Affective exhibition during the interpretation of statistical data

Tamires Queiroz, Carlos Monteiro, Liliane Carvalho, Karen François

Affective exhibition during the interpretation of statistical data

Tamires Queiroz[1], Carlos Monteiro[1], Liliane Carvalho[1] and Karen François[2]

1 Federal University of Pernambuco, Recife, Brazil, tamires_nq@hotmail.com, cefmonteiro@gmail.com, lmtlcarvalho@gmail.com

2 Free University of Brussels – VUB, Brussels, Belgium, karen.francois@vub.ac.be

In recent years, research on the processes of teaching and learning of Statistics emphasises that the interpretation of data is a complex process that involves cognitive and technical aspects. However, the interpretation of statistical data is a human activity that also involves affective aspects. However these aspects were not sufficiently investigated. This paper discusses some elements from an empirical pilot study that explores the idea of affective expression during the interpretation of statistical data by undergraduate students of statistics and pedagogy. Although the participants had different academic background their interpretations of media statistics data did not follow patterns that were influenced by the university courses they attended.

Keywords: Affective exhibition, interpretation of statistical data, statistics education.

INTRODUCTION

Several studies on statistics education have given evidence that the interpretation of data is a complex activity which is associated with several aspects. Gal (2002) argues that *statistical literacy* consists of two types of components: cognitive and dispositional. The cognitive components, such as mathematical and statistical knowledge, are related to reasoning aspects that enable readers to technically interpret data. The dispositional components refer to more subjective elements related, for example, to beliefs, attitudes and emotional aspects of the individual that is reading the data. Based on Gal´s perspective, readers can discuss and communicate their conclusions about statistical data using their expertise on specific area, as well as their personal and social experiences.

Arcavi (2003) emphasises that the interpretation of (statistical) data is not the result of a simple decoding of graph components. Although Arcavi is mainly working on mathematics learning he also refers to allied disciplines (e.g., data handling or statistics) and the way of data representations and date graphing (Arcavi, 2003, p. 217).

Since "we live in a world where information is transmitted mostly in visual wrappings, and technologies support and encourage communication which is essentially visual" (Arcavi, 2003, p. 215) human being is encouraged to interpret visual representations. He emphasises the complexity of the phenomenon of interpreting visual representation, e.g. graphical display. It is not only related to what comes "within sight", but we are also encouraged and aspire to 'see' what we are unable to see. Referring to Goethe he explains the quote "We don't know what we see, we see what we know" (Arcavi, 2003, p. 230) stressing the last part of the expression "We see what we know". Arcavi argues that the same visual objects may have different meanings in different contexts. He therefore proposes to classify three types of 'difficulties', see cognitive, sociological and cultural.

Lima (1998) and Monteiro (1998) suggest that during the interpretation of data, the way readers use their mathematical and/or statistical knowledge is a complex aspect which is not the result from only one aspect, such as their academic background. Lima (1998) analyses the interpretation of data developed by designers and mathematics teachers. The author concluded that the interpretations of the participants were different in the way they read the data, however both groups were similar in the use of mathematical knowledge during the interpretations. Monteiro (1998) investigates the processes of interpreting graphs of printed media by a group of businessmen with different academic backgrounds and a group of economists. The author did not identified differ-

ences in relation to the strategies of problem solving, although the group of economists tended to produce more estimates in their interpretations.

Monteiro (2005) develops the idea of *critical sense* during the process of interpretation of statistics media graphs that is related to mobilisations and balance of several elements. The term mobilisation (Monteiro & Ainley, 2003) is related to the possibility of re-using or re-sourcing (Adler, 2000) previous knowledge and experiences during the process of interpretations of media graphs. This mobilisation seems to be a process in which readers explore the data, confronting it with their own perspective, and their previous experiences related to the data interpreted. However, the process of mobilisation in interpretation of media graphs does not 'naturally' happen. In order to mobilise their previous knowledge and experiences to interpret a media graph, readers need to establish a certain level of engagement in the task, which then leads to the articulation in which they make a recontextualisation of the knowledge and experiences mobilised, comparing them to the data. The reader also needs to balance different elements. Therefore, there is no direct application of knowledge and experiences for the process of interpretation. This complex process of mobilise and balance different elements during the interpretation of statistical data displayed in the graph is called *critical sense*.

