IV. EFFECTIVE DEVELOPMENT OF TEACHERS' SKILLS IN THE AREA OF ICT AND E-LEARNING

BIG DATA IN EDUCATION. ATTITUDES OF TEACHERS TOWARDS THE APPLICATION OF BIG DATA TECHNIQUES

Paloma Paniagua Martín & Sixto Cubo Delgado
University of Extremadura
paniaguamartinp@gmail.com, sixtocubo@gmail.com

Abstract: We live surrounded by huge amounts of data. We have been producing more information in the last two decades than in thousands of previous years of Humanity's existence. And the quantities multiply exponentially. Thanks to Big Data technologies companies we can predict our buying behavior, analyze our mood or establish new models of interaction between users. However, we have not yet been able to introduce all this technology into the educational world. Is this useful and viable? What do teachers think about it? The objective of this study is to know the opinions and attitudes that the University faculty and a group of teachers from different educational levels have towards the use of Big Data technologies in their classrooms.

Keywords: Big Data, Innovation, Teachers, Learning Analytics, Educational Technology.

1. INTRODUCTION

1.1 What is Big Data?

In 1998 the American John Mashley, theoretical computer scientist, published an article in which he predicted the immense wave of data that was coming. Subsequently, in 2013 the term "Big Data" was included in the Oxford dictionary and was defined as: "Extremely large data set that can be analysed computationally to reveal patterns, trends and associations, especially in relation to human behaviour and interactions".
Google Executive Director, Eric Schmidt said in a "Techonomy" conference that "we generate more information, in two days, than in our entire history until 2003". Such a volume of information represents an authentic revolution that can completely transform the world.

But... How do companies use Big Data? An IBM article called "The use of Big Data in the real world" (2012), talks about how the concept of "Big Data" has become a business priority, given its ability to profoundly influence trade at global scale. Companies that want to succeed and grow must adapt to these new models of complex algorithms and data analysis (Evans, 2015).

One of the most important uses of Big Data is the "monetization" of the data, (García, 2017). How many times have you searched a product in google and from that moment all the advertising does not stop chasing you? This concept is called "reorientation" and its objective is to remind users interested in a specific product that the product is there with an interesting offer for them (Abad, 2015).

1.1.2 What does Big Data know about you?

Currently, consumers surf the web leaving a trace on: who they are, what they are interested in, whom they are related to, where and when they buy, (Gázquez, 2016). All data is collected to become purchase patterns. In 2017, the newspaper "El Mundo" published an interview about Martin Hilbert, who said: "With 150 likes the automated learning algorithm can detect your personality, with 200 you know more than your partner and with 250 likes you can know more about yourself than yourself". The objective of Big Data is to convert the data into information, an information that facilitates the making of important decisions, even in real time (López, 2014).

However, all those apparent wonderful advantages can also work against us. It has created a huge business of companies that "sell" the data circulating on the networks, data that citizens have given and that, properly packaged and selected for what and for whom, are sold at the price of gold to other companies or institutions (Francese Valls, 2017). In exchange for browsing for free, we give away our data, free of charge, only for others to use our fingerprint. For example: Can the use of "Big Data" make a president win the elections? Remember the Donald Trump Election as President of the USA, the English Referendum favourable to Brexit, etc. Martin Hilbert (2017) stated that during the American presidential elections, one in five messages on Twitter or Facebook were false messages. For all this, and many other security problems, governments are being forced to regulate all this data transit, without authorization, from users, with specific laws for the use of Big Data. Two years ago, Europe created Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and the free movement of personal data. This data and repealing Directive 95/46/EC, General Data Protection Regulation, which is being applied since May 25, 2018 in the 28 countries of the European Union. In Spain, we have the Spanish Data Protection
Association (AEOD) that regulates the data protection law both in companies and in educational centres.

1.1.3 Big Data in education

Everything students do is made up of data (Kalota, 2015). What they learn and what they do not, the correct or incorrect exercises, the number of times they participate in class, the students that drop out, the frequency and causes, the scores, the times of the day in which most receivers are found, etc. And if we are able to interpret them correctly, the educational system can be enormously strengthened. In the United States, a pioneering movement of schools based on learning with technologies, called "AltSchool", has emerged. These schools collect data on what students do in their training centres and extract useful information that allows them to maximize their learning. Currently, it is necessary to create new methods based on technology to monitor students, predict academic risks or simply understand the behaviour of school groups, among many others benefits. Now we have methods such as: Adaptive learning, Flipped classroom, Combined learning, E-learning, Mobile learning, and platforms such as: Moodle, ClassDojo, Google Classroom ... among others (Salazar, J 2016).

