Teachers' Perceptions on Traditional and Non-Traditional Data Visualization for Pedagogical Decision-Making

Maria Cavalcanti[1], Ranilson Paiva[2], Patrícia Espinheira Ospina[3],

Ig Ibert Bittencourt[1], Esther de Freitas[1], Sérgio Amorim[1], Alan Pedro da Silva[1]

[1] Computing Institute – IC
Federal University of Alagoas – UFAL
Maceió, Brazil

[2] Computing and Systems Department – DSC
Federal University of Campina Grande – UFCG
Campina Grande, Brazil

[3] Center of Exact Sciences and Nature – CCEN
Federal University of Pernambuco – UFPE
Recife, Brazil

[1] {mgcs, ig.ibert, ensf, sata, alanpedro}@ic.ufal.br
[2] ranilson@copin.ufcg.edu.br
[3] patespipa@de.ufpe.br

***Abstract*—** In 2012 in the US, about 6.7 million students were enrolled in an online course. In 2013, this number increased 6%. In Brazil, in the same period, the number of enrolled students increased 52%. These facts result in more data (quantity and variety), which are the basis of the pedagogical decisions teachers make to guide their students' learning experience. This mentioned, it is necessary to extract information from these data, but to present them in spreadsheets, tables and static graphics, is not an effective practice. In this context, some authors suggest using data visualization to communicate information clearly and efficiently from the point of view of users, helping them analyze and reason about the data. However, people react differently to different types of visualizations, which made us classify them in traditional or non-traditional, and evaluate how they influence users, and what features could determine the users' preference between these two classes. In this paper, we surveyed 235 teachers in order to evaluate how the different visualizations (traditional and non-traditional) affect the way they evaluated data from an online course. They had to check the visualizations and identify which item contributed the most, and which item contributed less for the performance of the students. The answers were evaluated considering the teachers': age, gender, experience, education and perception on the usefulness of each visualization. Our purpose was to create a model to help in the recommendation of visualizations that help making pedagogical decisions.

Keywords— distance education, data overload, data visualization, pedagogical decision-making and teachers' model

I. INTRODUCTION

We are experiencing a change in the educational paradigm, where lessons at fixed times and available only for students of a particular geographical area, are giving way to online learning environments where a large part of class activities takes place in flexible schedules, and available to those with internet access. This new paradigm is called AAA Learning [1]. It relies on information and communication technology (ICT) to promote learning anywhere, anytime and to anyone.

The increase in supply and demand for courses mediated by online learning environments is an evidence of this change. According to the Babson Survey Research Group, in 2012 approximately 6.7 million students were enrolled in at least one online course. This number increased by approximately 6% in the following year (i.e. in 2013, there were 7.1 million students) [6]. In Brazil, the numbers are significant: according to the Brazilian Distance Education Census (Censo EAD), released by the Brazilian Association of Distance Education (ABED[1]), in 2012, 5.8 million Brazilian students were enrolled in distance learning courses. In 2013, this number increased 52%, resulting in 8.8 million students [4].

However, these courses produce a vast amount and variety of data [7] [16]. Part of these data, generally, is used by teachers in order to make pedagogical decision, which are used to guide their students' learning experience [16].

It is necessary to extract information from these data, but simply presenting them to teachers, in the form of spreadsheets, tables or static graphics, in order to allow them to gather some information it contains, is not an effective practice for some purposes. Some authors suggest creating "the story of your data" to help people visualize the data and its meaning [3]. For this purpose, we can use Data Visualization, defined as the study of the way of representing data using an artistic and interactive approach over more traditional approaches, transforming data into meaningful and useful information [9].

[1] ABED: http://www.abed.org.br/site/pt/institucional/quem_somos/

The goal of data visualization is to communicate information clearly and efficiently from the point of view of users (in this study, the users are the teachers), helping them analyze and reason about the data and the evidence in the visualization [9] [15]. Nevertheless, people react differently to different types of data visualization, which made us classify them into traditional or non-traditional, and evaluate how they influenced users and their capacity to identify pedagogical problems represented in them. We also evaluated which characteristics could drive the users' preference for one of these classes of visualization.