McLeod (1992) did an extensive review of the literature in mathematics education that addresses the affective domain. The author states that, among other factors, most of these studies did not impact mathematics education because they were focused only on stable aspects of affectivity. In other words, these studies were more concerned with the products and not with the processes involved. Most studies reviewed by McLeod seem to conceive narrowly affectivity, merely investigate more specific concepts such as beliefs and attitudes. In addition, the studies did not make a link between affective and cognitive factors. Therefore, similar to those studies that investigated only cognitive aspects, the studies that addressed only affective aspects seem to have no impact on the learning and teaching of mathematics curriculum content. McLeod suggests that beyond a deepening of theoretical questions about the definition of affectivity and their relation to cognition, studies need to be based on research approaches that combine quantitative and qualitative methods.

Based on this literature review we can conclude that Gal (2002) provided statistical literacy principles comprised of cognitive and dispositional elements. Arcavi (2003) emphasised the importance of people's social and cultural background in the interpretation of data, which is also exemplified by Lima (1998) and Monteiro (1998). Finally, Monteiro (2005) added the notion of 'critical sense' based on statistical literacy principles, however emphasising the complex interrelation between components and processes involved. Research on the specific affective influences in statistics interpretation is still lacking. Looking at the field of mathematical literacy, affectivity has the rather narrow meaning of attitudes and beliefs. With the concept of affectivity we will include the dispositional elements that are take part at the interpretation of statistical data.

In the field of statistics education, it is necessary to develop a similar literature review that investigates the number of studies focused on the affective aspects, and how they can make impact on the teaching and learning of school statistics. Besides we need empirical evidence based on quantitative and qualitative approaches to better understand the interpretation of statistical data.

RESEARCH QUESTION

The research question in this exploratory study is if the affective expressions of students in situations of interpretation of statistical data are related to the students' background in the teaching and learning of statistics. Therefore we investigated the interpretations of students from two different backgrounds in the teaching and learning of statistics, viz. (i) bachelor in statistics and (ii) degree in pedagogy. We expected that the differences related to the type of course in which the participants were enrolled might influence their affective expressions on their interpretations.

METHODOLOGY

This pilot study was a qualitative investigation based on standardized open-ended interviews. In order to investigate aspects about the affective expression and statistics literacy, we invited first year students from two different university courses from the same Brazilian Federal University: two undergraduate students from an education course (P1 and P2) and two students from a statistics bachelor course (S1 and S2).

We chose a group of students from education because they are pre-service primary school teachers, and also because in this course has disciplines, such as Psychology which addresses affective aspects. The first year education students already attended those disciplines. The choice for students from bachelor in statistics was because this course has a curriculum focused on disciplines such as probability, data analysis, and stochastic phenomena. The first year students already had such disciplines and they will not have any discipline that approaches the affective aspects as part of their course. The first author contacted the students in their classroom and explained the research. The students interviewed were volunteers. Data collection was conducted in November and December 2013. The Pedagogy students P1 (57 years old, female) and P2 (33 years old, male) completed two disciplines that approaches affective aspects associated with teaching and learning. They did not attend any statistics course during their first year. The statistics students S1 (17 years old, female) and S2 (18 years old, female) completed five specific disciplines related to statistics.

DATA COLLECTION

The data collection was developed from individual standardized open-ended interviews which were composed by four tasks related to statistics data from different publications. Statistical data had the following topics: (1) on mammography examination, (2) on traffic accidents, (3) on life and health insurance, and (4) on high school students´ handguns. (Due to lack of space in this paper, we report on 2 and 4).The main reason to choose these cases was that the topics had a certain level of polemic. We expected that this type of data would motivate participants to make comments related to technical aspects as well as to their emotional reaction to the data. During the interviews, each task was presented to the student on a printed sheet. The researcher read the specific questions for each task. Interviews were recorded, and transcribed. The protocols were originally in Portuguese, fragments for this paper were translated by the authors.

Task 2 was comprised of questions about a line graph (Figure 1) that shows the percentages of deaths among Brazilian youth population caused by traffic accidents between 1998 and 2008. This graph was originally published in a report from the Brazilian ministry of justice (Waiselfisz, 2011).

Interview questions related to task 2 (Figure 1) were:

1) What can you conclude from the results presented in this graph?

2) If you could ask a question to whoever built this graph, would you do? Which one(s)?

3) What would you say? What elements which would you emphasise?

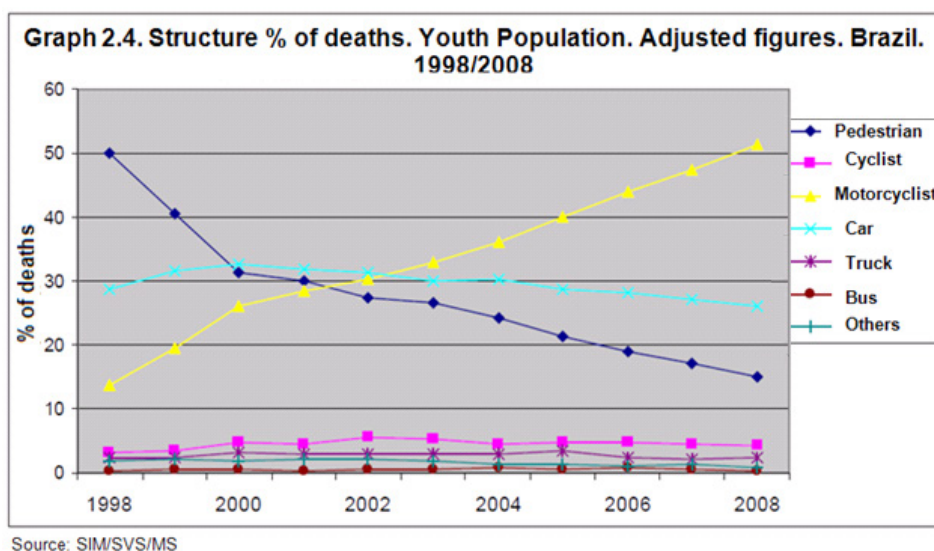4) What do you think these data are between 2008 and 2013? Why do you think that?

Figure 1: The percentages of deaths among youth Brazilian from 1998 until 2008

Task 4 was used by Watson (2005). It is a fragment of media news about the use of handguns by high school students:

> About 6 in 10 United States high school students say they could get a handgun if they wanted one, a third of them within an hour, a survey shows. The poll of 2,508 junior and senior high school students in Chicago also found 15% had actually carried a handgun within the past 30 days, with 4% taking one to school.

Interview questions related to task 4 were

1) What can you conclude about this research?

2) What do you think about the data collected in this research?

3) Do you have any questions or comments about this research? Which one(s)?

THE ANALYSES OF PARTICIPANTS' RESPONSES

The transcriptions of audio records were organized in protocols which were base to the data analyses. The participants' responses were analysed by a categorisation which emerged from a qualitative approach. Initially, we categorised their responses for each task and item. In order to discuss aspects from the data collected in the following sections we exemplify with some extracts from the interviews.

TASK 2 – DEATHS AMONG BRAZILIAN YOUTH POPULATION

In order to answer the first item of this task (What can you conclude from the results presented in this graph?) all participants made initially general comments related to the graph itself. However, they also made some observation which could make explicit their personal reaction to the data displayed. The extract from the interview with P1 can exemplify this:

> S1: Here, the person [who reads the graph] realizes that motorcyclists [figures] rose so much. The pedestrians, I think, they became more conscious in relation to that, before used to be something that nobody cared. Nowadays have footpath, have road sign, have zebra cross-

ing, have more... much care. While, the motorcyclists [figures]... is getting worst day by day. While the cars has a decrease but it is not much. While the ones [figures] of pedestrians remain all in one pattern, a little less.

The second question (If you could ask a question to whoever built this graph, would you do? Which one(s)?) motivated P2, S1 and S2 to ask questions related to the survey itself. They asked questions about the data collection, sampling and where the survey was conducted. Those questions indicated concern about the reliability of the research, e.g. S2 asked how the data were collected, as illustrated below:

> Researcher: If you could ask a question to whoever built this graph, would you do?
> S2: I would do... How the data was collected ... Only.
> Researcher: What do mean?
> S2: Kind of... If it was a street survey... Asking how many accidents you have in one year or if it was a survey... For example... Within the IML [Institute of Legal Medicine] which is the agency responsible for deaths.

The questions formulated by participants suggested that they were mainly concerned about technical aspects related to the data. Maybe they needed this type of information to be more comfortable to make other comments about the data. In some way, we could say that their questions were also expressions of their feelings about the data, because they made explicit some scepticism about the data displayed on the graph.

The responses to question 3 (What would you say? What elements which would you emphasize?) were complementary to those of question 2. The questions motivated the students to make more observations related to the data displayed. Generally, the participants now responded as a rereading of the graph. They pointed to specific figures (as they did when answering question 1). The most frequent responses were related to the increase in deaths of motorcyclists, followed by references to the decrease in pedestrians' deaths. P1 developed a more extensive response, questioning more explicitly, as illustrated in the following extract of her interview:

Researcher: Ok. It is... if you were going to comment to someone about this graph, which points would you think that should be important to discuss about it? What are the points that you would emphasise, that you would discuss more about?

P1: I think that it would be really to make relations between these categories... of pedestrians, motorcyclists and drivers. Why these categories are most affected in the traffic? Look at the difference of others, of the trucks, of the cyclists... Cyclists also do not have an incentive, don't they?... for the use of cycleway. But, it would be to make relation between these categories.

We can observe that P1 emphasises the importance of analysing the graph as a whole, as P1 attributes importance to the relationships between the categories. P1 was interpreting beyond the data displayed when she was referring to specific issues which are part of her daily experiences and observations (e.g. when she talked about the situation of cyclists). P1 seems to be preoccupied with the effects on people rather than the performance effect of the figures.

From the analyses of protocols we identified more variety of participants' responses to the fourth question (What do you think these data are between 2008 and 2013? Why do you think that?). All participants justified their answers based on *contextual reference* (Monteiro, 2005), that is when they contextualise the data displayed on the graph making references which are related to their formal knowledge in different areas and their opinion.

On the one hand, P2 and S2 justified their answers talking about a possible increase in numbers of accidents based on information that they had from two media reports.

Researcher: In the case, these results were from 1998 until 2008, right? What do you think these data are between 2008 and 2013?

S2: I think that tends to increase the number of deaths of cyclists [but really referring to motorcyclists], well... for my

knowledge of the world, right? Because... kind... the rates of IPVA [Brazilian tax on the ownership of motor vehicles], these things have decreased and more people are buying cars and... by this graph, it tends to increase even [the number of accidents].

On the other hand, P1 and S1 were more positive, they referred to the effectiveness of dry law that prohibit people to drive after drink any amount of alcoholic. The following extract gives an example of this type of approach:

Researcher: In this case, this... This report was made from 1998 to 2008. What do you think these data are between 2008 and 2013?

S1: Guy, I think it must have dropped. Slowly, but it is dropping. Our conscience is more... we have the dry law then we are taking more care, aren´t we? I believe that is a little bit better than two years ago.

From our analysis of the question 4 responses, we can infer that these participants also expressed different feelings about the same data displayed. These different affective expressions certainly are related to individual aspects from those who interpret the data. Hence there is also evidence that the interpretation of statistical data is composed of affective elements which need to be considered.

TASK 4 – THE USE OF HANDGUNS BY HIGH SCHOOL STUDENTS

Most of participants' responses to the first question (What can you conclude about this research?) tended to express feelings about the survey. The following extract from P1 interview is an example:

P1: I was shocked with the facility of armaments and weapons that the American population has. Because it is not just at school, no, any citizen, isn't it? Have in their home one or two weapons. ...The poll of 2,508 junior and senior high school students of the first and the last year of high school students... ...said that 15% of them had a gun in the last 30 days,

with 4% having taken to school... [reading the task]. From these 15%, 4 [%] led to the school, didn't they? Other day I was commenting on that... so... some of them have their jets, their imported cars, their helicopter; others live picking up litter, don't they? And, for the world to arm itself is also very fast. It's just you do your atomic bomb, to make intimidations with each other. It's shocking. I find shocking.

P1 tended to interpret the news report basing on an emotional reaction to the theme, although she considered the statistical figures related to the poll. S1 and S2 interpretations were quite similar to P1 when they responded this question. We can infer that the theme related to this task was more explicitly polemic, and may have influenced these participants' interpretations. Only P2 had a more descriptive reading of the data which did not seem to have had any personal reaction about the news report. As we can see the excerpt from his interview:

P2: I conclude that... it was done a piece of research, a poll... that from 2.508 high school students from the United States... it was detected that 60% of these students say they could get a gun if they wanted. Then, moreover, this 60%... it is... a third of these students could get a gun within an hour, and moreover, it... yet... students from first and last year of high school [rereading the task]... that means... it has here... at first, the research was done with high school students. Making a correction! And... this research has reached that percentage of 6 out of 10 could have access. After that, it was done a research investigating this number of students 2.508, of the final years of high school and among these ... it is... it was found that 15[%] of them carried a handgun... carried a handgun in the last 30 days and 4% of them had already led gun to school.

In this part of the interview, P2 was trying to understand the details about the procedures related to the data collection, and other details of the news data. His

concerns to specific aspect of the news did not allow him to question or expose his point of view.

The second question of task 4 intended to explore more specifically the participants' interpretation about statistical data involved (What do you think about the data collected in this research?). The participants' responses seemed to be complementary to their comments on the first question. Therefore, P1, S1 and S2 who did not make observations about the figures, responded here by making observations concerning quantitative data as well as expressing the personal point of view, e.g. the following extract form the interview:

Researcher: What do you think about the data collected in this research?

S1: I found it very serious. Because if among... 2,500 students, 15% of them have revolver, that is too many... if 4% can take it to school, imagine how many people within that school have a revolver. Any time something happens... Nowadays, in the ways those things are... any... "step on somebody's toes", you are already assaulting somebody... this is a very dangerous thing and has to have drastic measures.

On the other hand, the interpretation of P2 was more explicitly related to his personal opinion about the data.

P2: These data here, it shows and proves about the reality of a country that has a policy of well open access to weapons, right? And at the same time, there is no much control and no much oversight, right?

Finally, the third question was an opportunity to the participants to make final remarks about the data (Do you have any questions or comments about this research? Which one?).

P2 and S1 questioned about the survey sampling, the sites where the poll was taken and how the research was done, pointing out some possible biases. According to them, the information provided about the survey was insufficient to answer their questions, however, none of them expressed they didn't trust

the research. The comments of P1 and S2 were based on their opinions about the situation, taking into account their knowledge about social, economic and political aspects which might be related to the theme, and which could eventually influence the interpretation of the data. These two participants did not ask for further information.

FINAL CONSIDERATIONS

This study explored aspects of the exhibition of affective components in the interpretation of statistical data. Specifically, the empirical research data was generated from situations in which the participants interpreted statistical data. The participants were students engaged in different processes of teaching and learning related to affectivity and statistics. Students of Pedagogy attended courses only related to affectivity, and students of statistics had approached various contents related to data analysis, and they did not attend courses that comprise the theme of affectivity as curriculum content.

The research tasks were associated with statistical data of controversial themes. Our expectation was to propose research situations that could provide possibilities for the participants to present more personal expressions during their interpretation of the data. We also expected that participants' academic background could influence how they would interpret and to express their affectivity in relation to the data.

The analyses of protocols suggest that the interpretation of statistical data is a dynamic process, which do not follow predictable patterns. Participants responded in different ways during their processes of interpretation to most of the questions. They mostly expressed their opinions and feelings in responding to the questions, or at some moments they mixed their objective analysis related to the statistical knowledge with the subjectivity of their impressions about the data. Even more technical oriented questions (see task 2 question 2 and task 4 question 2) can be interpreted as expressions of feelings (e.g. scepticism) or expressions of a personal point of view.

In line with the theoretical investigations on interpreting statistical data, these preliminary data gave evidence that the process of interpretation is a complex process consisting of cognitive and dispositional components. Based on the preliminary research findings, we have no clear evidence that these components were determined by the curriculum background of the students (pedagogical versus statistical). However our findings make clear that a broader discussion about the processes of data handling and the interrelation with affective aspects is an important issue in the further development of teachers' understanding of statistical literacy. Further investigations have to reveal possible differences between first year students and students in a final stage of the courses.

ACKNOWLEDGEMENT

REFERENCES

Adler, J. (2000). Conceptualising resources as a theme for mathematics teacher education. *Journal of Mathematics Teacher Education, 3*, 205–224.

Arcavi, A. (2003). The role of visual representation in the learning of mathematics. *Educational Studies in Mathematics, 52*, 215–241.

Gal, I. (2002). Adults' Statistical Literacy: Meanings, Components, Responsibilities. *International Statistical Review, 70*, 1–25.

Lima, L. (1998). *Interpretação de gráficos de quantidades veiculados pela mídia impressa: Um estudo exploratório* [Interpretation of print media graphs of quantities: an exploratory study] (Unpublished master's dissertation). Federal University of Pernambuco, Recife, Pernambuco, Brazil.

McLeod, D.B. (1992). Research on affect in mathematics education: A reconceptualization. In A. D. Grouws (Ed.), *Handbook of Research on Mathematics Teaching and Learning* (pp. 575–598). New York, NY: Macmillan.

Monteiro, C. (1998). *Interpretação de gráficos sobre economia veiculados pela mídia impressa* [Interpretation of print media graphs about economics] (Unpublished master's dissertation). Federal University of Pernambuco, Recife, Pernambuco, Brazil.

Monteiro, C. (2005). *Investigating critical sense in the interpretation of media graphs* (Unpublished PhD thesis). University of Warwick, Coventry, England.

Monteiro, C., & Ainley, J. (2003). Developing Critical Sense in Graphing. In M.A. Mariotti (Ed.), *Conference of the European Society in Mathematics Education. Vol. 3.* (pp. 1–10). Bellaria, Italy. Paper retrieved from http://www.dm.unipi.it/~didattica/CERME3/proceedings/Groups/TG5/TG5_list3.html

Watson, J.M. (2005). Developing reasoning about samples. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227–294). Dordrecht, Netherlands: Kluwer Academic Publishers.

Waiselfisz, J.J. (2011). *Mapa da Violência 2011. Os Jovens do Brasil*. [Map of Violence 2011. The Young people of Brazil]. Brasília, Ministério da Justiça, Instituto Sangari.