

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

Ranilson Paiva¹, Ig Ibert Bittencourt¹, Maria Cavalcante¹, Patrícia Ospina²

¹Computing Institute – Federal University of Alagoas (UFAL)
Maceió, AL – Brazil

²Computing and Systems Department – Federal University of Pernambuco (UFPE)
Recife, PE – Brazil

{ranilsonpaiva,ig.ibert}@ic.ufal.br, mariadasgracascsilva@gmail.com, patespipa@de.ufpe.br

Abstract. *From 2012 until 2016, the number of US students enrolled in an online course increased 14.68%, resulting in more work for online teachers, who are responsible for planning and making pedagogical decisions to guide students. Interactions in such courses can generate data (quantity and variety), where relevant information in the educational context can be extracted, assisting teachers managing their classes. However, to present these data in spreadsheets, tables and graphics, is not enough. 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 visualization, which we categorized in two broad groups: traditional or non-traditional. We evaluated how users reacted to these types of visualizations and what users' features are associated with their preferences for one category or the other. In this paper, we surveyed 235 teachers to evaluate how these two categories of visualizations affect the way participants evaluate data from an online course. They had to check the visualizations and identify which item contributed the most, and which item contributed the least to the performance of the students. The answers (correct or incorrect) were evaluated regarding the teachers': age, gender, experience, education and perception on the usefulness of each visualization. Our ultimate purpose was to create a model to recommend visualizations according to the teachers' profile.*

Keywords: *Data Visualization, Model Teachers' Perceptions, Traditional Visualizations, Non-Traditional Visualizations.*

1. Introduction

The increase in supply and demand for courses mediated by online learning environments is an evidence of a change in education [Chrysafiadi and Virvou 2013]. For instance, according to [Seaman et al. 2018], in 2012, there were 5,425,406 students enrolled in an online graduate or undergraduate course. In 2016, this number increased to 6,359,121 students; a 14.68% difference. This new paradigm is called AAA

Learning and it relies on information and communication technology (ICT) to (ideally) promote learning anywhere, anytime and to anyone [Bittencourt et al. 2009].

Courses in this new paradigm produce a vast amount and variety of data [Deogun et al. 1997, Paiva et al. 2015]. It is possible to extract relevant information from these data [Witten et al. 2016] to guide pedagogical decisions [Paiva and Bittencourt 2017b, Paiva and Bittencourt 2017a], but simply presenting the output to teachers, in the form of spreadsheets, tables or graphics, is not enough to allow them to understand this information. Thus, authors suggest helping people visualize the data and their meaning [Jannach et al. 2010]. For that, we can use Data Visualization, which is 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 [Krum 2013].

The goal of data visualization is to communicate information clearly and efficiently from the point of view of the users (in this study, teachers), helping them analyze and reason about the data and the evidence they contain [Krum 2013, Friendly 2008]. Nevertheless, people react differently to different types of data visualization. We categorized them into two broad types: (1) traditional or (2) non-traditional, and evaluated how they affected users and their capacity to identify the pedagogical problems displayed in them.

Studies investigate the possibility of creating visualizations considering the characteristics and skills of the users [Conati and Maclaren 2008, Toker et al. 2013]. Based on that, we surveyed 235 teachers, where 168 of them completed it. The 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 analyze the visualizations created with these data, and identify the item with higher impact on the students' performance, as well as the item with lower impact.

We evaluated their answers considering: age, gender, experience, academic background and perception on the helpfulness of each visualization. Our objective was to create a model in order to recommend the appropriate type of visualization, based on the teachers' profile (a set of teachers' characteristics). We also evaluated which characteristics could drive the users' preference for one of these categories of visualization.

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 for one of the two types of visualization, and teacher's profile. We also measured the effect that each characteristic had on these preferences.

2. Background

2.1. Data Visualization

Images have been used as a communication mechanism long before the formalization of written language [Ward et al. 2015], which indicates the importance of its role to communicate facts and ideas. According to [Ward et al. 2015], data visualization is 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 [Krum 2013], data visualization is a visual representation of a range of different numerical data, which allows **making decisions, comparing data and identifying 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.

