stored and are evaluated by the developers. The individual systems are different in size and flexibility of their "knowledge databases" as well as in the performance of their control programs.
Some chatbots only compare a part of an input to the pattern database, and other chatbots analyze all parts of the input that are combined with the answers. Most chat bots protocol their conversations and propose new keywords or recognition patterns based on them. However, the decision is made by humans whether they are included in the knowledge database or not. This method is called supervised learning.
Some systems independently expand their database based on the protocols. If there is no recognition pattern for the input, some chatbots can also use external databases [4]. The success of a chatbot system depends mainly on the quality of knowledge and the amount of knowledge in the knowledge base. Creating the knowledge base is a time-consuming job. However, by improving the algorithms of the system (for example, improved analysis of inputs), this time-consuming work can be reduced [3].

3 Architecture of Prototype

The IDeRBot-Chatbot is a self-programmed webapp based on a client-server archi tecture. For the development of the IDeR-ChatBot standard technologies HTML, CSS, JavaScript and the Open Source Framework Java Spring were used. The com munication between client and server is done by using the WebSocket protocol [7].
The messages exchanged between the client and the server are transmitted as JSON strings. JSON is a file format that can be used to transfer structured data between client and server and to generate content based on this data [8].

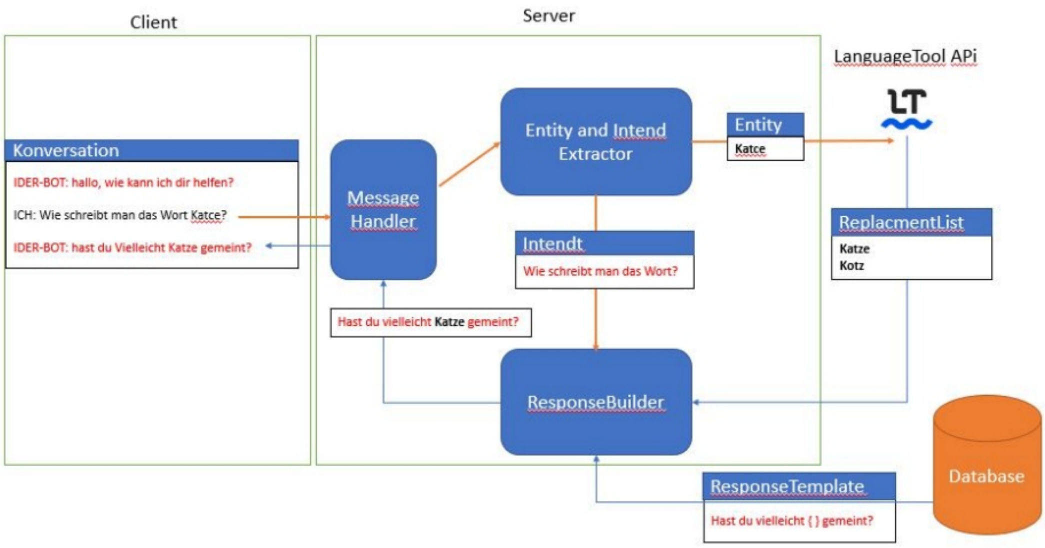

Fig. 1. Architecture of prototype with German dialog

3.1 Client

On the client side is the web frontend user interface, on which the users can write text and messages to the server and receive them from the server.

Figure 2 shows the input and output masks (communication interface) used for the message exchange between the chatbot and the users. The window left labeled as Text Editor displays the input field where the user can write text, but this doesn't mean that the text is correct. Before the user can send text for correction to the tutor or teacher, she/he has the possibility to check the text for misspelling by using Language ToolAPI, which can provide correction suggestions. In the right side there are windowsplaced through which user can communicate with chatbot and ask it for corrections if she/he is not sure that corrections suggestion made by Language Tool API are right

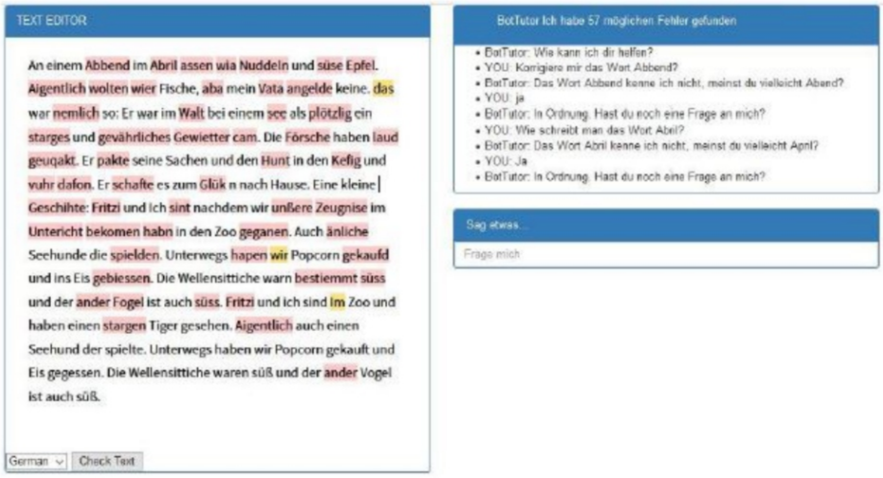

Fig. 2. Webfrontend communication interface for German language

3.2 Backend

To receive messages form users in backend a permanent socket connection must be available for all connected clients. For this implementation the WebSocket protocol of Java Spring Framework was used.1

From the received message first full the user's intention must be recognized, this was achieved by using the Levenshtein distance algorithm, the next step is to extract words that contains the spelling mistake from the message. To achieve this the Lan guage Tool API was used. The free open-source application Language Tool helps both non-native speaker and native speakers to recognize spelling and grammar mis takes and supports an appealing writing style [9]. In this Prototype those spelling mistakes are passed from the backend to the Language Tool API endpoint via REST API. The REST API returns a list of possible correction for each spelling mistake as a JSON String to the backend. This list includes, similar words or correction suggestions that can be associated with the spelling error.

The figure 3 shows all processes from receiving messages to generating responses in the backend.

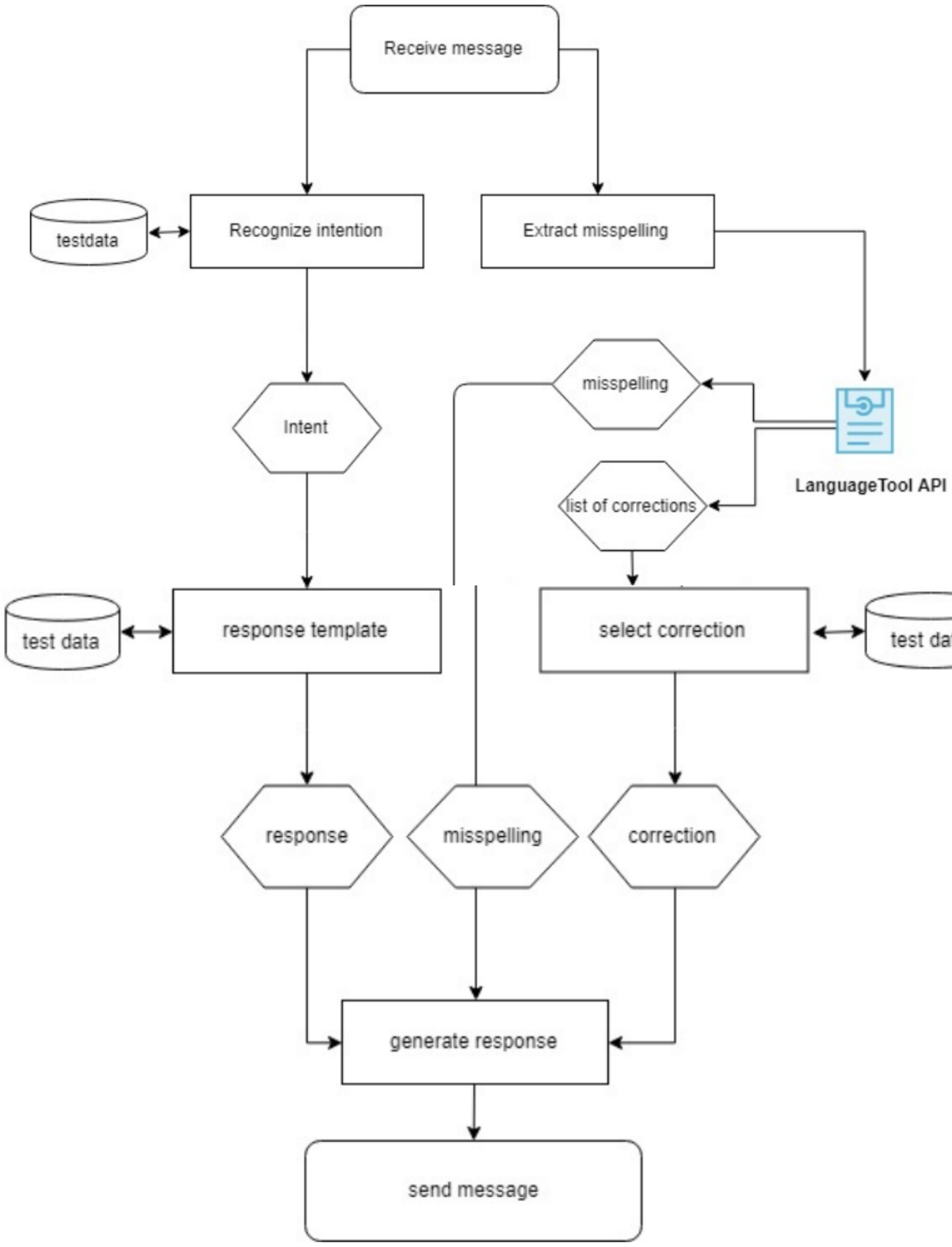

Fig. 3. Backend processes and message flow

3.3 Intent recognition

Recognizing a user intent is the most essential part of the IDeRBlog chatbot. The Levenshtein distance algorithm contained in the Java library (java-string-library) was used to recognize the intent. The Algorithm compares two strings and identifies com mon patterns when these strings do not match exactly. [10].

To detect user intent from message the distance () function takes the message as a parameter and runs through the list of user expressions from the test data (Train ingPhrases) and selects a TrainingPhrases with lowest distance.

3.4 Extracting misspelling

To extract a word(entity) from intent that contains a spelling error, the following procedure was followed. After the intent of the message has been detected by the Levenshtein distance algorithm, the message should be split into words. The individual words from the message that match the words from the intent would be ignored. The remaining words in the message that do not occur in the intent are candidates that may contain a spelling mistake, for which the user wants to ask the Bot for correction suggestions.

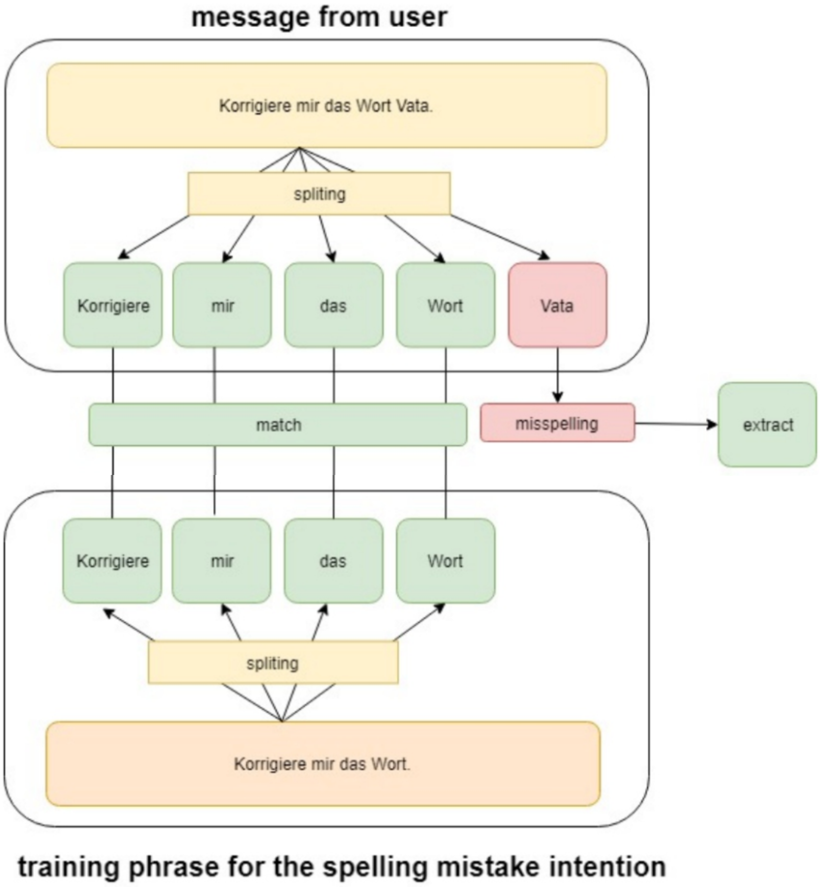

Fig. 4. Extraction of a spelling mistake

The extracted word must now be checked for spelling mistakes. This is done in the next step by using the Language Tool API.