Recent studies investigate the possibility of creating visualizations considering the characteristics and skills of the users [2] [11] [12]. Based on that, we surveyed 333 people. In total, 235 teachers answered the questionnaire, where 168 teachers did it appropriately/completely. The survey objective was to check how different visualizations (1 traditional and 2 non-traditional), affected the way these teachers evaluated data from an online course. They had to check the visualizations created with these data, and identify which item most contributed, and which item contributed less for the students' performance.

We evaluated their answers considering: age, gender, experience, academic background and perception on the helpfulness of each visualization. Our objective was to create a teacher model in order to recommend the appropriate type of visualization, according to the teachers' profile. These profiles are created based on the questionnaire analysis.

The results show that some characteristics affect whether teachers choose a traditional or non-traditional visualization. It was possible to identify the relation between the preferences regarding traditional or non-traditional visualization, and a certain teacher profile (a set of teachers' characteristics). We also measured the effect that each characteristic had on these preferences.

This work is organized as follows: in **Section II**, we present some background knowledge necessary to understand this study; in **Section III**, we explain the motivation, preparation and application of the survey (questionnaire) used in this study; in **Section IV**, we present and discuss the results from the survey analysis; and in **Section V**, we talk about our conclusions, limitations and future work.

II. BACKGROUND

A. Data Visualization

Images have been used as a communication mechanism long before the formalization of written language [13], which indicates the importance of its role to communicate facts and ideas.

The authors in [13] define data visualization as the communication of information using graphical representations. Once, a single image can contain a lot of information, and this information can be processed much faster compared to reading a page containing only words.

According to Krum [8], data visualization is a visual representation of a range of different numerical data, which allows making decisions, comparing data and the identifying possible trends. Through data visualization it is possible to identify, improve and recognize patterns that speed up perception, with decreased effort in understanding the data analyzed.

B. Perception

According to Ward [13], "perception" is a process by which we interpret the world around us, forming a mental representation of the environment. This representation is not seen the same way by all, but is subject to many different matching and error. In this process, the brain produces assumptions for the world, trying to overcome the ambiguity inherent in the sensory data it obtained in response to the proposed tasks.

As concluded by Ward [13], perceptions deal with the human senses, and involves **recognition**, **organization** and **interpretation** of sensory information. The **recognition** step, provides the reader awareness of what is shown to him/her; The **organization** stage is responsible for gathering and storing the main inferences about a certain information; Finally, the **interpretation** stage promotes the connection of inferences to the knowledge of what was shown.

C. MeuTutor-ENEM

The work by Neto [9], defines *MeuTutor*[2] as a gamified educational web environment that focus on personalized learning. It performs the monitoring of the activities of their students in a personalized way.

MeuTutor offers the following school subjects: mathematics, Portuguese, Literature, Biology, Chemistry, Physics, History, Geography, English and Spanish. In order to promote student-content interactions, the environment offers several educational resources such as: video-lessons, problem-solving, interactive quizzes and social networking integration for students to share their accomplishments [10].

We collected data from *MeuTutor* in order to create the 3 visualizations used in the questionnaire. The intention was to provide teachers with real data.

III. THE SURVEY

A. Data Visualization Selection

For the development of this experiment (a survey by printed questionnaire), we chose three (3) different visual representations: one traditional[3] and two non-traditional[4] visualizations. The first visualization was created based on the bar chart technique, also known as column chart. The second was designed based on the bubble hierarchy technique. The third was created from two technics: the unit graph technique, and the line graph technique.

It is important to mention that we are not seeking to evaluate the effectiveness of the techniques used, but rather to understand which type of visualization teachers perceive as

[2] MeuTutor: available at http://meututor.com.br/
[3] Also known as "conventional" meaning something that is commonly used.
[4] Term used in this work as the opposite of "traditional/conventional" terms.

being more helpful in making them notice problems/demands within a group of students, allowing them to make the necessary pedagogical decisions based on that.

B. Research Questions

This article intends to aid teachers with their (pedagogical) decision-making. Considering this principle, we guided our research efforts, according to the following research questions: (RQ1) do teachers' characteristics influence their perception regarding the type of visualization (traditional or non-traditional)? If so, (RQ2) which characteristics influence teachers' perceptions? And, if so, (RQ3) how much do the characteristics influence teachers' perceptions?