2.2. Perception

According to Ward [Ward et al. 2015], “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 concludes Ward [Ward et al. 2015], 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.

3. The Survey

3.1. Data Visualization Selection

For the experiment (a survey using printed questionnaire), we chose three (3) different visual representations: one traditional¹ and two non-traditional² visualizations. The first visualization was created based on a bar chart. The second was designed based on the bubble hierarchy technique [Krum 2013, Alexander et al. 2014]. The third was created from merging two technics: the unit graph technique, and the line graph technique.

It is important to mention that we are seeking to evaluate/understand which type of visualization teachers perceive as being more helpful in making them notice issues within a group of students, allowing them to make the necessary decisions to help these students, based on the information displayed in the visualization.

3.2. 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 (RQ3) how much do each characteristic influence teachers’ perceptions?

¹Also known as “conventional” meaning something that is commonly used.

²Term used in this work as the opposite of “traditional/conventional”.

3.3. Data Preparation for Creating the Questionnaire

In order to create the questionnaire, we used data from two hundred and thirty-one (231) students from an online learning environment called MeuTutor³. The data contained the students' performance and some interactions with the learning resources available in MeuTutor. The dataset contained students from 3 different cities in the state of Alagoas (in the northeast region 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 one 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, divided by the total amount of questions answered⁴. With the result, we categorized the students into 3 subgroups and assigned them a color. Students with a low performance (ratio from 0% to 50%) were colored red; students with intermediate performance (ratio from 51% to 70%), were colored yellow; Students with a high performance (ratio from 71% to 100%), were colored green.

We correlated 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 (login) made to the environment.

3.4. The Questionnaire

The questionnaire (Cronbach Alpha = 0.5) consisted of thirty-two questions, divided into 7 sections. Section 1 contained personal questions about the participants: name, sex, age and the experience (in years, as a teacher). Section 2 contained questions about the teachers' education (completion and the kind of institution they studied during elementary and high school). Section 3 contained questions about the teachers' training (if they attended or were attending a technical course, what kind of institution the course was offered, whether they had a higher education degree. Section 4 checked if the participant was colorblind (directly asked, and tested with a colorblindness ID picture). Section 5,6 and 7 contained, respectively, the visualizations for groups I, II and III. In these last three sections, the participants had to answer 7 questions (5 regarding the pedagogical situation displayed in the visualization, and two regarding their perception on the helpfulness provided by the visualization⁵).

3.5. 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,

³Available at: <http://meututor.com.br/>

⁴We will refer to this ratio as the students' "performance".

⁵The questionnaire is available at: <https://goo.gl/LGpqot>

Information	Variable Name	Possible Values
Gender	Gender	0 (Male) or 1 (Female)
Age	Age	—
Experience (e) as a teacher (years)	$e < 1$ (Exp1)	1 (yes) or 0 (no)
	$1 \leq e < 5$ (Exp2)	1 (yes) or 0 (no)
	$5 \leq e < 10$ (Exp3)	1 (yes) or 0 (no)
	$e \geq 10$ (Exp4)	1 (yes) or 0 (no)
Elementary School	Esc1	1 (private School) or 0 (public School)
High School	Esc2	1 (private School) or 0 (public School)
Technical course (completed or in progress)?	Form1	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)
High education (completed or in progress)?	Form4	1 (yes) or 0 (no)
High education in public institution?	Form5	1 (yes) or 0 (no)
High education in private institution?	Form6	1 (yes) or 0 (no)
High education in Exact sciences?	CurSup1	1 (yes) or 0 (no)
High education in the Humanities?	CurSup2	1 (yes) or 0 (no)
High 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)

Table 1. Variables – Teachers’ Personal Characteristics.

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 team inside the folders participants received in the accreditation.

3.6. Data Analysis

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

During the analysis, we used all variables collected from the 7 sections 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 1). For instance, 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).