1.1.4 What does Learning Analytics mean?

Students, as users of technological devices and social networks, are contributing, consciously or unconsciously, valuable data, the so-called "fingerprint", which after being analysed and processed statistically, serve to identify trends and predict behaviours. That data provides valuable information to design and customize the educational offer. This is called "learning analytics" and is the result of the application of Big Data to improve training and learning (Martínez, Raidell, & Duart, Josep M, 2016).

Education today has passed the transmission of knowledge, because it is very extensive and fully accessible thanks to internet, to another more creative education, based on preparing people to learn, to be autonomous in the process of access/selection of relevant information, and to adapt to the changing needs throughout life (Cobo and Moravec, 2011). The "Learning Analytics" is a technique that is still in its infancy but that provides the educator with a better vision of how students develop and progress, offers a personalized support, allows to evaluate which activities and exercises can be more effective, detect problems and above all look for solutions.

With this background, the research presented poses the following objectives and hypotheses:

1.2 Objectives

General objectives are as follows:

- Deepen the concept of Big Data and its applications.
• Know some opinions and attitudes towards the Big Data.

Specific objectives are as follows:

• Show the attitudes of university teachers and educational centres in Extremadura towards the use of new Big Data technologies.

• Analyse the acceptance and viability of the Big Data in education.

• Analyse the previous knowledge of the teaching staff towards the term Big Data.

1.3 Hypotheses

There were seven hypotheses stated in this research:

1. The faculty of University knows more about the concept "Big Data" than teachers of other educational levels.

2. The University faculty considers the "Big Data" technology more useful and viable than teachers of other educational levels.

3. Teachers, under the age of 40, think that it is more useful and viable than older teachers.

4. There are no differences of opinion between men and women about Big Data technology.

5. Teachers with more than 16 years of experience have a less favourable opinion about Big Data technology.

6. Teachers of higher levels (Secondary Education, Baccalaureate and University), consider that the use of Big Data in Education will bring situations with greater school success compared to the teachers of lower levels (Infant and Primary).

7. The specialists of Therapeutic Pedagogy (TP) and Hearing and Language (LA) consider that the use of Big Data technologies could reduce the difficulties in learning compared to the teachers of Early Childhood Education and Primary Education.

2. MATERIAL AND METHODS

2.1 Participants

The sample consisted of 119 teachers. The age range was from 22 to 65 years old from the University of Extremadura and from 6 educational centres of Extremadura. The type of sampling is of convenience, with the selection of centres according to their availability to participate in the study, with a similar percentage between the participating university faculty and the teaching staff of educational centres.
"Big Data" in Education. Attitudes of Teachers Towards the Application …

Table 1.

Analysis of the sample.

<table>
<thead>
<tr>
<th>GENDER</th>
<th>PROVINCE</th>
<th>EDUCATIONAL LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Men</td>
<td>Cáceres</td>
<td>TP/LA 7</td>
</tr>
<tr>
<td>Women</td>
<td>Badajoz</td>
<td>Pre-Primary Education 20</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>AGE</th>
<th>TEACHING EXPERIENCE</th>
<th>EDUCATIONAL LEVEL</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; 30 years old</td>
<td>0-5 years</td>
<td>Baccalaureate 6</td>
</tr>
<tr>
<td>From 30-40</td>
<td>6-15 years</td>
<td>University 45</td>
</tr>
<tr>
<td>From 41-50</td>
<td>16-25 years</td>
<td></td>
</tr>
<tr>
<td>&gt; 50 years old</td>
<td>26 or more</td>
<td></td>
</tr>
</tbody>
</table>

Source: Own work.

2.2 Instrument

After carrying out an extensive bibliographical review, no instrument was found that was related to the object of the research. Therefore, a questionnaire was designed to measure the attitudes of the teaching staff towards the Big Data and subsequently, 11 experts were asked and experts the validation of it. Once the evaluations were received, the items that had received a lower score by the validators were modified or eliminated. The validation occurred at the end of May 2018.