C. Data Preparation for Creating the Questionnaire

In order to create the questionnaire, we used data from two hundred and thirty-one (231) students from *MeuTutor* (Section II.C). The data regarded the performance of students engaging with the learning resources available in *MeuTutor*. The dataset contained students from 3 different cities in the state of Alagoas, in the northeast of Brazil. We separated students into three groups, according to their location. Group I contained one hundred forty-seven (147) students; Group II contained fifty-five (55) students; Group III contained twenty-nine (29) students.

After that, we created three different visualizations, one for each group, being a traditional and two non-traditional visualizations, displaying the performance of the students, based on their interactions with *MeuTutor*'s learning resources. In order to measure the performance, we used the ratio between the number of questions answered correctly, by the total amount of questions answered. With the result, we categorized the students (from each group) into 3 subgroups and assigned them a color. Students with low performance (ratio from 0% to 50%) were colored red; students with performance about the average (ratio from 51% to 70%), were colored yellow; Students with superior performance (ratio from 71% to 100%), were colored green.

We evaluated the students' performance considering their interactions with the following learning resources: (1) the number of videos watched; (2) the number of questions answered correctly; (3) the number of questions answered incorrectly; and (4) the number of accesses made to the environment.

D. The Questionnaire

The questionnaire (Cronbach Alpha = 0.5) consisted of thirty-two questions, divided into seven blocks. The first block contained questions about the participants' (teachers) personal information (name, sex, age and the experience, in years, as a teacher). The second block contained questions about the teachers' education (regarding completion and the kind of institution they studied elementary school and high school). The third block contained questions about the teachers' training (if the participant attended or is attending a technical course, what kind of institution this course was/is offered, whether the participant is attending or completed higher education (college/university) and the institution this course was/is offered, and the name of the course(s). The fourth block identified if the participant is colorblind (directly asked, and tested with a colorblindness ID picture). Finally, the fifth, sixth and seventh blocks contained, respectively, the visualizations for groups I, II and III. In these last three blocks, the participants had to answer seven questions (five regarding the pedagogical situation displayed in the

TABLE I. VARIABLES – TEACHERS' PERSONAL CHARACTERISTICS

Information	Variable Name	Possible Values
Gender	Gender	0 (Male) or 1 (Female)
Age	Age	---
Experience as a teacher (time)	Less than 1 year - Exp1	1 (yes) or 0 (no)
	From 1 to 5 years - Exp2	1 (yes) or 0 (no)
	From 6 to 10 years - Exp3	1 (yes) or 0 (no)
Elementary School	Esc 1	1 (private institution) or 0 (public institution)
High School	Esc2	1 (private institution) or 0 (public institution)
Attending or completed technical course?	Fom1	1 (yes) or 0 (no)
Technical course in public institution?	Form2	1 (yes) or 0 (no)
Technical course in private institution?	Form3	1 (yes) or 0 (no)
Higher education (completed or in progress)?	Form4	1 (yes) or 0 (no)
Higher education in public institution?	Form5	1 (yes) or 0 (no)
Higher education in private institution?	Form6	1 (yes) or 0 (no)
Higher education in Exact sciences?	CurSup1	1 (yes) or 0 (no)
Higher education in the Humanities?	CurSup2	1 (yes) or 0 (no)
Higher education in Biological Sciences?	CurSup3	1 (yes) or 0 (no)
Are you color blind?	Dalt	1 (no) or 0 (yes)
Color blindness test	Dalt1	1 (passed) or 0 (failed)

visualization, and two regarding the teachers' perception on the helpfulness provided by the visualization[5].

E. Questionnaire Application

The questionnaire was applied in the *IV Brazilian Congress of Informatics in Education and X Latin American Objects and Learning Technologies*, held in Maceió, Alagoas - Brazil, from the 26th to the 30th of October, 2015. One thousand (1000) copies of the questionnaire were handled to the event's participants, under the authorization of the organizing staff inside the folders participants received in the accreditation.

During the five days of the event, we reminded participants to answer the questionnaire. In the final day of the event, we gave away (15) books and some small gifts, to those who handled the questionnaire back to us (answered).

F. Data Analysis

We collected a total of three hundred thirty-three (333) questionnaires. However, ninety-eight (98) participants were not teachers so, given the nature of our research, we could not consider their questionnaires, which left us with two hundred thirty-five (235) questionnaires. From this amount, fifty (50) were not completely answered and seventeen (17) had problems in one or more answers. The remaining 168 questionnaires were used for the statistical analysis.

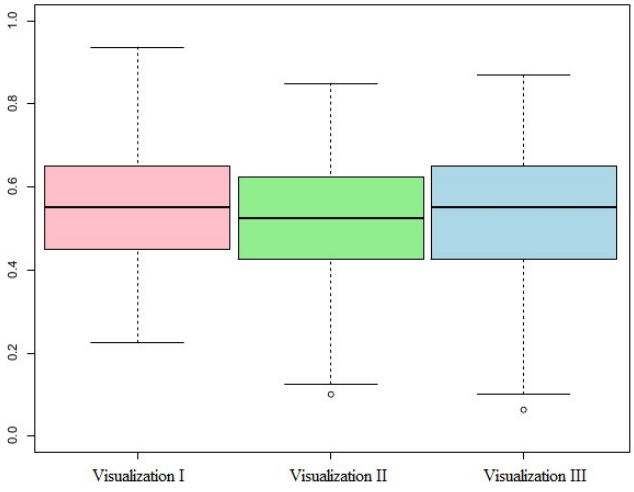

Fig. 1. Scores (correct answers) for each of the three visualizations.

During the analysis, we used all variables collected from the seven (7) blocks of the questionnaire. The variables related to the teachers' characteristics were coded with values zero (0) and one (1), except for the age, which we decided to leave it as informed by the teachers (according to Table I). Teachers with more than 10 years of experience (Exp4), were coded as follows: zero in variable Exp1 (these teachers did not have less than 1 year of experience), zero in Exp2 (they did not have from 1 to 5 years of experience) and zero in Exp3 (they did not have from 6 to 10 years of experience).

[5] The questionnaire is available at: http://www.ranilsonpaiva.com/www/publications/2016/fie/questionnaire.docx

We applied beta regression models to investigate the effect of these variables. According to Ferrari [5], the analysis with beta regression is indicated whenever it is possible to model a variety of uncertainties coming to a model with a degree of accuracy over the data variables analyzed. The intention was to overcome cold start, during a recommendation process, by modeling teachers' perception of helpfulness, based on their personal characteristics, for the types of visualization traditional and non-traditional, to visually assist pedagogical decision-making.

IV. Results and Discussion

Initially, we conducted a descriptive analysis of the response variables, which separately considered the correct answers of the participants (blocks 5, 6 and 7 of the questionnaire) for each of the three visualizations. Figure 1 the scores for each visualization (visualization 1 = left; visualization 2 = middle; visualization 3 = right).

Based on Figure 1, we notice the scores are very similar for all three visualizations. However, visualization I (traditional) shows a higher score (more correct answers regarding the pedagogical situation presented in the corresponding visualization) compared to visualizations II and III (non-traditional). This suggests traditional visualization provided more assistance compared to non-traditional visualizations. We could also notice that the two non-traditional visualizations had similar score between themselves (range and interquartile range), suggesting they provided equivalent assistance.

In blocks 5, 6 and 7 from the questionnaire, we asked the participants how helpful (question 6 – **Q6**) the visualization was, and how fast (question 7 – **Q7**) they understood the data displayed on them, for the pedagogical situation they presented. The answers were in a Likert scale (0 = very low; 25 = low; 50 = average; 75 = high; 100 = very high).

As we can see in Table II, for visualization I, 94 out of 168 participants (about 56%) considered **visualization I** helpful (above the average, i.e., High or Very High answers for Q6) and 88 participants (about 52%) were able to quickly understand (below the average, i.e., Very Low or Low answers for Q7) the situation displayed in it.

For **visualization II**, 72 out of 168 participants (about 43%) considered it helpful (above the average, i.e., High or Very High answers for Q6) and 81 participants (about 48%) were able to quickly understand (below the average, i.e., Very Low or Low answers for Q7) the situation displayed in it.

For **visualization III**, 77 out of 168 participants (about 46%) considered it helpful (above the average, i.e., High or Very High answers for Q6) and 80 participants (about 48%) were able to quickly understand (below the average, i.e., Very Low or Low answers for Q7) the situation displayed in it.

It is important to notice that the answers High and Very High, for question 6, were indications that the respective visualization was considered helpful by the participants. On the other hand, answers Very Low and Low, for question 7, were indications that the respective visualizations were understood quickly by the participants.

TABLE II. VARIABLES – TEACHERS' PERSONAL CHARACTERISTICS

	Visualization I		Visualization II		Visualization III	
	Q6	Q7	Q6	Q7	Q6	Q7
Very Low	25	11	21	13	22	19
Low	42	77	52	68	45	71
Average	15	35	23	45	24	43
High	70	35	57	37	61	30
Very High	16	10	15	5	16	5

Applying the beta regression model, we were able to associate the performance (correct answers regarding the pedagogical situation presented in each visualization), with the teachers' characteristics. After several attempts, we come to the following models for each type of visualization.

According to Figure 2, the predominant personal characteristics, for teachers that had a better score using visualization I (traditional), were:

- **Gender**: Male;
- **Experience**: More than ten years;
- **High School**: Yes | In a public institution;
- **Technical Training**: Yes | In a public institution;
- **Higher Education**: Yes | In a public institution;
- **Higher Education Field**: Exact sciences.
- **Color Blind**: No;

Model - Visualization I

Parameters	Variabel	Estimats	p_value
β_1	Constante	−1.61	0.000***
β_2	Gender	−0.13	0.011**
β_3	Exp4	0.09	0.093·
β_4	Esc2	0.13	0.024*
β_5	Form2	0.13	0.022*
β_6	Form5	0.18	0.002***
β_7	Dalt1	0.50	0.005***
β_8	CurSup3	0.26	0.016*

· significant to 10%. * significant to 5%. **signicant to 1*

***significant over 1%

Fig. 2. A model created based on visualization I.

According to Figure 3, the predominant personal characteristics, for teachers that had a better score using visualization II (non-traditional), were:

- **Gender**: Male;
- **Experience**: More than ten years;
- **Elementary School**: In a public institution;
- **High School**: In a private institution;
- **Technical Training**: Yes | ICT[6] | In a public institution;
- **Higher Education Field**: Humanities;
- **Color Blind**: No.

Model - Visualization II

Parameters	Variabel	Estimats	p_value
β_1	Constante	−1.39	0.000***
β_2	Gender	−0.15	0.011*
β_3	Esc1	−0.14	0.047*
β_4	Esc2	0.22	0.002***
β_5	Exp3	−0.08	0.290*
β_6	Exp4	0.17	0.007
β_7	Dalt1	0.58	0.005***

* significant to 5%. **signicant to 1*

***significant over 1%

Fig. 3. A model created based on visualization II.

According to Figure 4, the predominant personal characteristics, for teachers that had a better score using visualization III (non-traditional), were:

- **Gender**: Female;
- **Experience**: Less than one year;
- **Elementary School**: Yes | In a public institution;
- **High School**: Yes | In in a private institution;
- **Higher Education**: Yes | In a public institution;
- **Higher Education Field**: Exact sciences.

Model - Visualization III

Parameters	Variabel	Estimats	p_value
β_1	Constante	−1.73	0.000***
β_2	Gender	0.07	0.059·
β_3	Exp1	0.13	0.070·
β_4	Esc1	−0.10	0.037*
β_5	Esc2	0.09	0.046*
β_6	CurSup1	0.10	0.008***
β_7	Form6	−0.07	0.099·

· significant to 10%. * significant to 5%. **signicant to 1*

***significant over 1%

Fig. 4. A model created based on visualization III.

V. CONCLUSION

In this study we surveyed teachers regarding their perceptions for traditional and non-traditional data visualization, evaluating whether their profile (set of characteristics from a group of teachers) influenced, or not,

[6] ICT: Information and Communication Technology (field).

their choice for a traditional or non-traditional kind of visualization.

The results show that, discarding answers indicating average results, more favorable results confirming the helpfulness of both kinds of data visualization.

The analysis shows that the teachers' profiles influence the way they perceive the two types of visualization considered in this study, positively answering our first research question: (RQ1) do teachers' characteristics influence their perception regarding the type of visualization (traditional or non-traditional)? Yes, according to the models in Figure 2, 3 and 4.

Continuing the analysis, we wanted to evaluate: (RQ2) which of these characteristics influenced teachers' perceptions? We observed that all characteristics influenced the choice between traditional and non-traditional visualizations. However, the magnitude of the effects varied, and some characteristics were relevant for a particular kind of visualization, but not for another.

Finally, we wanted to know: (RQ3) how much do the characteristics influence teachers' perceptions? The answer is shown in Figures 2, 3 and 4, where the characteristics (Variable), their effect (Estimates) and the significance level of the result.

The participants had a higher number of correct answers when the traditional visualization was used to identify the pedagogical problems, compared to the two non-traditional visualizations. However, the difference was modest and some participants informed (personally) the non-traditional visualizations were useful and helped them detect and react to the pedagogical problem exposed (Table II).

We intend to use these findings to (1) provide teachers with visualizations that match their preferences and personal profile; (2) overcome the cold start barrier, requiring only some of the teachers' personal characteristics; (3) empower and improve pedagogical decision-making; (3) associate these results with methods from other studies, for example those that consider the cognitive capabilities in order to recommend different kinds of visualization.

We also intend, as a future work, evaluate the data using techniques other than the Beta Regression Model, in search for different patterns and/or trends.

Some of the limitations of this study include the amount and nature of the data asked from the participants may be considered "too personal". It is also possible to ask different personal information; in this case this study's method may be applied on the data. We classified data visualization into traditional and non-traditional visualizations, which may be considered an oversimplification.

VI. Acknowledgements

We would like to thank *MeuTutor*® team for providing us access to the data needed to plan, implement and test our ideas. We would like to thank the organizing team of the CBIE & LACLO 2015 for allowing us to run our experiment during the event. We also acknowledge the financial support provided by FAPEAL and CAPES / CNPq.

References

[1] BITTENCOURT, I.I. et al. A computational model for developing semantic web-based educational systems, Knowledge-Based Systems, v. 22, p. 302-315, 2009

[2] Conati, Cristina, and Heather Maclaren. "Exploring the role of individual differences in information visualization." Proceedings of the working conference on Advanced visual interfaces. ACM, 2008.

[3] Dietmar Jannach, Markus Zanker, Alexander Felfernig, and Gerhard Friedrich. Recommender systems: an introduction. Cambridge University Press, 2010.

[4] EAD CENSO. Br (2013). Relatório Analítico da Aprendizagem a Distância no Brasil, 2012.

[5] Ferrari, Silvia, and Francisco Cribari-Neto. "Beta regression for modelling rates and proportions." *Journal of Applied Statistics* 31.7 (2004): 799-815.

[6] I Elaine Allen and Jeff Seaman. Grade change: Tracking online education in the united states, 2013. Babson Survey Research Group and Quahog Research Group, LLC. Retrieved on,3(5):2014, 2014.

[7] Jitender S Deogun, Vijay V Raghavan, Amartya Sarkar, and Hayri Sever. Data mining:Research trends, challenges, and applications. Roughs Sets and Data Mining: Analysis of Imprecise Data, pages 9–45, 1997.

[8] KRUM, R. Cool infographics: Effective communication with data visualization and design.[S.l.]: JohnWiley & Sons, 2013.

[9] Michael Friendly and Daniel J Denis. Milestones in the history of thematic cartography, statisticalgraphics,anddatavisualization. URLhttp://www.datavis.ca/milestones,2001.

[10] Neto O., "Joint-de: Sistema de mapeamento objeto-ontologia com suporte a objetos desconectados", 2014.

[11] Toker, Dereck, et al. "Towards adaptive information visualization: on the influence of user characteristics." User Modeling, Adaptation, and Personalization. Springer Berlin Heidelberg, 2012. 274-285.

[12] Toker, Dereck, et al. "Individual user characteristics and information visualization: connecting the dots through eye tracking." proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, 2013.

[13] WARD, M. O.; GRINSTEIN, G.; KEIM, D. Interactive data visualization: foundations, techniques, and applications. [S.l.]: CRC Press, 2010.

[14] Alexander, Michael, Mico Yuk, and Stephanie Diamond. Data visualization for dummies. John Wiley & Sons, 2014.

[15] Friendly, Michael. "A brief history of data visualization." In Handbook of data visualization, pp. 15-56. Springer Berlin Heidelberg, 2008.

[16] Paiva, Ranilson O. Araújo, Ig Ibert Bittencourt, Alan Pedro da Silva, Seiji Isotani, and Patricia Jaques. "Improving pedagogical recommendations by classifying students according to their interactional behavior in a gamified learning environment." In Proceedings of the 30th Annual ACM Symposium on Applied Computing, pp. 233-238. ACM, 2015.