We applied beta regression models to investigate the effect of these variables.

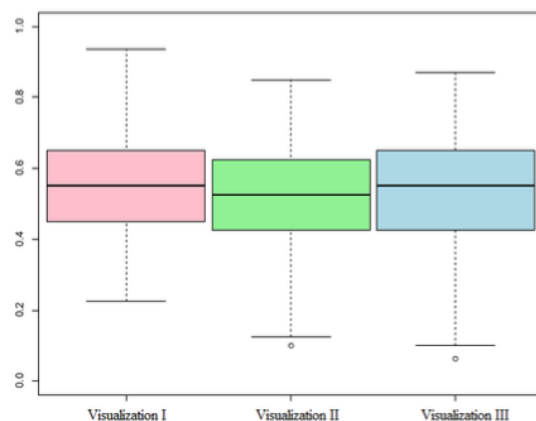


Figure 1. Score (correct answers in the questionnaire) for each visualization.

According to Ferrari [Ferrari and Cribari-Neto 2004], 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 creating a model of the teachers' perception of helpfulness for the types of visualization traditional and non-traditional, based on their personal characteristics. The objective is to visually assist teachers' (pedagogical) decision-making.

4. Results and Discussion

Initially, we conducted a descriptive analysis of the response variables, which separately considered the correct answers of the participants (sections 5, 6 and 7) for each of the three visualizations. Figure 1 shows the scores⁶ 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 scores (range and interquartile range), suggesting they provided equivalent assistance.

We asked the participants how helpful (question 6 - Q6) the visualizations were, and how fast (question 7 - Q7) they were able to understand the data visualized. The answers were in Likert scale (very low; low; average; high; very high). If a participant assigned High or Very High for Q6 and Q7, we considered the visualization was helpful and that they were able to quickly understand the information, respectively.

As seen in Table 2, 94 out of 168 participants (about 56%) considered **visualization I** helpful and 88 participants (about 52%) were able to quickly understand

⁶The percentage of correct answers made by the participants in the questionnaire.

	Viz 1		Viz 2		Viz 3	
	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

Table 2. Teachers' perceptions regarding the usefulness of the visualization (Q6) and if teachers could quickly understand the information displayed in the visualization (Q7).

the situation displayed in it. For **visualization II**, 72 out of 168 participants (about 43%) considered it helpful and 81 participants (about 48%) were able to quickly understand the situation displayed in it. For **visualization III**, 77 out of 168 participants (about 46%) considered it helpful and 80 participants (about 48%) were able to quickly understand the situation displayed in it.

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 Table 3, the predominant personal characteristics for teachers that had a better score using visualization I (traditional) were: (1) Gender: Male; (2) Experience: $e \geq 10$ years; (3) High School: Yes (public institution); (4) Technical Training: Yes (public institution); (5) Higher Education: Yes (public institution); (6) Higher Education Field: Exact sciences; (7) 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%. ** significant to 1%
*** significant over 1%

Table 3. Model for the teachers' perceptions regarding visualization 1.

According to Table 4, the predominant personal characteristics, for teachers that had a better score using visualization II (non-traditional), were: (1) Gender:

Male; (2) Experience: $e \geq 10$ years; (3) Elementary School: Yes (public institution); (4) High School: Yes (private institution); (5) Technical Training: Yes | ICT⁷ (public institution); (6) Higher Education Field: Humanities; (7) 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%. **significant to 1%

***significant over 1%

Table 4. Model for the teachers' perceptions regarding visualization 2.

According to Table 5, the predominant personal characteristics, for teachers that had a better score using visualization III (non-traditional), were: (1) Gender: Female; (2) Experience: $e < 1$ year; (3) Elementary School: Yes (public institution); (4) High School: Yes (private institution); (5) Higher Education: Yes (public institution); (6) 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%. **significant to 1%

***significant over 1%

Table 5. Model for the teachers' perceptions regarding visualization 3.