Secondly, after modifying the questionnaire, it was sent by "e-mail" in the month of June 2018, to the professors of the University of Extremadura and to teachers from 6 educational centres, obtaining a response rate with a volume of 119 subjects, which have been taken into account in the sample for the investigation. This questionnaire was answered anonymously, following the ethical values required in research with people, being informed of the objective of the study. The questionnaire consisted in the visualization of a video of 3 minutes, created expressly for the participants, in which was detailed what is the Big Data and what it is being used for (Link to the video: https://www.youtube.com/watch?v=uacXhStmu1Q).

Next, there were 5 sociographic questions and 10 Likert-type questions about the opinion, the usefulness, the viability of Big Data in Education and, finally, a last section of comments. The reliability and internal consistency was ascertained by the Cronbach alpha statistic whose results show high reliability, with a resulting data of 0.86.
3. ANALYSIS AND RESULTS

3.1. Descriptive analysis

To carry out the descriptive analysis, graphs of all the items of the questionnaire were made. In the following graphs we can see the results obtained in each item of the questionnaire:

Figure 1. Graphs of the items Sex, Age and Province

*Source: Own work*
Figure 2. Graphs of the items Experience, educational level, question 1, 2, 3 and 4

Source: Own work
The most significant results of these graphs are the following:

In the first question "Did you know what Big Data is before viewing the video?" Most of the respondents did not know this concept. However, that is the only graph whose values have been more negative. The rest of the graphs are quite similar, the scores begin to rise, from 3 to 5 on the Likert scale, coinciding with positive
opinions towards the use and implementation of "Big Data" techniques. For example, we can find high scores in the opinion that Big Data could facilitate the performance as teachers, help create personalized training and relevant content, reduce difficulties in student learning, reduce abandonment school and boost school success, and in general, the majority of respondents would like to be able to apply Big Data techniques in their teaching activity.

On the other hand, they were also asked about the educational levels they considered most viable for the application of Big Data techniques and the results showed, with higher scores, that the most optimal educational levels to use the Big Data would be: Primary Education, Secondary Education, Baccalaureate and University.

3.2. Inferential Analysis

For the inferential analysis, the same procedure was followed for each of the 7 hypotheses formulated. First, a graph was made about the variables involved in each hypothesis and then, when the variables were quantitative, continuous and at least interval, K-S, Rachas and Levene tests were applied to check if the data allowed to apply parametric tests. Once the values were obtained and the type of test that was applied was decided, the tests were carried out. The tests were T-Student for independent samples in case of parametric tests and Mann-Whitney U for nonparametric tests.

**Table 2. Inferential analysis of hypotheses.**

<table>
<thead>
<tr>
<th>HYPOTHESIS</th>
<th>MODEL</th>
<th>DEGREE OF SIGNIFICANCE</th>
<th>RESULT</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. University faculty knows more about the Big Data concept than teachers from other educational levels</td>
<td>U de Mann-Whitney</td>
<td>0,000</td>
<td>H₀ is rejected (p &lt;0.05), thus accepting the working hypothesis. There are statistically significant differences. The faculty of University knows more the concept Big Data than the professors of other educational levels.</td>
</tr>
<tr>
<td>2. University teachers consider the Big Data technology more useful and viable than the teachers of other educational levels.</td>
<td>U de Mann-Whitney Utility: 0,095 Viability: 0,047</td>
<td>&quot;We accept the H₀ (p &gt; 0.05) for the utility analysis and we reject the H₀ (p &lt; 0.05) for the feasibility analysis, that is, there are no statistically significant differences for&quot; utility &quot;but if they exist in the case of &quot;viability&quot;.</td>
<td></td>
</tr>
</tbody>
</table>
3. Teachers under the age of 40 think that it is more useful and viable than professors older than 40 years.

<table>
<thead>
<tr>
<th>HYPOTHESIS</th>
<th>MODEL</th>
<th>DEGREE OF SIGNIFICANCE</th>
<th>RESULT</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Teachers under the age of 40 think that it is more useful and viable</td>
<td>U de Mann-Whitney</td>
<td>Utility: 0.289</td>
<td>We accept the $H_0$ ($p &gt; 0.05$), so no statistically significant</td>
</tr>
<tr>
<td>than professors older than 40 years.</td>
<td>Viability: 0.374</td>
<td></td>
<td>differences are observed between the different ages and what they</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>think about the usefulness and viability of the Big Data.</td>
</tr>
<tr>
<td>4. There are no differences of opinion between men and women about Big</td>
<td>T-Student for</td>
<td>0.002</td>
<td>$H_0$ is rejected ($p &lt; 0.05$), that is, there are statistically</td>
</tr>
<tr>
<td>Data technology.</td>
<td>Independent samples</td>
<td></td>
<td>significant differences between what men and women think about the</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>use of Big Data technologies.</td>
</tr>
<tr>
<td>5. Teachers with less than 16 years of experience have a more favorable</td>
<td>T-Student for</td>
<td>0.535</td>
<td>The $H_0$ is accepted ($p &gt; 0.05$), so no statistically significant</td>
</tr>
<tr>
<td>opinion about Big Data technology.</td>
<td>Independent samples</td>
<td></td>
<td>differences are observed between the years of experience and the</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>teachers’ opinion towards the use of Big Data technologies.</td>
</tr>
<tr>
<td>6. Teachers of higher levels (Secondary Education, Baccalaureate and</td>
<td>U de Mann-Whitney</td>
<td>0.094</td>
<td>The $H_0$ is accepted ($p &gt; 0.05$). Therefore, it is observed that</td>
</tr>
<tr>
<td>University), consider that the use of Big Data in Education will bring</td>
<td></td>
<td></td>
<td>there are no statistically significant differences between the</td>
</tr>
<tr>
<td>situations with greater school success compared to teachers at lower</td>
<td></td>
<td></td>
<td>educational level and the belief that the use of Big Data in</td>
</tr>
<tr>
<td>levels (Infant and Primary).</td>
<td></td>
<td></td>
<td>Education will bring about situations with greater school success.</td>
</tr>
<tr>
<td>7. The specialists of Therapeutic Pedagogy (TP) and Hearing and Language</td>
<td>U de Mann-Whitney</td>
<td>0.587</td>
<td>The $H_0$ is accepted ($p &gt; 0.05$). Therefore, it is observed that</td>
</tr>
<tr>
<td>(LA) consider that the use of Big Data technologies could reduce the</td>
<td></td>
<td></td>
<td>there are no statistically significant differences between the</td>
</tr>
<tr>
<td>difficulties in learning compared</td>
<td></td>
<td></td>
<td>educational level and the belief that the use of Big Data in</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Education will bring about reductions in learning</td>
</tr>
</tbody>
</table>
3.3. Analysis of frequencies of the comments of the participants

The survey included a section dedicated to comments for all Education professionals who had something to object about the study, whether they were opinions or suggestions, etc. Of 119 subjects surveyed, 61 answered this section with very different comments. For this analysis, a frequency chart of the comments of the participants has been designed. These comments have been grouped into 4 categories: The use of Big Data techniques seems to me ... very interesting, interesting, uninteresting and not at all interesting. In the following figure 4 you can see the results.

Figure 4. Graph of the frequencies of the comments of the participants

Source: Own work
CONCLUSIONS

In accordance with the proposed objectives, we can deduce the following conclusions:

In the first place, after an extensive literature review it can be affirmed that the concept of “Big Data” has been deepened and its applications at a general level and more specifically at an educational level.

Regarding the concept that teachers have towards the use of Big Data techniques, we can affirm that, thanks to this study, it has been possible to know the attitudes of University professors and Educational Centres in a more concrete way.

In third place, after analysing data of the results obtained in the questionnaire, we can conclude that the acceptance and feasibility of Big Data in education has been analysed. After the results obtained in the graphs we can affirm that in general there is a great acceptance and that the majority of the participating professors considers it viable.

And finally, after the study, it has been possible to analyse the previous knowledge of the teaching staff towards the term Big Data, which was not known by the majority of the teaching staff.

Regarding the hypothesis that were raised at the beginning of the investigation, it can be concluded that:

The first hypothesis suggested that the faculty of the University knew more about the concept of Big Data than the teachers of other educational levels, after the analysis of the data obtained through the questionnaire, this hypothesis can be confirmed since the most favourable results in this question could be related to the greater investigative spirit that is supposed to the faculty of the University and to be more up to date in the new tendencies.

The second hypothesis suggested that university faculty considered Big Data technology more useful and viable than faculty from other educational levels. However, this hypothesis has been partially accepted, since the faculty of the University sees it as more viable than the faculty of other educational levels, but, on the other hand, it sees it as less useful than the faculty of other educational levels. No studies have been found to support this result, so it could be inferred that it may be due to the variability of the sample.

The third hypothesis proposes that age is a determining factor for the conception of Big Data as useful and viable. This hypothesis suggested that teachers under the age of 40 think that it is more useful and viable than older teachers. However, after the analysis of the data, it can be stressed that, in this case, there are no significant differences so that age is not an influential factor to determine a more or less positive conception, the utility and the viability of the Big Data. This statistical equality in the criterion towards the usefulness and viability of the Big Data would be determined by the scarce knowledge on the subject and the lack of information
on the benefits of its application in the classrooms. However, in both cases, they consider that this issue, as a technological renovation, will end up being necessary and inexorably implemented in the classrooms. On the other hand, the sample used contemplated a much higher percentage of people over 40 years compared to those less than 40 years.

The fourth hypothesis proposed that there would be no differences of opinion between men and women about "Big Data" technology. However, it can be affirmed that there are differences of opinion regarding gender. According to the graph, women are slightly more optimistic. This may be due to the fact that in the questionnaire the percentage of women who participated is higher than that of men, as it is evident in the current education system, where female teachers far outnumber the number of men, as reflected in the study carried out by the Junta of Extremadura in 2008 (Training Guide for the incorporation of equality in public administration).

The fifth hypothesis proposed that teachers with more than 16 years of experience would have a less favourable opinion about Big Data technology. However, it has been verified that, since there were no differences of opinion regarding the Big Data, the years of experience have not been a determining factor in this case. As it happened in hypothesis 3, the experience, normally a consequence of age, does not show a special predilection towards the uses of Big Data, which may be due to the variability of the sample and the insufficient information that teachers have yet on the subject.

The sixth hypothesis suggested that teachers at higher levels (Secondary Education, Baccalaureate and University) would consider the use of Big Data in Education as a tool that would bring situations with greater school success compared to the teachers at lower levels (Children and Primary). In this case, after analysing the data, this hypothesis is rejected. It could be inferred that the information that gives us the use of Big Data is currently much higher in higher education and university levels since higher level students have greater access to computers, social networks, smartphones, video games, etc. and all those "Actions" give us a large volume of data and behavioural reflexes. Therefore, the rejection of this hypothesis could be due to the variability of the sample used.

Finally, the seventh hypothesis stated that the specialists of Therapeutic Pedagogy (PT) and Hearing and Language (LA) considered that the use of Big Data technologies could reduce the learning difficulties, in comparison to the teaching staff of Education Infant and Primary Education. However, this hypothesis is rejected as well, since, according to the data, it is the teachers of Early Childhood and Primary Education who turn out to be more optimistic in the use of Big Data technologies. Contrary to the result of this hypothesis, it is considered that one of the great benefits that Big Data can bring to the education system is in favour of the personalization of learning. Therefore, this particular result may reflect poor information on the subject.
Regarding the analysis of the comments that the participants wrote in the questionnaire, it could be said, on the one hand, that more than 66% of those 61 subjects considered that it is a very interesting, innovative topic and it is worth putting it into practice; however, others were more reluctant, (the remaining 34%) considering the implementation of training plans for teachers in this matter of vital importance.

Other teachers thought that the educational system should be reformulated to use methodologies based on Big Data and, finally, a smaller group of teachers said that Big Data has many disadvantages, considering this type of education as impersonal and saying that it loses the personal discovery factor.

However, there is great uncertainty and in some cases fear of the use of Big Data in the educational context, since some subjects considered that managing a personalized education for all students of an educational centre is practically impossible and would require a significant increase in specialized teaching staff. In any case, the majority thought that it is a very interesting topic but for this it is necessary to obtain more information about its uses and possibilities and to study in depth the ethical aspect and the protection of data before putting it into school practice.

To conclude with this study, we can point out that there is a long way to go with the Big Data. It is necessary to continue researching in this field and promoting educational plans, new methodologies, teacher training and new resources and tools that allow teachers to know and implement Big Data techniques in their classes in order to boost school success and personalized education. Therefore, we can suggest various lines of research in the future.

The first future line of research would be to replicate this study with a more representative sample. On the other hand, it would also be convenient to replicate the research incorporating the qualitative methodological approach, which would provide a more holistic perspective on teachers' attitudes towards the use of Big Data technologies. And finally, carry out an intervention project on the Big Data in education, to put into practice what in theory this concept proposes and contributes, verifying if the results and attitudes of the students are improved, or not, significantly, as well as the teaching work.

REFERENCES