3.5 Proposed corrections

In order to generate correction suggestions for the extracted word, it must first be checked for spelling mistakes. If a spelling error is found, the Language Tool API will also return a list of suggested corrections for the misspelled word. It is possible that wrong correction suggestions are given in the output list, especially for orthographic terms.

The algorithm must compare these correction suggestions with the orthographic terms from the test data and make the right suggestion to the user. To achieve this, the database was filled with test data wich contains some orthographic terms. Each term is associated with a correction suggestion by a frequency factor as shown in Figure 5.

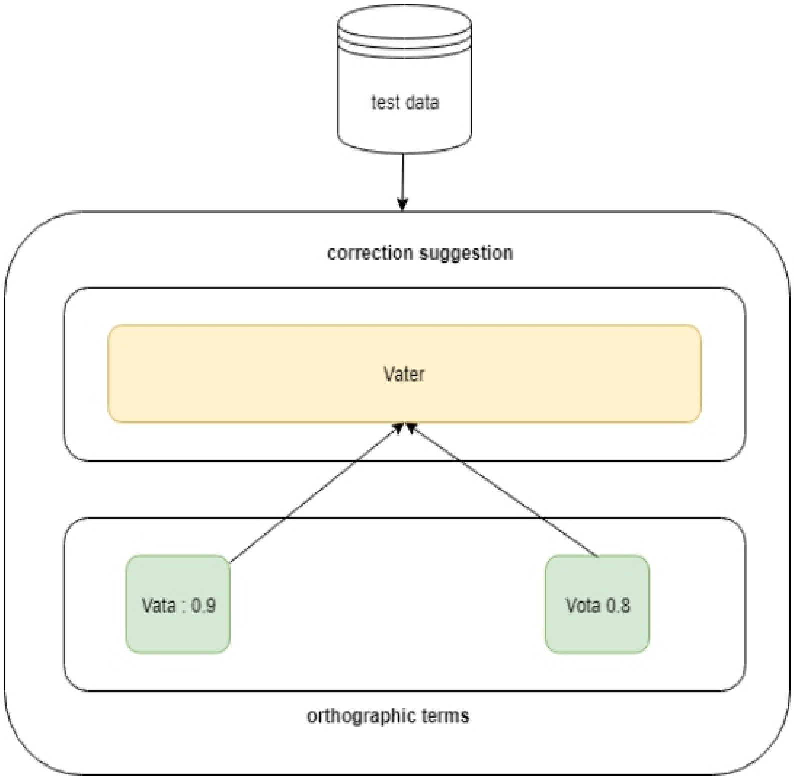

Fig. 5. Frequency factor

The suggested correction with the highest frequency factor is proposed to the user first. If the user accepts the suggested correction, the frequency factor for the suggested correction is recalculated and will be updated in database. This frequency factor ensures more precise suggested corrections for the upcoming proposals.

3.6 Generate message

After the intention of the user has been detected, a response template is automatically selected from the test data in the database. The response template contains the text message (response) and the necessary slots, which serve as placeholders for the spelling mistake and the correction suggestions. The number of slots for a specific response template must be predefined in order to generate an exact response for the user.

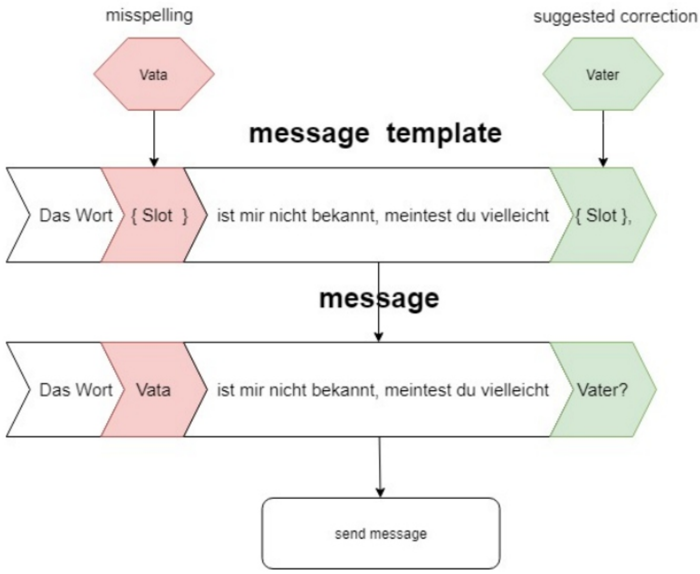

Fig. 6. Message generation using slots

4 Testing the Prototype

Finally, the prototype and the IDeRBlog chatbot were evaluated by experts. Their questions can be separated into three categories: A greeting, small talk, questions about spelling mistakes and questions about spelling mistakes/terms.

4.1 Test category 1: Greeting and small talk

Question to the IDeRBlog chatbot
"Hallo!" (Hello!)
"Wie heißt du?" (What's your name?)
"Wie geht es dir?" (How are you?)

All intentions from the messages were recognized and correct greetings for the test person were generated.

4.2 Test category 2: Questions about spelling mistakes

Question to the IDeRBlog chatbot
"Wie viele Fehler habe ich in meinem Text? "
(How many spelling mistakes do I have in my text?)
"Schreibt man Appril mit zwei p? " (Does Appril have a double p?)
"Habe ich Walt richtig geschrieben? " (Did I spell Walt right?)

All intentions from the messages were also recognized during the spelling error tests. The correct number of errors was detected, and all spelling mistakes were corrected correctly.

4.3 Test category 3: Questions on orthographic errors / terms

Question to the IDeRBlog Chatbot:	Response of IDeRBlog Chatbot:
„Korrigiere mir das Wort **Vata** (Correct me the word Vata?)	Hast du vielleicht **Data** gemeint? (Did you maybe mean Data?)

During the test run, the Language Tool API's correction proposal failed because **"Data"** is not the right correction proposal for the orthographic term or colloquial expression **“Vata”**. Correct it would be **“Vater”** which means **“father”** in German. These orthographic terms or colloquial expressions like **Vata** are used often by chil dren and teenagers when writing texts in schools, especially children who have learned the German language only by listening and do not have German as their na tive language. After we inserted this orthographic term into our test data, and put this in conjunction with the word father, the next test made the correct suggestion proposal to the user. Each time the user accepts the suggested correction, the frequency factor for this orthographic is increased. This ensures that the suggested corrections are becoming more precise. To conclude, it can be said, that the algorithm seems to workquite well but depends strongly on the amount of test data.

5 Discussion

After testing, the Chatbot IDeRBlog has performed well in detecting intentions and detection of spelling errors has been satisfactory. Correction statements for the selected spelling mistake have been correctly reported to the user. Due to the fact that the Chabot is just a prototype in development, there are still some things that can be done to make the Chatbot smarter:

• Feeding the database with more orthographic terms and linguistic variations, re sults in a hearing recognition rate
• To increase accuracy, the Chatbot IDeRBlog should also be used frequently.

In summary, it should be noted that the correct generation of responses by a chatbot can is based on the amount of intents that chatbot can detect in a message and the learning ability of the algorithms. At the moment there are no algorithms that can detect intentions without having some training data.It also turned out that the error detection rate was heavily dependent on the existing test data. A further optimization of the IDeRBlog chatbot would be the error detection of grammatical errors.

6 Conclusion

The aim of this research was to develop a chatbot for the IDeRBot learning platform of the TU Graz and to integrate it. The chatbot should act as a tutor for the spell checker and communicate with the pupils. Through the use of the chatbot, the users should receive an added value in learning the German language. In addition, the inter action should encourage users to write more texts in the German language.
First, an overview of chatbots in general was given by using the literature. It was used to define the term chatbot and to identify tests to determine quality standards. The historical development and functioning of chatbots was also discussed. Areas of application in which chatbots have already been successfully used were also shown. The IDeRBlog chatbots were then discussed by describing the technologies, development environments and libraries used.
The development of this prototype has shown that chatbots can also relieve teach ers or tutors in the classroom. These positive experiences encourage us to continue in the research and development in this

area. In conclusion it can be sad that, chatbots are quite capable of performing such tasks, but they will never be able to replace a teacher or a tutor completely. Because chatbots are just algorithms and machines do not have the emotional side.

7 Acknowledgements

We gratefully thank the EU-Erasmus+-Project IDeRBlog-ii for financial support and our project partners for the cooperation. The IDeRBlog-ii project are funded by the European Commission in the framework of Erasmus+. IDeRBlog-ii: VG-IN-SL18-36-047317, 2018-2021.

8 References

[1] Braun, Alexander (2003): Chatbots in der Kundenkommunikation. Springer-Verlag Berlin Heidelberg.

[2] Willmes, Anna (2014): Chatbots: Was ist das und wie funktionieren sie? In: https://famiville.wordpress.com/2014/11/24/chatbots-definition-funktionsweise-einsatzgebiete-starken-schwachen/#more-139.

[3] Möbus, Claus; Eißner, Andreas; Feindt, Jan; Janser, Claudia; Krefeldt, Jens; Sieverding, Sven; Sölbrandt, Stefan; Stumpe, Jörg; de Vries, Holger; Willer, Stefan (2006): Web Kommunikation mit OpenSource. Chatbots, Virtuelle Messen, Rich-Media-Content. Springer-Verlag Berlin Heidelberg. https://doi.org/10.1007/3-540-29093-1

[4] Berger, R., Ebner, M. & Ebner, M. (2019) Conception of a Conversational Interface to Provide a Guided Search of Study Related Data. Internationale Journal of Emerging Technologies (i-JET). 14(7). pp. 37-47. https://doi.org/10.3991/ijet.v14i07.10137

[5] Storp, Michaela (2002): Chatbots. Möglichkeiten und Grenzen der maschinellen Verarbei tung natürlicher Sprache. In: https://www.mediensprache.net/networx/networx-25.pdf.

[6] Deshpande, Aditya; Shahane, Alisha; Gadre, Darshana; Deshpande Mrunmayi; Prod. Dr. Joshi, Prachi M (2017): A Survey of Various Chatbot Implementation Techniques. Interna tional Journal of Computer Engineering and Applications, Volume XI. In: https://pdfs.semanticscholar.org/8e60/5c49d4a7cba9bf077d97b401ba78aafe693f.pdf.

[7] Weßendorf Matthias (2011): WebSocket: Annäherung an Echtzeit im Web. In: https://www.heise.de/developer/artikel/WebSocket-Annaeherung-an-Echtzeit-im-Web 1260189.html?seite=all (besucht am 25.03.2019).

[8] Ackermann Philip (2018): JavaScript. Das umfassende Handbuch. 2. aktualisierte und erweiterte Auflage. Bonn: Rheinwerk Verlag.

[9] Language Tooler GmbH (2019): Funktionen und Hilfe. In: https://LanguageTool.org/de/(besucht am 25.03.2019).

[10] Lisbach Bertrand (2011): Linguistisches Identity Matching. Paradigmenwechsel in der Suche und im Abgleich von Personendaten. 1. Auflage. Germany Wiesbaden GmbH. https://doi.org/10.1007/978-3-8348-9791-6_1

9 Authors

Jeton Arifi is Junior Researcher at the Institute of Interactive Systems and Data Science at Graz University of Technology, Graz, Austria. (e-mail: jeton.arifi@tugraz.at)

Markus Ebner is Researcher at the Department Educational Technology at Graz University of

Graz, Austria. (e-mail: markus.ebner@tugraz.at)

Martin Ebner is head of the Department Educational Technology at Graz University of Technology, Graz, Austria. (e-mail: martin.ebner@tugraz.at). He is responsible for all university wide e-learning activities. He holds an Adjunct Prof. on media informatics (research area: educational technology) and works also at the Institute of Interactive Systems and Data Science as senior researcher. For publications as well as further research activities, please visit: http://martinebner.at

The Ecology of Analytics in Education Stakeholder Interests in Data-Rich Educational Systems

Ulf Dalvad Berthelsen, Morten Tannert
Danish School of Education, Aarhus Universitya.

A B S T R A C T

In this paper, we suggest an ecological perspective on the role of analytics in education. We discuss how different stakeholder positions in educa tion points to different interests in using analytics. As a point of reference, we examine the Danish case of ICT integration in primary and lower secondary school (Danish: Folkeskolen) in order to study cases of emerging and at times conflicting stakeholder interests. On this basis, we discuss how this complexity of the educational ecosystem affects different stakeholder positions within the field.

Keywords—Learning analytics, stakeholders, educational ecosystems, techno logical ecosystems, digital ecosystems

1 Introduction

In recent years, large investments in digital educational infrastructure, and digital teaching and learning environments has led to an extensive interest in learning analyt ics from both public and private stakeholders. This digital transformation of education has evoked an unprecedented level of datafication of learning and teaching with large amounts of data collected on all levels of the educational system [1], [2]. This data is collected from a wide variety of systems and applications and includes information on everything from attendance to test results, student interaction, digitized student products and now even biometric data. To an increasing extent, data on behavior and per formance generated in digital teaching and learning environments has become the filter through which learning and teaching is monitored and evaluated. These techno logical developments challenge traditional views of learning and teaching as well as our understanding of the role of analytics in education.

According to Ferguson [3], there are several drivers behind the growing interest in applying data analytical methods in education. These drivers are motivated by a range of political, educational, academic, and economical factors. Thus, the field of analytics in education has become a complex field of different stakeholder positions and interests. New partnerships between public and private actors emerge and challenge the traditional power balance in public education [4], [5]. This is, among other things, because policy reforms on digitalization rely on business actors to help schools fulfill the demands [6]. These new dynamics open for a much more complex conception of the role of analytics in education, because new emerging stakeholders hold different and at times competing or conflicting interests.

In this paper, we propose an ecological approach to the conception of analytics in education. An ecological approach allows us to describe the complexity and interre latedness of the different levels of the educational ecosystem. Our aim is to show that learning analytics is only one out of several different interests in educational data, and we suggest purpose and success criteria as important standards for distinguishing between different types of stakeholder interests in educational data. As a point of reference, we examine the Danish case of ICT integration in primary and lower sec ondary school (Danish: Folkeskolen) in order to study cases of emerging and at times conflicting stakeholder interests.

On this basis, we discuss how the complexity of the educational ecosystem affects different stakeholder positions within the field.

2 Background

Over the last decade, the field of learning analytics has emerged as a subdisciplinary field of educational research [7], [8], that focuses on utilizing computational techniques and tools for analyzing large amounts of data generated in educational contexts for purposes of understanding and optimizing learning and learning envi ronments [9], [10], [11], [12]. Early learning analytics tended to focus on analyses that could inform strategies at an institutional level [13], e.g. by predicting overall aca demic performance or student retention. However, with the continuous increase of available digital data generated by individual students, learning analytics has moved the analytic scope towards the level of the individual learner [14], [15], in order to examine learning as it occurs in the context of formal education. This shift in focus is, among other things, caused by the emergence of e.g. digital learning platforms and learning management systems, online learning environments, digital learning resources, student programming and multimodal text production, that all leave behind digital traces containing rich information about both learning activities and learner behavior [16], [17], [18], [19].

The increased interest in data on individual learning progression has not meant, however, a decrease in the interest in other types of data generated in educational contexts. On the contrary, what we currently see is an overall increase in the generation and collection of data for multiple purposes in all parts of the educational system.Obviously, doing learning analytics, i.e. collecting and analyzing data for optimizing learning, is still a major interest. It is equally obvious, however, that the accelerating datafication of education presents an opportunity to do analytics for various other purposes than optimizing learning.

3 The Ecosystem of Education

3.1 Social contexts as ecologies

In order to describe an understand the increased complexity of educational analyt ics, we employ an ecological perspective to explore how different stakeholders in various parts of the ecosystem pursue different interests. In doing so, we follow a growing trend in recent years of framing social context and human interaction in terms of ecology. Pioneered by, among others, Gibson [20], [21] and Bronfenbrenner [22], notions such as ecosystem, environment, diversity and change have entered the terminology of various disciplines such as psychology, anthropology, and science and technology studies. Even though Gibson focused almost entirely on human interaction with the physical environment, his ecological account has been very influential in understanding how the surrounding environment shapes and enables human action and interaction. Bronfenbrenner focused more explicitly on human interaction in so cial contexts as he developed his ecological systems theory in order to account for human development. This account too has had great impact and inspired research with an ecological perspective in many fields.

Being conceived in the 70's, neither of the two approaches pay much attention to the impact of (digital) technologies on social ecosystems. Later, an important attempt to include technologies is made by Nardi & O'Day [23]. In order to describe technol ogy-rich social contexts, they introduce the notion of information ecology, which they define as an ecosystem consisting of people, practices, values, and technologies. Thus, in their account, technologies are put on a par with actors and their values and practices. For the purpose of describing the complexity of contemporary data-rich educational systems, we need, however, to be able to look at the interconnected technologies as an independent system. Therefore, following García-Peñalvo [24], we understand the ecosystem of educational technologies (in the broadest sense) as an independent part of the larger ecosystem of education. This is important because digital technologies, as opposed to non-digital technologies such as books or pencils, can disperse data and information across platforms and interact in certain ways that create synergy and cross-sectional effects.

3.2 Analytics in educational ecosystems

On these grounds, we will work with three levels of description

- The ecosystem of institutions
- The ecosystem of technologies
- The analytics practices in technology-rich educational ecosystems.

By ecosystem of institutions, we mean the entire range of public and private institutions and companies, including all the different groups of people that make op the institutions, e.g. politicians, policy makers, leaders, practitioners, parents, and students. In other words, the ecosystem of institutions is the sociocultural level of the educational ecosystem, i.e. the level of action, interaction, decision making etc.

By ecosystem of technologies, we mean the interconnected system of technologies (both hardware and software) that make up the entire cross-sectional network of ad ministrative systems, communication systems, test systems, learning management systems, digital learning resources, digital student products etc. It is important to no tice the asymmetry between the ecosystem of institutions and that of technologies:

The ecosystem of institutions can exist without the ecosystem of technology, but not vice versa.

By analytics practices in technology-rich educational ecosystems, we mean the use of data generated within the ecosystem of technologies carried out by actors within the ecosystem of institutions with the purpose of doing analytics. By this definition, we mean to emphasize the dynamic relationship and mutual dependency of the three levels: Institutions implement new technologies > the ecosystem of technologies gen erates new types of data > new types of data enable new analytics practices > new analytics practices affect the behavior of actors in the ecosystem of institutions etc.

4 Ecosystems of Education: The Danish Case

Since the early 2000's the Danish public sector has been subject to a range of dif ferent strategies for digitalization [25], [26], [27], [28], [29]. The earliest strategies only targeted the educational sector moderately. However, in 2011, the Danish Government and KL - Local Government Denmark (interest organization of the 98 Danish municipalities) made a deal to set aside 500 million DKK (approx. 67 million EUR) to strengthen the use of ICT in primary and lower secondary school [30]. The funds were in particular allocated directly to schools to allow them to purchase digital learning resources and provide all students with well-functioning ICT. In addition, the fundswere to strengthen the development of digital learning resources in the private sector and further strengthen research and consultant-based projects focusing on applying ICT in teaching. As a result, the Danish primary and lower secondary school has seen an increase in the use of digital learning resources [31] and a general digitization of the practices of different agents from both within and outside the educational domain. In the following, we will take a closer look on how these changes have accommodated new stakeholder positions and interests within the Danish educational ecosystem.

4.1 Ecosystem of institutions

A key part of the 2011 digitalization strategy was to provide public funding for pri vate producers of learning resources. The idea was to aid the development of a range of digital learning resources that could utilize the affordances of the digital format to facilitate innovative and varied teaching practices [32]. The most recent outline of funded projects shows that many small private actors were provided with funding [33]. However, studies have shown that the Danish market for digital learning resources is largely dominated by big private actors who offer omnipotent online learning portals with traditional ready-to-use courses [34], [35]. For instance, Gyldendal (the biggest publishing house in Denmark) has developed an online learning platform that covers all school subjects and content areas, and it is the most frequently used digital learning resource for the teaching of Danish as L1 (mother tongue education)[31].

From a macro perspective, there is an apparent discrepancy in interests between government intentions of using small private actors as an initiator of innovative teaching on the one hand and the private market allowing big actors with business interests to dominate with more traditional courses and content on the other. This supports the idea put forward by Tomlinson [36] that in economic markets consumer sovereignty soon gives way to producer sovereignty, which allows the stakeholder interests of big private actors to dominate at the expense of national policies and intentions. This type of digital transformation – whether it is in the public or the private sector – is usually based on an economic logic that equals digitalization with either financial profit, savings or increased systemic efficiency [37], [38], [39], [40].

From a micro perspective, different stakeholder positions are emerging. Parents and teachers enter the equation as agents interested in the data generated by new digital technologies in education. However, the interests of teachers and parents may differ, as teachers tend to use performance and achievement data in a functional man ner to inform instructional strategies [41], whereas parents might have more implicit interests (e.g. choosing schools for their children). Further, school leaders might use data from technologies in the classroom to allocate financial means or teacher re sources to specific groups of students. This is in line with the recent trend of datadriven or data-informed school management, which is becoming increasingly popular within the Danish educational system [42], [43]. A somewhat recent Danish study shows that both school leaders and parents find quantitative data more legitimate than qualitative data, whereas teachers prefer to use qualitative data to inform classroom practice. However, at school level quantitative data are more often used as the basis of local decisions, because quantitative data is easier to collect and organize systemati cally [44]. Although these stakeholder interests are not necessarily conflicting, they pose different perspectives and add different uses to the data generated by digital technologies in schools.

4.2 Ecosystem of technologies

The changes described in the above section have direct implications for the ecosystem of technologies. Although Denmark in general is a highly digitalized country [45], the latest digitalization strategies have further increased the number of digital technologies in the classroom. In conjunction, these technologies form an ecosystem capable of generating data about several aspects of learning and teaching environ ments, including data on individual students' learning outcome. However, different technologies in the classroom accommodate new stakeholder positions. For instance, a growing number of Danish so-called iPad-schools have occurred across the country investing in tablets based on a 'one device, one student' strategy [46]. Such strategies open for further intersections between public education and private market actors, who have learning and financial profiting as their respective interests.

Specific hardware also affects the kinds of software that are implemented in educa tion. An example from the Danish case is the implementation of learning management systems, which is part of a larger 2014 collaboration between government, municipalities and private contractors (Danish: Bruger portal sinitiativet). Because all students have been equipped with either laptops or tablets, the idea is to further expand the digital infrastructure of Danish primary and upper secondary schools by providing teachers with online platforms that can help facilitate learning activities in the classroom. One of the features of the project is that data on student learning progression and well-being can be gathered and shared with teachers, parents, school leaders and governing institutions [47]. This allows the municipalities and the state to track pro gress on both an institutional (schools) and an individual (students) level, which po tentially allows data to be used for accountability or benchmarking purposes.

4.3 Analytics practices in technology-rich educational ecosystems

From the descriptions above, it becomes clear that the digitalizing of the Danish primary and lower secondary school has created a large and diverse ecosystem of technologies capable of producing large amounts of data. It is equally clear that this development has opened for new stakeholders with other interests in educational data than optimizing learning outcomes. As the educational ecosystem evolves, new ana lytics practices might emerge. Right now, however, we see three main types of analyt ics practices within the educational ecosystem.The first type of practice is institutional analytics. From a macro perspective insti tutional analytics focuses on accountability, benchmarking and institutional governance, and it is still a very common type of analytics despite the increased focus on the individual learner. From a micro perspective, this particular interest also manifests locally when school leaders use educational data in the interest of allocating financial means and teacher resources to specific areas of the school. Thus, we see institutional analytics as a dominant analytics practice with institutional governance as its purpose and with the distinct success criteria of increasing operational efficiency.

The second kind of practice we see is business analytics. If we turn to industry, we see a range of private actors interested in using the data created within the ecosystem of technologies for competing in the economic market more effectively [48]. Particu larly, the developers of digital learning resources are important actors because they dominate the teaching of specific school subjects and hence have large amounts of data from the classrooms available to them. These types of data can be utilized in the process of designing or redesigning products with the aim of aligning business planning and strategy with data-informed knowledge of consumer (i.e. students and teachers) behavior. Thus, business analytics has the purpose of consolidating and increas ing specific market positions and the distinct success criteria of ensuring revenue, and we see it as another dominant analytics practice within the Danish educational ecosystem.

The third kind of analytics we see is learning analytics. Learning analytics has the purpose of understanding and optimizing learning and learning environments and the distinct success criteria of improving students' learning outcome. As such, learning analytics is concerned with a wider variety of methodological and ethical issues, than is the case in the other two analytics practices. In addition, questions of data owner ship and the difficulties of operationalizing valid measures of learning also makes the practice of learning analytics the most complex [44], which might explain why learn ing analytics currently seems to be the least dominant analytics practice in the Danish context. We see very few examples of full-fledged implementations of learning ana lytics tools, and they are often met with skepticism from practitioners.

5 Discussion

According to Reyes [49], learning analytics – as opposed to other types of analytics – has the potential to inform and benefit most stakeholders in education, because it examines the one unit that most stakeholders share, namely the learning of students. However, as we have pointed out there are many other types of analytics in education. Learning analytics is, as such, not a replacement for other kinds of analytics, but rather it is a further addition to an already complex educational ecology. In the following, we will briefly address how this complexity affects different stakeholder positions within the field.

5.1 Implications for policy makers

At the levels of policy and governance, there is a need for a more reflected approach to analytics based on deep knowledge of the potentials and limitations of data analytical methods in educational contexts. This implies that policy makers take into account the dynamics of the complex educational ecology and carefully consider how bringing in new technologies and new stakeholders affect the different ecosystems and the ecology as a whole. This means that knowledge generated through analytics is not blindly integrated into educational policies as a replacement for teacher expertise. Rather, such knowledge should be a supplement to teachers' pedagogical thinking and professional judgement [50].

5.2 Implications for schools and private stakeholders

As data magnitude and stakeholder numbers increase, so does the opportunities for data breaches. In addition, as more stakeholders from the private sector contribute to the ecosystem of technologies, the question of data ownership becomes more apparent [51]. This only adds to the high requirements of security and transparency that are already assigned to the field. Schools have a major challenge coping with this com plexity. As pointed out by Rao, Ding & Gudivada [52], successful analytics require a highly sophisticated reference architecture consisting of many different human and technical resources. As such, neither schools nor private stakeholders are left alone to cope with the many challenges of analytics, but are required to collaborate and find common solutions and workflows.

5.3 Implications for teachers

Teachers might come under heavy pressure when different kinds of analytics enter the educational ecology. Thus, data literacy skills need to become an integral part of teacher training and education [53]. Data literacy is, in this sense, not only a matter of being able to understand and interpret data, but it also includes the ability to act upon data in a critical and constructive manner. A part of this is being aware of how differ ent stakeholder positions within the ecology implies specific interests in educational data. Thus, teachers must be able to interact with data in a way that balances datainformed decision-making, and pedagogical freedom and reasoning.

5.4 Implications for students and parents

It is arguably an open question whether students can be seen as stakeholders in analytics, especially if the students are children. Students are, however, one of the main contributors and creators of data, as it is their activities and behavior that generate a big part of the data used in analytics. Still, new evaluation formats (especially in formal education) require students to be able to understand how the resulting data are interpreted. Data literacy is, thus, a key competence for students at all levels of the educational system. Similarly, parents are faced with many types of data about their children – particularly through primary and lower secondary school (assessment data, data on well-being, performance data etc.). Hence, data literacy skills are also important for parents in order for them to gain agency.

6 Conclusion

From the description of the Danish case, we have seen that stakeholder positions in education point to different and specific interests in doing analytics. Thus, it is an important point that not all types of analytics in education are learning analytics, i.e. analytics with the purpose of understanding and optimizing student learning and learning environments. Analytics concerned with institutional governance, bench marking, accountability, business profits etc. are all relevant and for the most part legitimate positions seeking to take advantage of the digital traces left behind by dif ferent actors in education. These different positions do not necessarily pose a prob lem. However, when stakeholder interests are conflicting, we might see counterproductive initiatives where the use of different types of analytics are working in opposite directions. As such, taking into account the different stakeholder positions and interests in analytics is crucial for further development and manifestation of the field of educational analytics.

7 References

[1] Klašnja-Milićević, A., Ivanović, M. & Budimac, Z. (2017). Data science in education: Big data and learning analytics. Computer Applications in Engineering Education, 25(6), 1066-1078. https://doi.org/10.1002/cae.21844

[2] Zhang, W. & Jiang, L. (2018). Algorithm Analysis for Big Data in Education Based on Depth Learning. Wireless Personal Communications, 102(4), 3111-3119. https://doi.org/10.1007/s11277-018-5331-3

[3] Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. Interna tional Journal of Technology Enhanced Learning, 4(5/6), 304-317. http://dx.doi.org/doi:10.1504/IJTEL.2012.051816

[4] Ball, S. (2009). Privatising education, privatising education policy, privatising educational research: network governance and the 'competition state'. Journal of Education Policy, 24(1), 83-99. https://doi.org/10.1080/02680930802419474

[5] Selwyn, N. (2018) Data points: exploring data-driven reforms of education. British Journal of Sociology of Education, 39(5), 733-741. https://doi.org/10.1080/01425692.2018.1469255

[6] Williamson, B., Rensfeldt, A. B., Player-Koro & Selwyn, N. (2018). Education recoded: policy mobilities in the international 'learning to code' agenda. Journal of Education Poli cy. https://doi.org/10.1080/02680939.2018.1476735

[7] Lodge, J. M., Horvath, J. C. & Corrin, L. (2019). Introduction: Learning analytics in the classroom. In. Lodge, J. M., Horvath, J. C. & Corrin, L. (Eds.). Learning Analytics in the Classroom. Translating Learning Analytics Research for Teachers. New York: Routledge.https://doi.org/10.4324/9781351113038

[8] Reimann, P., Bull, S., Kickmeier-Rust, M. D., Vatrapu, R., & Wasson, B. (Eds.). (2016). Measuring and Visualizing Learning in the Information-Rich Classroom. New York & London: Routledge.

[9] Romero, C., & Ventura, S. (2010). Educational data mining: A review of the state of the art. IEEE Transactions on Systems, Man and Cybernetics, 40 (6), 601-618. DOI: 10.1109/TSMCC.2010.2053532

[10] Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. American Behav ioral Scientist, 57 (10), 1380-1400. https://doi.org/10.1177/0002764213498851

[11] Chatti, M. A., Dyckho, A. L., Schroeder, U., & Thues, H. (2013). A reference model for learning analytics. International Journal of Technology Enhanced Learning, 4(5-6), 318-331. https://doi.org/10.1504/IJTEL.2012.051815

[12] Siemens, G., & Gasevic, D. (2012). Guest editorial - learning and knowledge analytics. Educational Technology & Society, 15(3), 1-2.

[13] Mirriahi, N., Dawson, S., Gasevic, D. & Long, P. (2014). Widening the Field and Sparks of the Future. Journal of Learning Analytics, 1(3), 1-3. https://doi.org/10.18608/jla.2014.13.1

[14] Tabak, I. & Radinsky, J. (2014). Knowing the Learner: Multiple Lenses on Diagnostic Capacities, Failures, and Uncertainties, in Support of Learning. Journal of the Learning Sciences, 23(4), 487-489. https://doi.org/10.1080/10508406.2014.972796

[15] Papamitsiou, Z. & Economides, A. (2014). Learning Analytics and Educational Data Min ing in Practice: A Systematic Literature Review of Empirical Evidence. Educational Technology and Society, 17(4), 49-64.

[16] Bertland, M., Martin, T., Benton, T., Smith, C. P. & Davis, D. (2013). Using Learning Analytics to

Importance and impact of 5G network is critically discussed in this study. Several types of challenges and mitigation process of those challenges are critically evaluated here. For this reason, this study should be beneficial for each and every company in future.

REFERENCES

detection*[1] Rahman, Sawsan Abdul, et al. "Internet of things intrusion detection: Centralized, on-device, or federated learning?." IEEE Network 34.6 (2020): 310-317.*

[2] Gupta, Rajesh, et al. "Blockchain-envisioned multi-swarming UAVs to tackle COVID-I9 situations." IEEE Network 35.2 (2020): 160-167.

[3] Zhuang, Weihua, et al. "SDN/NFV-empowered future IoV with enhanced communication, computing, and caching." Proceedings of the IEEE 108.2 (2019): 274-291.

[4] Oliver, Kathryn, Anita Kothari, and Nicholas Mays. "The dark side of coproduction: do the costs outweigh the benefits for health research?." Health research policy and systems 17.1 (2019): 1-10.

[5] Berndt, Andrea E. "Sampling methods." Journal of Human Lactation 36.2 (2020): 224-226.

[6] Campbell, Steve, et al. "Purposive sampling: complex or simple? Research case examples." Journal of research in Nursing 25.8 (2020): 652-661.

[7] Henry, L.A., Buyl, T. and Jansen, R.J., 2019. Leading corporate sustainability: T he role of top management team composition for triple bottom line performance. Business Strategy and the Environment, 28(1), pp.173-184.

[8] Slamnik-Kriještorac, N., Kremo, H., Ruffini, M. and Marquez-Barja, J.M., 2020. Sharing distributed and heterogeneous resources toward end-to-end 5G networks: A comprehensive survey and a taxonomy. IEEE Communications Surveys & Tutorials, 22(3), pp.1592-1628.

[9] Products. Everything you need to know about 5G. Qualcomm. (2020) Available: https://www.qualcomm.com/5g/what-is-5g#:~:text=5G%20wireless%20technology%20is%20meant,experiences%20and%20connects%20new%20industries. Accessed 11th January, 2023.

[10] Husain, Syed, Andreas Kunz, and JaeSeung Song. "3GPP 5G Core Network: An Overview and Future Directions." Journal of information and communication convergence engineering 20.1 (2022): 8-15.

[11] Supriyanto, Agus, et al. "Teacher professional quality: Counselling services with technology in Pandemic Covid-19." Counsellia: Jurnal Bimbingan dan Konseling 10.2 (2020): 176-189.

[12] Patzold, Matthias. "5G Sets sail around the globe [Mobile radio]." IEEE Vehicular Technology Magazine 14.2 (2019): 5-11.

[13] Oughton, Edward J., et al. "Revisiting wireless internet connectivity: 5G vs Wi-Fi 6." Telecommunications Policy 45.5 (2021): 102127.

[14] Home. The Impact of 5G on the Internet of Things. Davra. (2023). Available: https://davra.com/5g-internet-of-things/. Accessed 11th January, 2023.

[15] Lööw, Joel, Lena Abrahamsson, and Jan Johansson. "Mining 4.0—The impact of new technology from a work perspective." Mining, Metallurgy & Exploration 36.4 (2019): 701-707.

[16] Hanaysha, J. R., and H. M. Alzoubi. "The effect of digital supply chain on organizational performance: An empirical study in Malaysia manufacturing industry." Uncertain Supply Chain Management 10.2 (2022): 495-510.

[17] Zhang, Jun, Xianghao Yu, and Khaled B. Letaief. "Hybrid beamforming for 5G and beyond millimeter-wave systems: A holistic view." IEEE Open Journal of the Communications Society 1 (2019): 77-91.

[18] Wang, Y., Paquette, L. & Baker, R. (2014). A Longitudinal Study on Learner Career Advancements in MOOCs. Journal of Learning Analytics, 1(3), 203-206. https://doi.org/10.18608/jla.2014.13.23
[19] Martin, T. & Sherin, B. (2013). Learning Analytics and Computational Techniques for Detecting and Evaluating Patterns in Learning: An Introduction to the Special Issue. Journal of the Learning Sciences, 22(4), 511-520. https://doi.org/10.1080/10508406.2013.840466
[20] Gibson, J.J. (1977). The Theory of Affordances (pp. 67–82). In R. Shaw & J. Bransford (eds.). Perceiving, Acting, and Knowing: Toward an Ecological Psychology. Hillsdale, NJ: Lawrence Erlbaum.
[21] Gibson, J. J. (1979). The Ecological Approach to Visual Perception. Boston: Houghton Mifflin.
[22] Bronfenbrenner, U. (1979). The Ecology of Human Development: Experiments by Nature and Design. Cambridge, MA: Harvard University Press.
[23] Nardi, B. A. & O'Day, V. (1999). Information ecologies: using technology with heart. Cambridge: MIT Press.
[24] García-Peñalvo, F. J. (2016). Technological Ecosystems. IEEE Revista Iberoamericana de Technologicas del Aprendizaje, 11(1), 31-32. https://doi.org/10.1109/rita.2016.2518458
[25] Projekt Digital Forvaltning (2002). På vej mod den digitale forvaltning – Vision og strategi for den offentlige sektor [Towards digital management – Vision and strategy for the public sector]. Copenhagen: Den Digitale Taskforce. Retrieved from https://digst.dk/media/12700/digitaliseringsstrategi-2001-2004.pdf
[26] Regeringen, KL, Amtsrådsforeningen, Københavns Kommune and Frederiksberg Kom mune (2004). Den offentlige sektors strategi for digital forvaltning 2004-06 – realisering af potentialet [The public sector's strategy for digital management 2004-06 – realizing the potential]. Copenhagen: Den Digitale Taskforce. Retrieved from https://digst.dk/media/12702/digitaliseringsstrategi-2004-2006.pdf
[27] Regeringen, KL and Danske Regioner (2007). Strategi for digitalisering af den offentlige sektor 2007-2010. Mod bedre digital service, øget effektivisering og stærkere samarbejde [Strategy for the digitization of the public sector 2007-2010 – Towards better digital ser vice, increased efficiency and stronger collaboration]. Copenhagen: Den Digitale Taskfor ce. Retrieved from https://digst.dk/media/12701/digitaliseringsstrategi-2007-2010.pdf
[28] Regeringen, KL and Danske Regioner (2011). Den digitale vej til fremtidens velfærd. Den fællesoffentlige digitaliseringsstrategi 2011-2015 [The digital path to future welfare. The common public strategy for digitization 2011-2015]. Copenhagen: Økonomistyrelsen. Re trieved from https://digst.dk/media/12704/digitale_vej_til_fremtidens_velfaerd.pdf.https://doi.org/10.6027/anp2013-746
[29] Regeringen, KL and Danske Regioner (2016). Et stærkere og mere trygt digitalt samfund. Den fællesoffentlige digitaliseringsstrategi 2016-2020 [A stronger and more secure digital society. The common public strategy for digitization 2016-2020]. Copenhagen: Digitaliseringsstyrelsen. Retrieved from https://digst.dk/media/12811/strategi-2016-2020-enkelttilgaengelig.pdf
[30] The Danish Government, Danish Regions and Local Government Denmark (2011). The digital path to future welfare. eGovernment strategy 2011-2015. Copenhagen: The Danish Agency for Digitization. Retrieved from https://digst.dk/media/12703/tilgaengelig_engelsk_strategi.pdf
[31] Bundsgaard, J., Buch, B. & Fougt, S. S. (2017). De anvendte læremidlers danskfag belyst kvantitativt [The L1 subject of the learning resources from a quantitative perspective]. In. Bremholm, J., Bundsgaard, J., Fougt, S. S. & Skyggebjerg, A. K. (Eds.). Læremidlernes danskfag (pp. 28-54) [The L1 subject of the learning resources]. Aarhus: Aarhus University Press.
[32] Ministeriet for Børn, Undervisning og Ligestilling (2016). Udmøntningsmodel for udviklingspulje

til digitale læremidler [Facilitation model for a development pool for digital learning resources]. Copenhagen: Sekretariatet for IT i Folkeskolen. Retrieved from https://www.uvm.dk/puljer-udbud-og-prisuddelinger/puljer/puljeoversigt/tidligere-udmeldte-puljer/grundskole/udviklingspulje-til-digitale-laeremidler https://doi.org/10.7146/lt.v1i1.107618

[33] Undervisningsministeriet (2017). Oversigt over støttede projekter – udviklingspulje til di gitale læremidler [Outline of supported projects – development pool for digital learning re sources]. Copenhagen: Sekretariatet for IT i Folkeskolen. Retrieved from https://www.uvm.dk/puljer-udbud-og-prisuddelinger/puljer/puljeoversigt/tidligere-udmeldt e-puljer/grundskole/udviklingspulje-til-digitale-laeremidler

[34] Bremholm, J., Bundsgaard, J., Fougt, S. S. & Skyggebjerg, A. K. (Eds.). Læremidlernes danskfag [The L1 subject of the learning resources]. Aarhus: Aarhus University Press[35] Gissel, S. T. & Skovmand, K. (2018). Kategorisering af digitale læremidler. En undersø gelse af didaktiske, digitale læremidlers karakteristika [A categorization of digital learning resources. An examination of the characteristics of didactic, digital learning resources]. Odense: Læremiddel.dk

[36] Tomlinson, J. (1986). Public Education, Public Good. Oxford Review of Education, 12(3), 211-222. https://doi.org/10.1080/0305498860120301

[37] Dehning, B., Richardson, V. J., & Zmud, R. W. (2003). The Value Relevance of An nouncements of Transformational Information Technology Investments. MIS Quarterly, 27(4), 637–656. https://doi.org/10.2307/30036551

[38] Agarwal, R. et al (2010). Research Commentary—The Digital Transformation of Healthcare: Current Status and the Road Ahead. Information Systems Research, 21(4), 796-809. https://doi.org/10.1287/isre.1100.0327

[39] Baker, M. (2015). Digital Transformation. Buckingham Business Monographs.

[40] Rogers, D. L. (2016). The Digital Transformation Playbook: Rethink your business for the digital age. Columbia Business School Publishing. https://doi.org/10.7312/roge17544

[41] Rangel, V. S., Monroy, C. & Bell, E. R. (2016). Science teachers' data use practices: A descriptive analysis. Education Policy Analysis Archives, 24(86), 1-38. https://doi.org/10.14507/epaa.24.2348

[42] Klausen, M. S. (2018). Skoleledelse med data – i praksis [School management with data –in practice]. Copenhagen: Samfundslitteratur.

[43] Olsen, H. S., Marcher, H. & Hornskov, S. B. (2018). Når data kommer på bordet: ledelse med data, der giver mening [When data is available: management with data that makes sense]. Frederikshavn: Dafolo.

[44] Hestbæk, A.-D. & Friis-Hansen, M. (Eds.) (2017). Elevers læring og progression – Rap port fra et udviklingsprojekt om visualisering af data om elevers læring og progression[Student learning and progression – Report from a development project about data visuali zation of student learning and progression]. Copenhagen: VIVE – Det Nationale Forsk nings- og Analysecenter for Velfærd. https://doi.org/10.7146/aul.194.141

[45] OECD (2017). OECD Digital Economy Outlook 2017. Paris: OECD Publishing. http://dx.doi.org/10.1787/9789264276284-en

[46] Meyer, B. (2015). iPads in inclusive classrooms: Ecologies of learning. In. Isaías, P., Spec tor, M., Ifenthaler, D. & Sampson, D. G. (Eds.). E-learning systems, environments and approaches (pp. 25–37). New York: Springer International Publishing. https://doi.org/10.1007/978-3-319-05825-2_3

[47] KL (2015). Referencearkitektur for brugerportalen EFTER HØRING [Reference archictecture for the digital portal efter hearing]. Retrieved from https://www.kl.dk/ImageVaultFiles/id_76760/cf_202/Refer encearkitektur_for_brugerportalen_EFTE

[48] Chahal, H., Jyoti, J. & Wirtz, J. (2019). Business Analytics: Concept and Applications. In. Chahal, H., Jyoti, J. & Wirtz, J. (Eds.). Understanding the Role of Business Analytics (pp. 1-8). Singapore: Springer. https://doi.org/10.1007/978-981-13-1334-9_1

[49] Reyes, J. A. (2015). The skinny on big data in education: Learning analytics simplified. TechTrends, 59(2), 75-80. https://doi.org/10.1007/s11528-015-0842-1

[50] Kansanen, P., Tirri, K., Meri, M., Krokfors, L., Husu, J. & Jyrhämä (2000). Teachers' Pedagogical Thinking. Theoretical Landscapes, Practical Challenges. New York: P. Lang. https://doi.org/10.3726/978-1-4539-1514-1

[51] White, G. (2019). The Application of Ethical Themes for Responsible Data-Driven Deci sion Making in Education. In. Khan, B. H., Corbeil, J. R. & Corbeil, M. E. (Eds.). Responsible Analytics and Data Mining in Education. Global Perspectives on Quality, Support, and Decision Making. New York & London: Routledge. https://doi.org/10.4324/9780203728703

[52] Rao, Ding & Gudivada (2019). Supporting Data Analytics in Education. Human and Tech nical Resources for Collecting, Storing, Analyzing, and Mining Data. In. Khan, B. H., Corbeil, J. R. & Corbeil, M. E. (Eds.). Responsible Analytics and Data Mining in Educa tion. Global Perspectives on Quality, Support, and Decision Making. New York & Lon don: Routledge. https://doi.org/10.4324/9780203728703-2

[53] Mandinach, E. B. & Jimerson, J. B. (2016). Teachers learning how to use data: A synthesis of the issues and what is known. Teaching and Teacher Education, 60, 452-457. https://doi.org/10.1016/j.tate.2016.07.009

8 Authors

Ulf Dalvad Berthelsen is an Associate Professor at The Danish School of Educa tion, Aarhus University, Jens Chr. Skous Vej 4, building 1483-543, 8000 Aarhus C, Denmark, E-mail: ulda@edu.au.dk.

Morten Tannert is a research assistant at The Danish School of Education, Aarhus University, Denmark, E-mail: mota@edu.au.dk.

Gamification and Student Engagement with a Curriculum-Based Measurement System

Yu Yan (*), Simon Hooper, Shi Pu
The Pennsylvania State University, University Park, United States
ACT Inc., Iowa City, United States

A B S T R A C T

In this study, we employed a random control experiment to evaluate the effectiveness of gamification (e.g. scores, goal, progressive bar, etc.) and initial task difficulty on college student engagement with computer-based assessments. A group of Chinese college students (N=97) were randomly assigned to four groups obtained by crossing the two independent variables: gamification (with or without) and entry level difficulty (low or normal). Students completed several English reading tests (maze tests) for 35 minutes. Student engagement was measured by the average off-task time between two maze tests. The results showed that both gamification and low-difficulty entry level reduced students' off-task time. However, the gamification effect was only significant for male students but not for female students. The study also demonstrated that the maze test can be a potential method to predict general English proficiency with Chinese English language learners.

Keywords—Gamification, engagement, computer-based assessment, curriculum-based measurement, English language learners, task difficulty, gender

1 Introduction

In the late 20th century, notions of effective instruction changed as the theoretical underpinnings of learning expanded from behaviorism to constructivism and socialcultural theory [1]. Our understanding of the role educational assessments likewise shifted from summative assessment to formative assessments, emphasizing that the function of assessment is to provide students with details of their progress to support their learning [2]. Technology plays a critical role in supporting this change by im proving efficiency, reducing cost, and assisting in the development of adaptive tasks [3], [4].

This manuscript presents results from the study of a computer-based formative assessment system. This section presents the background and context of the study, describes curriculum-based measurement and gamification, provides a rationale for the current research, and concludes with our research questions.

1.1 Background and context

Echoing the transition from summative to formative assessment, China's Ministry of Education initiated a curriculum reform to encourage computer-based formative assessment in college English teaching and learning. In 2004, the Chinese Ministry of Education (CMOE) began to reform the college English curriculum by implementing two main changes: increasing computer-assisted learning in English to individualize students' learning plans; and, adding formative assessments to the extant summativeheavy assessment system to motivate and inform students about their progress. In response to these reforms, many Chinese universities have developed computersupported formative assessments to support teaching English—an approach that has been praised by both teachers and students. Formative assessment can assist learning by providing feedback on students' progress. One formative assessment approach involves the use of curriculum-based measurement (CBM). CBM uses reliable, sim ple, and brief standardized assessments that provide teachers and students with func tional performance information that can be used to guide instructional decision mak ing. CBM uses commonplace measurement tasks (e.g., maze tests, reading aloud from a text, and written word sequences), and standardized scoring rubrics to assess each task.

Originally developed to test the effectiveness of an individualized educational plan for special education, CBM is now used widely in elementary and secondary educa tion to monitor students' progress in math, literacy, and language learning in the United States [5]. CBM helps teachers monitor students' learning progress over time and customize instructional interventions for students based on individual perfor mance. Technical adequacy is an important characteristic of CBM [6]. Previous research has demonstrated the reliability and validity of diverse CBM assessments in reading [7], writing [8], and mathematics [9].

Computer administration of CBM can reduce teachers' workload by auto-scoring assessments, improving scoring accuracy, and individualizing feedback. However, to deliver sound and timely information on students' progress, assessments must be valid, reliable, and efficient. Previous research has demonstrated the reliability and validity of using computerized CBM, and teachers have responded favorably to computerized assessment administration [10], [11].

1.2 Curriculum-based measurement and English language learners

Although initially developed to monitor students' math and literacy progress in special education, CBM use has expanded to general education and second language learners [5]. For example, CBM-Reading has been used in the US with English lan guage learners (ELLs) whose native languages are Spanish, Chinese (Mandarin), Japanese, and Arabic, and research has demonstrated that CBM-Reading is sensitive to ELLs' growth in reading skills [12]–[14].

In 2014, 99% percent of Chinese undergraduate students enrolled in regular highereducation institutes were taking English classes [15], [16]. English proficiency plays the role of "gatekeeper" for advanced degrees. Potential graduates are not awarded bachelor's degrees if their English performance does not meet the national standard.

Efficient formative assessment approach is needed for this large population.

However, traditional norm-referenced standardized tests may not be valid for lan guage minority students (i.e., students whose primary or home/native language is not English) due to cultural and linguistic factors [14]. For ELLs, CBM appears to be a promising assessment approach. Thus, it is important, to investigate whether using CBM with Chinese ELLs helps teachers identify students with language learning difficulties. Most studies on CBM and ELLs focus on bilingual students who are learning English in English-speaking countries (for an overview, see Sandberg & Reschly [14]). One recent study was found to investigate the efficacy of using CBM in a non-native English-speaking country. Chung and Espin [17] examined the validi ty of using CBM with Dutch students who were learning English. They reported alternate-form reliabilities of maze scores ranging from 0.44 to 0.88 as well as correla tions between maze scores and English course scores ranging from 0.20 to 0.79. The results suggest that CBM has the potential to be a reliable and valid predictor of ELLs' English proficiency. Because of its small sample size, however, the study can only be considered exploratory. More research is needed to confirm the findings.

1.3 Research on gamification and students' engagement

Encouraging active participation and sustaining motivation while completing com puter-based assessments are critical to maintaining the efficacy of formative assess ments. Many studies have researched the relationship between students' motivation and their test performance [18], [19]. A literature review of 12 empirical studies indicated that unmotivated students scored more than one-half a

standard deviation lower than highly motivated students [20]. Students' motivation also influences the effectiveness of computer-based assessment. For example, research has demonstrated that the extent to which students come to "accept" an assessment system influences their willingness to use CBA [21].
Many strategies have been used to increase students' motivation in online learning and assessment environments. One commonly used approach, gamification, has attracted considerable attention from the education and business sectors [22], [23].
Gamification is the use of game elements in non-gaming contexts to improve user experience and motivation [24]. Results from many gamification studies suggest that gamification has a positive influence on students' motivation and performance. Using gamification in CBM may increase students' participation and help sustain motivation.

1.4 Research questions

The current study addressed the use of a gamified CBM assessment system with Chinese ELLs. We examined the effect of gamification on students' engagement when using a CBM system. In addition, given the benefits of CBM with at-risk school-age students [6], we were interested in whether CBM would help teachers to identify students who experience difficulties learning a second language.

The study included two formal research questions:
RQ1: Do gamification and low-difficulty entry level improve students' engage ment when taking assessments?
RQ2: Can the maze test results predict ELLs' English course grades?

2 Method

2.1 Participants

Participants were recruited from a four-year, second-tier regional college in the Si chuan province of southwestern China. The college had an enrollment of approximately 12,000 undergraduate students, of whom 68.4% were female. All and sophomores were required to take a weekly 90-minute English language course each semester. Students from two English classes taught by one English teacher (N=103) were recruited for the pilot study. However, six students chose not to partici pate in the study, yielding a final sample of 97 students. The students received extra credit for taking part in the study.

2.2 Materials

The materials included two versions (gamified vs. none-gamified) of a web-based CBM system, known as Avenue: PM. A full description of the software (including the management system, CBM assessments and scoring rules, teacher scoring interfaces, and data performance charts) is presented elsewhere [25]. Students participating in this study only took the maze test in Avenue: PM. Maze assessments: A maze test is a common CBM assessment used to monitor students' reading progress [26], [27]. In a maze task, students are given a text passage with a blank space for every seventh word, excluding the words in the first sentence.
The maze passage includes three choices: one is the correct answer and the other two are distractors. The distractors are taken from a different, randomly selected part of speech rather than the part of speech of the correct answer. For example, if the correct answer were a noun, the two distractors may be verbs, adjectives, determiners, con junctions, or prepositions. The incorrect choices are randomly generated from a dis tractors' pool in Avenue: PM. The maze passage is timed to one minute. Students read the passage and use the mouse to select the word for each blank.

Test passages are written at 12 reading levels, with every two levels corresponding to one grade of American students' reading proficiency in primary school. The success criterion (a score from 3–14) is the total score (i.e., correct – incorrect items) required for a student to move forward within a level. Students begin with passages presented at a predetermined reading level (i.e., 1–12), but move up or down levels according to their performance. Each level includes six steps, and students move up or down a step depending on whether they achieve the success criterion for a passage. Reaching step 6 moves the student up a level and falling below step 1 moves the student down a level (in both cases re-starting at step 2). The system uses an algorithm to select passages in a random order without repetition. After completing a passage, the student receives a score and correct/ incorrect feedback. The system contains 237 passages written by subject-matter experts, and each level includes between 13 and 40 passages.

Two versions of the maze: were developed for this study. Both present students with maze assessments, but they differ in the presence or absence of gaming elements designed to enhance student motivation and engagement.

Version 1: Maze with gamification features. Effectively designed game environments create a sense of flow [28], resulting in improved concentration, joy, and in volvement [29]. The game elements employed in the maze assessment aim at generat ing positive emotional experiences by "shift[ing] students' perceptions about the tasks from 'testing' environments to 'practice' or 'gaming' environments" [25].

The gamified version of the software includes four elements (see Figure 1). The first game element is progressive difficulty levels. The test passages become more complex as the levels increase. Students move up or down between levels according to their current reading performances in the system, similar to that which occurs in adaptive assessments. Students receive notices when their current level. The second element is the presence of visually appealing images of animal characters that repre sent the levels. Low-difficulty levels use characters that are lower on the food chain (e.g., a crustacean or jellyfish) and high-difficulty levels use more advanced animals (e.g., an elephant or lion). Third, a progress bar is presented within each level. The progress bar is similar to the energy graphic that appears in many online games. Students can see where they are in each level and how many more assessments they need to pass before moving on to the next level. The fourth element is immediate feedback on the score. After completing a maze test, students automatically see the correct answers and their total score.

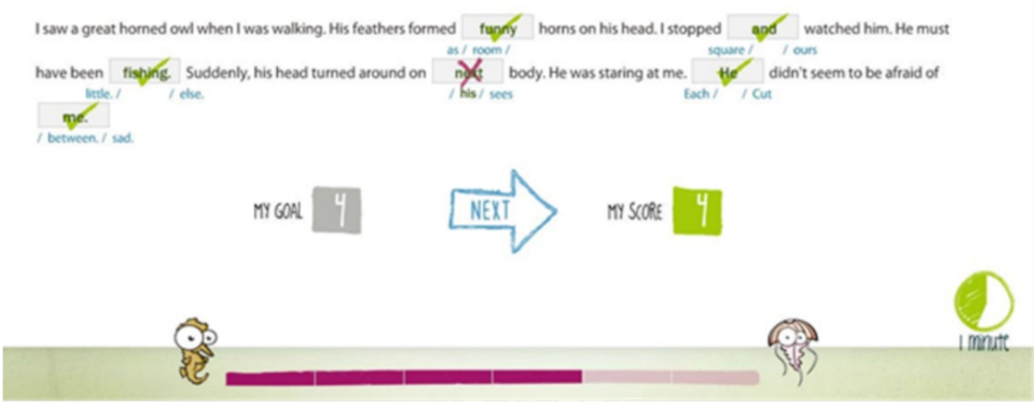

Fig. 1. Maze with gamifications

Version 2: Maze without gamification features: The maze test includes embedded game elements. To test the effect of gamification, a non-gamified version of maze was developed and compared with the original maze test with gamification features. Figure 2 includes examples of the non-gamified version. The level notification that is shown in the gamified system is hidden in the non-gamified version. Table 1 shows the differences between the two versions of maze tests.

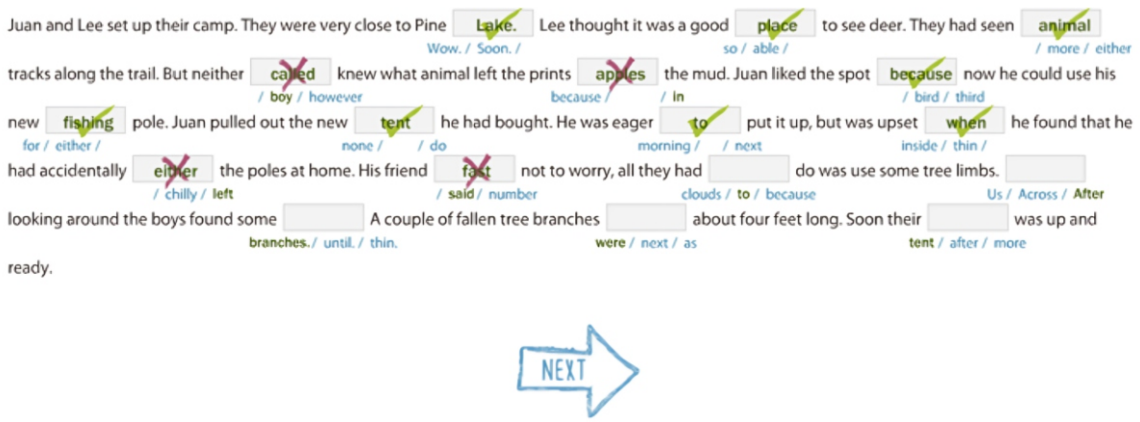

Fig. 2. Maze without gamifications

Students using non-gamified maze test do not know their current level, their progress within a level, or their score for each passage. Despite the two versions' different interfaces, the progress mechanism remains the same. Thus, even though students using the non-gamified version are unaware of their progress, they move forward and backward in levels.

Table 1. Table 1. Interface Differences between Two Versions of the Maze

	With gamification	Without gamification
Animal characters indicating levels	Yes	No
Progress bar	Yes	No
Goal for passing each level	Yes	No
Correct score feedback	Yes	No

Entry level difficulty: There are 12 difficulty levels for the maze test. Test passages in successive levels become progressively difficult in both lexical complexity and the success criterion. As described in the previous session, students begin with passages presented at a predetermined level (i.e., 1–12), and move up or down levels according to their performance. Reading passages used in levels 1 – 4 are of low diffi culty (i.e., grade equivalent reading level 1 – 2.5), levels 5 – 8 are of normal difficulty (i.e., grade equivalent reading level 3 – 4.5), and levels 9 – 12 are of high difficulty (i.e., grade equivalent reading level 5 – 6.5).

Criterion test: The final exam in the English course was used as the criterion test in this study. The final exam assesses students' English-language ability in different areas, including reading, grammar, vocabulary, and writing. It is a one-hour, paperbased exam taken at the end of the semester. The exam is scored using a range of 0 100, with 100 representing the highest possible score.

2.3 Procedure

Students were randomly assigned into four groups obtained by crossing the two independent variables, gamification and entry level difficulty. They used the gamified or non-gamified Avenue: PM system, depending on their group assignment. The students began the maze test in Avenue: PM at either level 3 (low-difficulty entry level) or level 6 (normal entry level). The two different entry levels were suggested by the students' English course teacher. The descriptive statistics covering the students in each treatment group are shown in **Fehler! Verweisquelle konnte nicht gefunden werden..**

The study was conducted during a 45-minute English class session in a computer lab. Students in the class received a printed handout of the system instructions. The students were given 35 minutes to use the system, but due to an unreliable Internet connection, not all students were able to use the entire 35 minutes for testing.

Table 2. Descriptive Statistics across Treatment Groups

	With gamification	Without gamification	Total
Low-difficulty entry level	*N* = 25	*N* = 24	*N* = 49
Normal entry level	*N* = 23	*N* = 25	*N* = 48
Total	*N* = 48	*N* = 49	*N* = 97

2.4 Data analysis

To examine the impact of gamification and entry level difficulty on students' engagement for research question one, students' average off-task time between maze assessments was used as the dependent variable. The average off-task time was measured by the total off-task time divided by the total number of maze passages completed. Total off-task time and total mazes completed were extracted directly from the database. Higher off-task time indicates lower student engagement. The independent variables included system version (i.e., with or without gamification), entry level difficulty (i.e., low-difficulty or normal), and gender (i.e., male or female).

For research question two, students' final exam scores served as the criterion vari able. Students' end level on the maze tests was used as the predicting variable. Pear son correlations were calculated between students' end level on the maze tests and their final score. In addition, differences in maze end level by gender was also exam ined to assess the validity of the maze test as an indicator of students' English ability.

3 Results

RQ1: Do gamification and difficulty affect students' engagement?

Table 3 reports the mean and standard deviations of off-task time for four treatment groups, and Table 4 reports the mean and standard deviation for the four treatment groups broken down by gender.

Table 3. Students' Off-task Time (in seconds) per Passage by Treatment Groups

Treatment groups	N	Mean	SD
Gamification & low-difficulty entry	25	11.56	6.72
Gamification & normal entry	23	13.60	5.46
No Gamification & low-difficulty entry	24	10.92	6.61
No Gamification & normal entry	25	16.31	6.00
Total	97	13.18	6.47

Table 4. Students' Off-task Time (in seconds) per Passage by Treatment Groups and Gender

Treatment groups	Male			Female		
	N	*Mean*	*SD*	*N*	*Mean*	*SD*
Gamification & low-difficulty entry	9	14.35	10.32	16	9.99	2.86
Gamification & normal entry	8	15.26	6.68	15	12.71	4.70
No Gamification & low-difficulty entry	1	17.64	0.00	23	10.92	6.61
No Gamification & normal entry	3	26.55	8.45	22	14.92	4.17
Total	21	16.60	9.12	76	12.24	5.22

Raw data for the off-task time were not normally distributed, therefore, a log trans formation was performed on the off-task time before conducting the subsequent ANOVA analysis. After the log transformation, normality and homogeneity of variance were achieved.

A three-way (2×2×2) ANOVA was used to examine whether gamification, difficul ty, and gender influenced students' engagement as measured by off-task time. The main effect for gamification was significant, [$F(1, 95) = 5.28$, $p < .05$], as was the main effect for low-difficulty entry level [$F(1, 95) = 7.44$, $p < .01$]. The results also showed a significant gender effect, [$F(1, 95) = 9.93$, $p < .01$], with lower off-task time for female students than for male students. The two- and three-way interactions were not significant for all independent variables (all $p > .05$).

To address the concern of "single observation" and the issue of the imbalanced sample in the three-way ANOVA, a follow-up two-sample t-test was conducted to examine the effect of gamification on male and female groups separately. Gamifica tion significantly reduced off-task time for male students ($t = 2.17$, $p < .05$), but not for female students ($t = 0.92$, $p > .05$). The results showed that gamification had a positive effect on reducing the off-task time for a specific gender group.

RQ2: Can maze tests performance predict students' English ability?

To examine the validity of the maze tasks, the correlation between students' end level on the maze tests and their final English score was examined. We also compared the mean differences of maze end level between male and female students. Table 5 reports the Pearson correlation between maze end level and course grade, which ranged from .42 to .48 in the overall sample, within male students, and within female students. All correlations were significant at the .05 level.

Table 5. Correlation between Maze End Level and English Course Exam Score

Gender	N	r	p
Male	21	0.42	0.04
Female	76	0.45	0.04
Total	97	0.48	0.00

The mean difference between gender groups was also examined. For the maze tasks, a statistically significant difference ($t = 4.63$, $p < .001$) was found between the two gender groups, with females achieving a higher ending maze level than male students. This result matched the difference in the final course grade between the two gender groups ($t = 6.04$, $p < .001$). Table 6 includes mean and standard deviations of students' English test score and maze end level by gender.

Table 6. Students' English Course Exam Score and Maze Tasks End Level by Gender

		English course exam score				Maze tasks' end level			
	N	*Mean*	*SD*	*Min*	*Max*	*Mean*	*SD*	*Min*	*Max*
Male	21	48.38	10.43	31	72	4.05	1.687	1	7
Female	75	63.70	10.25	34	83	5.81	1.522	2	9
Total	96	60.38	12.04	31	83	5.43	1.715	1	9

4 Discussion

This study examined the effect of gamification and initial difficulty on Chinese ELLs' engagement in a CBM assessment system. In addition, the study investigated the validity of using maze assessments to measure Chinese ELLs' English proficiency. This section discusses the main findings, implications, and limitations of the study as well as future directions.

Gamification improved engagement by reducing students' off-task time. Using the system with game features (i.e. progress bar, level, and scores) resulted in students spending significantly less off-task time between assessments. This result aligns with previous findings suggesting that gamification can increase students' effort during assessments [30] and improve students' motivation when using a tutoring system [31].

Despite the general positive effect, it is interesting that gamification only reduced male students' off-task time. Two explanations for this finding are offered. First, male students enjoy video games more than do female students [32], [33]; thus, they might be more stimulated and more engaged in the gamified system. Second, levels and points may be more effective for students with lower abilities [30]. In our sample, the female students had higher English proficiency (thus hypothetically, higher English ability) than male students as shown in Table 6, therefore, gamification may have less effect on female students than male students.

However, caution is urged in generalizing these results as the existing literature on the role of gender in gaming environments is mixed. For example, De Jean, Upitis, Koch, and Young [34] found that gender played a key role on learning outcomes and attitude in a gamified learning environment. However, other studies have found gamification to be equally motivating for male and female high school students' in a computer science course [35] or had no gender effect on fifth-graders' math performance and attitude [36]. Recent studies argued that male and female preferred different types of games [37], and females were motivated more by the social aspects of gamification [38]. Therefore, the gender differences in the effectiveness of gamification would be influenced by the design (e.g., the nature of the gamification elements) of the gamification environment.

The current study also demonstrated that low-difficulty entry level improved students' engagement by reducing off-task time while using the system. Students who started at easier levels displayed significantly less off-task time than those starting from normal entry level. Starting with an easier task appears to be a good strategy for improving students' engagement when completing maze tests. Previous research suggested that gaming environments can effectively motivate students by providing "challenging but not overwhelmingly difficult" experiences [39]. When the challenge and difficulty are appropriately balanced, a student will be immersed in a "flow" status [28], in which a person becomes engaged in an activity with deep concentration and enjoyment.

The validity analysis suggests that maze task score is a promising indicator of general English proficiency for Chinese ELLs. First, the correlation between ending maze level and the English course grade is significant for both the overall sample ($r = .48$) and the gender subgroups ($r = .45$ for female; $r = .42$ for male). These findings are consistent with a previous study by Chung and Espin [17], who reported correlations between maze scores and English course grades ranging from 0.19 to 0.79 with Dutch ELLs. In addition, the gender gap in the maze score is consistent with the gender gap in students'

English course scores. Gender differences in foreign language learning is a well-acknowledged phenomenon (e.g., [40]). Such differences may be attributable to a difference in brain functions or from the difference in learning strategies[41], [42]. Indeed, Chinese female students usually achieve higher than male students in both English vocabulary and general proficiency tests (e.g., [43]). Therefore, differences in maze performance between male and female groups appear to support the maze as a valid indicator of general English proficiency. In the current study, female students significantly outperformed male students in the English course exam. The analysis indicates that the maze task is sensitive to identifying students with different levels of English proficiency in general.

The current study has three notable strengths. First, the gamified system was not compared to a traditional paper-based assessment or a different computer-based assessment, but to a system identical to the gamified one with the exception of a few gamification features in the interface. Therefore, the differences in students' behaviors were thus due to the gamification effect, not the different format of the test (i.e., pa per-based test vs computer-based test). Second, the participants were not children, but college students who may not be easily motivated by gamification elements employed the system (e.g. visually appealing images). Arguably, the effect of the same gamified system on children or adolescents might be larger than the effect found in this study.

The third strengthen is that this study used engagement indicators (i.e., off-task time) derived from students' recorded behaviors in the assessment system, which may be more accurate and objective than data collected from subjective surveys.

Limitations of the study are also noted. The original research question did not consider the gender effect in the design stage; therefore, even though students were ran domly assigned to the four treatment groups, the distribution of male and female stu dents was not equal across the four groups. The gamification effect can be more relia ble with a balanced design. Furthermore, this study measured students' engagement over a short period of time (35 minutes). Previous research has pointed out that the perceived benefits of gamification decline with use [38]. Therefore, the findings of this study might not be applicable in a long-term study.

Implications for future research are offered. To begin, future study might investi gate the effects of voluntary participation1. Voluntary participation allows students to enter or leave a game at will. This mechanism can transform a challenging learning process into a more pleasurable experience. This study was only semi-voluntary: students were given 35 minutes during their English class to explore the system. Stu dents did not have an alternative activity available to them. It would be interesting to investigate whether students' behaviors would change without teachers and peers' presence. Second, future studies could consider the interaction effect of the teacher's presence and gamification. While both have the potential to boost student engage ment, teachers' presence might diminish the effect of gamification, especially in an educational culture where gaming in class is not a traditional instructional activity but associated with disruptive and dis-engaging behavior that requires teachers' interven tion. Third, it would be interesting to investigate the longer-term effects on students' motivation and behaviors while using the Avenue: PM system.

In conclusion, the present study provides support for using the maze task with Chi nese ELLs and demonstrates the complexity of designing educational assessment systems with game features. It also opens new directions for further research on using gamification in a CBM system.

5 Acknowledgement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. We would like to thank Mr. Zhenggang Shi, an English professor at Tianfu College of Southwestern University of Finance and Economics, China, for his assistance with the study set up.

6 References

[1] R. K. Sawyer, "The new science of learning," in The Cambridge handbook of the learning sciences, R. K. Sawyer, Ed. New York: Cambridge University Press, 2006, pp. 1–16.

[2] National Research Council, Knowing what students know: The science and design of educational assessment. Washington D.C.: National Academies Press, 2001.

[3] J. W. Pellegrino, "Measuring what matters? Technology and the design of assessments that support learning," in Handbook of Design in Educational Technology, R. Luckin, S. Puntambekar, P. Goodyear, B. L. Grabowski, J. Underwood, and N. Winters, Eds. Routledge, 2013, pp. 377–387.

[4] J. W. Pellegrino and E. S. Quellmalz, "Perspectives on the integration of technology and assessment," Journal of Research on Technology in Education, vol. 43, no. 2, pp. 119–134, 2010. https://doi.org/10.1080/15391523.2010.10782565

[5] S. L. Deno, "Curriculum-based measures: Development and perspectives," Assessment for Effective Intervention, vol. 28, no. 3&4, pp. 3–12, 2003. https://doi.org/10.1177/073724770302800302

[6] S. L. Deno, "Developments in curriculum-based measurement," The Journal of Special Education, vol. 37, no. 3, pp. 184–192, Nov. 2003.

[7] M. M. Wayman, T. Wallace, H. I. Wiley, R. Tichá, and C. A. Espin, "Literature synthesis on curriculum-based measurement in reading," The Journal of Special Education, vol. 41, no. 2, pp. 85–120, Aug. 2007. https://doi.org/10.1177/00224669070410020401

[8] C. A. Espin, J. Shin, S. L. Deno, S. Skare, S. Robinson, and B. Benner, "Identifying indi cators of written expression proficiency for middle school students," The Journal of Spe cial Education, vol. 34, no. 3, pp. 140–153, Oct. 2000. https://doi.org/10.1177/002246690003400303

[9] P. M. Stecker, L. S. Fuchs, and D. Fuchs, "Using curriculum-based measurement to improve student achievement: Review of research," Psychology in the Schools, vol. 42, no. 8, pp. 795–819, 2005. https://doi.org/10.1002/pits.20113

[10] N. Förster and E. Souvignier, "Curriculum-based measurement: Developing a computer based assessment instrument for monitoring student reading progress on multiple indica tors," Learning Disabilities: A Contemporary Journal, vol. 9, no. 2, pp. 65–88, 2011.

[11] M. Tsuei, "A web-based curriculum-based measurement system for class-wide ongoing as sessment," Journal of Computer Assisted Learning, vol. 24, no. 1, pp. 47–60, May 2008.https://doi.org/10.1111/j.1365-2729.2007.00242.x

[12] R. Domínguez de Ramírez and E. S. Shapiro, "Curriculum-based measurement and the evaluation of reading skills of Spanish-speaking English language learners in bilingual education classrooms," School Psychology Review, vol. 35, no. 3, pp. 356–369, 2006.

[13] E. Farmer, "Examining predictive validity and rates of growth in curriculum-based measurement with English language learners in the intermediate grades," Loyola University Chicago, 2013.

[14] K. L. Sandberg and A. L. Reschly, "English learners: Challenges in assessment and the promise of curriculum-based measurement," Remedial and Special Education, vol. 32, no. 2, pp. 144–154, 2011. https://doi.org/10.1177/0741932510361260

[15] National Bureau of Statistics of China, China statistical yearbook. China Statistics Press, 2014.

[16] Q. Wang, "The national curriculum changes and their effects on English language teaching in the People's Republic of China," in International Handbook of English Language Teaching SE - 8, vol. 15, J. Cummins and C. Davison, Eds. Springer US, 2007, pp. 87–105. https://doi.org/10.1007/978-0-387-46301-8_8

[17] S. Chung and C. A. Espin, "CBM progress monitoring in foreign language learning for secondary school students: Technical adequacy of different measures and scoring proce dures," Assessment for Effective Intervention, vol. 38, no. 4, pp. 236–248, Jun. 2013.

[18] J. Baumert and A. Demmrich, "Test motivation in the assessment of student skills: The effects of incentives on motivation and performance," European Journal of Psychology of Education, vol. 16, no. 3, pp. 441–462, 2001. https://doi.org/10.1007/bf03173192
[19] L. F. Wolf and J. K. Smith, "The consequence of consequence: Motivation, anxiety, and test performance," Applied Measurement in Education, vol. 8, no. 3, pp. 227–242, 1995.https://doi.org/10.1207/s15324818ame0803_3
[20] S. L. Wise and C. E. DeMars, "Low examinee effort in low-stakes assessment: Problems and potential solutions," Educational Assessment, vol. 10, no. 1, pp. 1–17, 2005.https://doi.org/10.1207/s15326977ea1001_1
[21] V. Terzis and A. A. Economides, "The acceptance and use of computer based assessment," Computers & Education, vol. 56, no. 4, pp. 1032–1044, May 2011. https://doi.org/10.1016/j.compedu.2010.11.017
[22] J. McGonigal, Reality is broken: Why games make us better and how they can change the world. Penguin, 2011.
[23] T. Reiners and L. C. Wood, Eds., Gamification in Education and Business, 1st ed. Springer International Publishing, 2015.
[24] S. Deterding, M. Sicart, L. Nacke, K. O'Hara, and D. Dixon, "Gamification: Using game design elements in non-gaming contexts," in CHI '11 Extended Abstracts on Human Fac tors in Computing Systems, 2011, pp. 2425–2428. https://doi.org/10.1145/1979742.1979575
[25] S. Hooper, C. Miller, and S. Rose, "Considering the design of an electronic progress monitoring system," in Handbook on Design in Educational Computing, R. Luckin, P. Goodyear, B. Grabowski, S. Puntambeker, J. Underwood, and N. Winters, Eds. Routledge, 2013. https://doi.org/10.4324/9780203075227.ch22
[26] J. Shin, S. L. Deno, and C. A. Espin, "Technical adequacy of the maze task for curriculum based measurement of reading growth," The Journal of Special Education, vol. 34, no. 3, pp. 164–172, Oct. 2000. https://doi.org/10.1177/002246690003400305
[27] R. Tichá, C. A. Espin, and M. M. Wayman, "Reading progress monitoring for secondary school students: Reliability, validity, and sensitivity to growth of reading-aloud and mazeselection measures," Learning Disabilities Research & Practice, vol. 24, no. 3, pp. 132–142, Aug. 2009. https://doi.org/10.1111/j.1540-5826.2009.00287.x
[28] M. Csikszentmihalyi, Flow: The psychology of optimal experience. 1990.
[29] D. Johnson and J. Wiles, "Effective affective user interface design in games," Ergonomics, vol. 46, no. 13–14, pp. 1332–1345, 2003. https://doi.org/10.1080/00140130310001610865
[30] Y. Attali and M. Arieli-Attali, "Gamification in assessment: Do points affect test perfor mance?," Computers & Education, vol. 83, pp. 57–63, 2015. https://doi.org/10.1016/j.compedu.2014.12.012
[31] G. T. Jackson and D. S. McNamara, "Motivation and performance in a game-based intelli gent tutoring system," Journal of Educational Psychology, vol. 105, no. 4, pp. 1036–1049, 2013. https://doi.org/10.1037/a0032580
[32] C. Chou and M. J. Tsai, "Gender differences in Taiwan high school students' computer game playing," Computers in Human Behavior, vol. 23, no. 1, pp. 812–824, 2007.https://doi.org/10.1016/j.chb.2004.11.011
[33] T. Karakus, Y. Inal, and K. Cagiltay, "A descriptive study of Turkish high school students' game-playing characteristics and their considerations concerning the effects of games," Computers in Human Behavior, vol. 24, no. 6, pp. 2520–2529, 2008. https://doi.org/10.1016/j.chb.2008.03.011
[34] J. De Jean, R. Upitis, C. Koch, and J. Young, "The Story of Phoenix Quest: How girls respond to a prototype language and mathematics computer game," Gender and Education, vol. 11, no. 2, pp.

207–223, 1999. https://doi.org/10.1080/09540259920708
[35] M. Papastergiou, "Digital game-based learning in high school Computer Science educa tion: Impact on educational effectiveness and student motivation," Computers & Educa tion, vol. 52, pp. 1–12, 2009. https://doi.org/10.1016/j.compedu.2008.06.004
[36] F. Ke and B. Grabowski, "Gameplaying for maths learning: Cooperative or not?," British Journal of Educational Technology, vol. 38, no. 2, pp. 249–259, 2007. https://doi.org/10.1111/j.1467-8535.2006.00593.x
[37] T. Hartmann and C. Klimmt, "Gender and computer games: Exploring females' dislikes," Journal of Computer-Mediated Communication, vol. 11, no. 4, pp. 910–931, 2006.https://doi.org/10.1111/j.1083-6101.2006.00301.x
[38] J. Koivisto and J. Hamari, "Demographic differences in perceived benefits from gamifica tion," Computers in Human Behavior, vol. 35, pp. 179–188, 2014. https://doi.org/10.1016/j.chb.2014.03.007
[39] R. M. Ryan, C. S. Rigby, and A. Przybylski, "The motivational pull of video games: A self-determination theory approach," Motivation and Emotion, vol. 30, no. 4, pp. 347–363, 2006. https://doi.org/10.1007/s11031-006-9051-8
[40] J. Sunderland, "Issues of language and gender in second and foreign language education," Language Teaching, vol. 33, no. 04, pp. 203–223, 2000. https://doi.org/10.1017/s0261444800015688
[41] Y. Aida, "Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety," The Modern Language Journal, vol. 78, no. 2, pp. 155–168, 1994. https://doi.org/10.1111/j.1540-4781.1994.tb02026.x
[42] J. M. Green and R. Oxford, "A closer look at learning strategies, L2 proficiency, and gen der," TESOL quarterly, vol. 29, no. 2, pp. 261–297, 1995. https://doi.org/10.2307/3587625
[43] Y. Gu, "Gender, academic major, and vocabulary learning strategies of Chinese EFL learners," RELC Journal, vol. 33, no. 1, pp. 35–54, 2002. https://doi.org/10.1177/003368820203300102

7 Authors

Yu Yan is a Ph.D. candidate in Learning, Design, and Technology at The Pennsylvania State University, 314 Keller Building, University Park, PA 16802, United States. Her research interests include gamification, student engagement, learning analytics, educational data mining, cognitive load, EEG and flipped classroom.

Simon Hooper is Professor, Learning, Design, and Technology at The Pennsylvania State University, where he designs, develops, and examines the impact of technol ogy to support teaching and learning. His current research interests include progress monitoring, interaction design, learning analytics, and data visualization.

Shi Pu is a data scientist in the AI and Machine Learning team at ACT, Inc, 500 Act Drive, Iowa City, Iowa 52243, United States. He received his Ph.D. degree in Higher Education at The Pennsylvania State University, United States.

Instructions for Authors

Essentials for Publishing in this Journal

1 Submitted articles should not have been previously published or be currently under consideration for publication elsewhere.

2 Conference papers may only be submitted if the paper has been completely re-written (taken to mean more than 50%) and the author has cleared any necessary permission with the copyright owner if it has been previously copyrighted.

3 All our articles are refereed through a double-blind process.

4 All authors must declare they have read and agreed to the content of the submitted article and must sign a declaration correspond to the originality of the article.

Submission Process

All articles for this journal must be submitted using our online submissions system. http://enrichedpub.com/ . Please use the Submit Your Article link in the Author Service area.

Manuscript Guidelines

The instructions to authors about the article preparation for publication in the Manuscripts are submitted online, through the e-Ur (Electronic editing) system, developed by **Enriched Publications Pvt. Ltd**. The article should contain the abstract with keywords, introduction, body, conclusion, references and the summary in English language (without heading and subheading enumeration). The article length should not exceed 16 pages of A4 paper format.

Title

The title should be informative. It is in both Journal's and author's best interest to use terms suitable. For indexing and word search. If there are no such terms in the title, the author is strongly advised to add a subtitle. The title should be given in English as well. The titles precede the abstract and the summary in an appropriate language.

Letterhead Title

The letterhead title is given at a top of each page for easier identification of article copies in an Electronic form in particular. It contains the author's surname and first name initial .article title, journal title and collation (year, volume, and issue, first and last page). The journal and article titles can be given in a shortened form.

Author's Name

Full name(s) of author(s) should be used. It is advisable to give the middle initial. Names are given in their original form.

Contact Details

The postal address or the e-mail address of the author (usually of the first one if there are more Authors) is given in the footnote at the bottom of the first page.

Type of Articles

Classification of articles is a duty of the editorial staff and is of special importance. Referees and the members of the editorial staff, or section editors, can propose a category, but the editor-in-chief has the sole responsibility for their classification. Journal articles are classified as follows:

Scientific articles:

1. Original scientific paper (giving the previously unpublished results of the author's own research based on management methods).
2. Survey paper (giving an original, detailed and critical view of a research problem or an area to which the author has made a contribution visible through his self-citation);
3. Short or preliminary communication (original management paper of full format but of a smaller extent or of a preliminary character);
4. Scientific critique or forum (discussion on a particular scientific topic, based exclusively on management argumentation) and commentaries. Exceptionally, in particular areas, a scientific paper in the Journal can be in a form of a monograph or a critical edition of scientific data (historical, archival, lexicographic, bibliographic, data survey, etc.) which were unknown or hardly accessible for scientific research.

Professional articles:

1. Professional paper (contribution offering experience useful for improvement of professional practice but not necessarily based on scientific methods);
2. Informative contribution (editorial, commentary, etc.);
3. Review (of a book, software, case study, scientific event, etc.)

Language

The article should be in English. The grammar and style of the article should be of good quality. The systematized text should be without abbreviations (except standard ones). All measurements must be in SI units. The sequence of formulae is denoted in Arabic numerals in parentheses on the right-hand side.

Abstract and Summary

An abstract is a concise informative presentation of the article content for fast and accurate Evaluation of its relevance. It is both in the Editorial Office's and the author's best interest for an abstract to contain terms often used for indexing and article search. The abstract describes the purpose of the study and the methods, outlines the findings and state the conclusions. A 100- to 250-Word abstract should be placed between the title and the keywords with the body text to follow. Besides an abstract are advised to have a summary in English, at the end of the article, after the Reference list. The summary should be structured and long up to 1/10 of the article length (it is more extensive than the abstract).

Keywords

Keywords are terms or phrases showing adequately the article content for indexing and search purposes. They should be allocated heaving in mind widely accepted international sources (index, dictionary or thesaurus), such as the Web of Science keyword list for science in general. The higher their usage frequency is the better. Up to 10 keywords immediately follow the abstract and the summary, in respective languages.

Acknowledgements

The name and the number of the project or programmed within which the article was realized is given in a separate note at the bottom of the first page together with the name of the institution which financially supported the project or programmed.

Tables and Illustrations

All the captions should be in the original language as well as in English, together with the texts in illustrations if possible. Tables are typed in the same style as the text and are denoted by numerals at the top. Photographs and drawings, placed appropriately in the text, should be clear, precise and suitable for reproduction. Drawings should be created in Word or Corel.

Citation in the Text

Citation in the text must be uniform. When citing references in the text, use the reference number set in square brackets from the Reference list at the end of the article.

Footnotes

Footnotes are given at the bottom of the page with the text they refer to. They can contain less relevant details, additional explanations or used sources (e.g. scientific material, manuals). They cannot replace the cited literature.
The article should be accompanied with a cover letter with the information about the author(s): surname, middle initial, first name, and citizen personal number, rank, title, e-mail address, and affiliation address, home address including municipality, phone number in the office and at home (or a mobile phone number). The cover letter should state the type of the article and tell which illustrations are original and which are not.

Note