⁷ICT: Information and Communication Technology (field).

5. Conclusion

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

Regarding (RQ1) do teachers' characteristics influence their perception regarding the type of visualization (traditional or non-traditional)? The models in Tables 3, 4 and 5 suggest that these characteristics influence their perceptions.

Regarding (RQ2) which of these characteristics influenced teachers' perceptions? We observed that all characteristics influenced the choice between traditional and non-traditional visualizations, but the level of the effects varied. Some characteristics were relevant for a particular kind of visualization, but not for another.

Regarding (RQ3) how much do the characteristics influence teachers' perceptions? According to Tables 3, 4 and 5, the characteristics (Variable), their effect (Estimates) and the significance level of the result.

The participants who used the traditional visualization had higher scores, compared to those who used 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, but they felt more confident using a visualization that was "familiar".

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 to make the first recommendations; (3) empower and improve pedagogical decision-making; (4) associate these results with methods from other studies, for example those that consider the cognitive capabilities in order to recommend different kinds of visualization; (5) evaluate, as a future work, 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.

As future works, we intend to: (1) check how some characteristics of the visualization influence teachers interacting with them (like: the colors used, the shapes used, the size of these shapes, etc.); (2) Create and check the effectiveness of visualizing the output of data mining algorithms; (3) Check how data visualization supports teachers' pedagogical decision-making; (4) Check how data visualization helps reducing teachers' work load, and other works where we can use data visualization to help teachers in different contexts.

6. Acknowledgements

We would like to thank MeuTutor® team for providing us access to the data we needed to test our research. 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

- Alexander, M., Yuk, M., and Diamond, S. (2014). *Data Visualization for Dummies*. John Wiley & Sons.
- Bittencourt, I. I., Costa, E., Silva, M., and Soares, E. (2009). A computational model for developing semantic web-based educational systems. *Knowledge-Based Systems*, 22(4):302–315.
- Chrysafiadi, K. and Virvou, M. (2013). Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, 40(11):4715–4729.
- Conati, C. and Maclaren, H. (2008). Exploring the role of individual differences in information visualization. In *Proceedings of the working conference on Advanced visual interfaces*, pages 199–206. ACM.
- Deogun, J. S., Raghavan, V. V., Sarkar, A., and Sever, H. (1997). Data mining: Research trends, challenges, and applications. *Roughs Sets and Data Mining: Analysis of Imprecise Data*, pages 9–45.
- Ferrari, S. and Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics*, 31(7):799–815.
- Friendly, M. (2008). A brief history of data visualization. In *Handbook of data visualization*, pages 15–56. Springer.
- Jannach, D., Zanker, M., Felfernig, A., and Friedrich, G. (2010). *Recommender systems: an introduction*. Cambridge University Press.
- Krum, R. (2013). *Cool infographics: Effective communication with data visualization and design*. John Wiley & Sons.
- Paiva, R. and Bittencourt, I. I. (2017a). The authoring of pedagogical decisions informed by data, on the perspective of a mooc. In *Anais dos Workshops do Congresso Brasileiro de Informática na Educação*, volume 6, page 15.
- Paiva, R. and Bittencourt, I. I. (2017b). Helping mooc teachers do their job. In *Researcher Links Workshop: Higher Education for All*, pages 52–67. Springer.
- Paiva, R. O. A., Bittencourt, I. I., da Silva, A. P., Isotani, S., and Jaques, P. (2015). 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*, pages 233–238. ACM.
- Seaman, J. E., Allen, I. E., and Seaman, J. (2018). Grade increase: Tracking distance education in the united states. *Babson Survey Research Group*.
- Token, D., Conati, C., Steichen, B., and Carenini, G. (2013). Individual user characteristics and information visualization: connecting the dots through eye tracking. In *proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pages 295–304. ACM.
- Ward, M. O., Grinstein, G., and Keim, D. (2015). *Interactive data visualization: foundations, techniques, and applications*. AK Peters/CRC Press.
- Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. (2016). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann.