

## INTERPRETATION OF STATISTICAL DATA: THE IMPORTANCE OF AFFECTIVE EXPRESSIONS

TAMIRES QUEIROZ

*Federal University of Pernambuco*  
*tamnqueiroz@gmail.com*

CARLOS MONTEIRO

*Federal University of Pernambuco*  
*carlos.monteiro@campus.ul.pt*

LILIANE CARVALHO

*Federal University of Pernambuco*  
*lmtlcarvalho@gmail.com*

KAREN FRANÇOIS

*Free University of Brussels*  
*karen.francois@vub.ac.be*

### ABSTRACT

*In recent years, research on teaching and learning of statistics emphasized that the interpretation of data is a complex process that involves cognitive and technical aspects. However, it is a human activity that involves also contextual and affective aspects. This view is in line with research on affectivity and cognition. While the affective aspects are recognized as important for the interpretation of data, they were not sufficiently discussed in the literature. This paper examines topics from an empirical study that investigates the influence of affective expression during the interpretation of statistical data by final-year undergraduate students of statistics and pedagogy. These two university courses have different curricular components, which are related to specific goals in the future professional careers of the students. The results suggest that despite differing academic backgrounds in both groups, the participants' affective expressions were the most frequent type of category used during the interpretation of research assignments.*

**Keywords:** *Statistics education research; Interpretation of statistical data; Affective aspects; Affectivity and cognition; Statistical literacy*

### 1. INTRODUCTION

Over the last few decades, an increasing number of studies supported the perspective that the interpretation of statistical data is not merely a technical procedure but a complex process in which people can establish relationships between data, infer information, and consequently construct knowledge (Ainley, Nardi, & Pratt, 1999; Shaughnessy, Garfield, & Greer, 1996). Several studies from the research field of statistics education provided empirical evidence that the interpretation of statistical data is a complex activity, which is also associated with affective aspects (e.g., Evans, 2000; Gal, 2002; McKnight, 1990;

Monteiro & Ainley, 2007). These findings are confirmed by literature reviews on emotions and cognition (Blanchette, 2006; Blanchette & Richards, 2010).

In this article we argue that the role of affectivity in interpreting statistical data is important to get a deeper understanding of how to foster statistical literacy and to improve statistics education. We therefore discuss the expressions of affective aspects involved in the interpretation of statistical data, based on a review on affect research in mathematics and statistical education, and on a broader literature review on affectivity and cognition. We further discuss an empirical study conducted with undergraduate pedagogy and statistics students. The aim was to explore and identify the affective expressions as part of the process of interpretation during different statistical tasks in order to get a deeper understanding of the interrelations of affectivity expressions in situations of data interpretation.

### 1.1. AFFECTIVITY AND STATISTICS EDUCATION

Lerman (2006) argued that for a long time a narrow perspective of the psychology of mathematics education prevailed, which did not take into account numerous aspects, including those related to affective elements involved in the teaching and learning of curricular contents of mathematics, which include statistical topics.

In some sense, most of the studies, which investigated statistical and mathematical topics, tend to emphasize the cognitive aspects involved. The term cognition is associated with mental processes that are usually linked to rationality and intelligence. Several psychological theories conceptualized cognition from different perspectives. For example, Piaget's approach emphasized the mental structures and the process of knowledge building by the epistemic individual (Piaget, 1974), whereas Vygotsky's socio-historical theories emphasized the cultural dimensions of knowledge acquisition processes (Vygotsky, 1978).

Similarly to the conceptualization of cognition, there are different descriptions of affectivity. In this article, we will use the construct *affect* as formulated by Philipp (2007): "a disposition or tendency or an emotion or feeling attached to an idea or object" (p. 259). According to this conceptualization, affect involves the following three aspects: *emotions* (feelings or states of consciousness); *attitudes* (manners of acting, feeling, or thinking that show one's disposition or opinion), and *beliefs* (understandings, premises, or propositions about the world that are thought to be true). Philipp (2007) emphasized that these three aspects of affect are not completely dissociated from cognition, and that belief is more cognitively influenced than emotions and attitudes are.

Evans (2000) claimed that a number of studies, which investigated learning and teaching mathematics tend to analyze non-cognitive factors, which are expressed by behaviours, physiological reactions, and visible feelings. According to Evans, those studies considered the relationship between cognitive and affective factors in terms of affect interfering with cognition, or sometimes supporting it. In order to develop a different perspective of the role of affective components in mathematics education, Evans analyzed college students' "maths life-histories" carrying out interviews composed of tasks in which the content might be viewed either as mathematical, or based on *out-of-school practices*. Those tasks also included the interpretation of graphs. The results indicated that misconceptions and memory lapses might be crucial and related to negative affective charges that are in many cases specific to school-mathematics practices. However, in using mathematics *out-of-school*, bad feelings associated with previous experiences do not seem to have interfered generally with the numerate aspects of the performance of the participants.

Hannula (2002) contended that the splitting of human mind into affectivity and cognition is only an analytical tool: “the interaction between the two is so intense, that neither can be fully understood in separation from the other” (p. 27). In the field of educational psychology, Wallon (1942) developed a theory in which affectivity is an essential feature of human beings. For Wallon, affectivity, cognition and learning cannot be conceptualized as dissociated elements in the human individual and his/her actions. Therefore, the analysis of both the affective and cognitive dimensions would be highly relevant in understanding the psychological processes involved in teaching and school learning.

McLeod (1992) completed 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 results and not with the processes involved. Most of the studies reviewed seem to conceptualize affectivity narrowly, merely investigating more specific concepts such as beliefs and attitudes. Additionally, these studies did not link affective and cognitive factors. Both types of studies, those that investigated only cognitive aspects and those that addressed only affective aspects, seem to have no impact on the learning and teaching of mathematics. McLeod suggested that beyond the extension of theoretical questions concerning the definition of affectivity and its relation to cognition, studies need to be based on research approaches that combine quantitative and qualitative methods. Philipp (2007) carried out a similar literature review to identify potential developments in the discussion concerning affectivity in mathematics education. The author concluded that over 15 years after the McLeod (1992) review, research studies began to make more explicit reference to the relationship between cognition and affectivity. These studies lead one to consider affective aspects also in the interpretation of statistical data.

The relation between affectivity and cognition from a broader perspective than just mathematics and statistics is the subject of a vast literature. The theories discussed in cognitive psychology about the relation between body, mind and emotions are rich and varied. However, they all maintain that there is no thought without emotion (Damasio, 1994). On this basis, it seems that it is to be expected that readers reveal emotional responses during their interpretation of statistical data. What is more important is that we now have evidence to investigate these affective aspects of the interpretation of statistical data.

In statistics education, only a few studies investigated affective aspects (Zieffler et al., 2011) and they mainly focused on attitudes towards statistics (e.g., Estrada, Batanero, & Lancaster, 2011; Gordon, 2004). Chiesi and Primi (2010) stated that the statistics achievements of university students seem to be the result of an interrelation between cognitive and non-cognitive factors. Therefore, an important pedagogical issue is the consideration of both in order to approach the complexity of teaching and learning statistics.

Another aspect to be studied is how the interrelation between emotions and reasoning works, and whether there is evidence that affect influences the interpretation or not. We have evidence from cognitive theory on the relationship between interpretation and reasoning; based on empirical findings with emotional and neutral conditional statements, Blanchette (2006) concluded that emotions did not affect interpretation. However, in their review article on behavioural research that has empirically examined the interaction between the affective system and higher-level cognition, Blanchette and Richards (2010) propose a dynamic interaction between cognitive and affective variables. The partition between these two fields may have originated from early distinctions between rationality

and emotions, between reasoning and passion. In line with Blanchette and Richards (2010), we advocate that statistical reasoning – as one form of higher-level cognition – has to be approached as a dynamic process taking into account the context in which the statistical data appear.

## 1.2. INTERPRETATION OF DATA AS A COMPLEX ACTIVITY

In contemporary society, statistical data is used in different contexts. Wild and Pfankuch (1999) suggested that in contextual inquiry people act as “data producers” and usually have to interpret their own data and report their findings using graphs (e.g., researchers and statisticians). Gal (2002) labelled as “reading contexts” those everyday situations in which people see and interpret graphs (e.g., watching TV, reading newspapers, looking at advertisements while shopping, visiting internet sites, etc.). Reading contexts are related to situations in which people are interacting with media or mass communication, which is associated with social vehicles of communication such as radio, television, and widely distributed print media. Other important areas in which interpretations of graphs are developed are school contexts (e.g., Monteiro & Ainley, 2007). In these contexts the interpretations acquire specific characteristics, which make them unlike enquiry and reading contexts. For example, when graphs are used in statistics classrooms, their common purpose is to teach elements of school curricula.

Gal (2002) stated that the differentiation of contexts does not necessarily mean that they are homogeneously defined because individuals can develop different kinds of participation. For example, people engaged in a reading context can be actors, speakers, writers, readers, listeners, or viewers, either in passive or active roles. Monteiro (2002) investigated these processes of interpretation of statistical data among different groups of professionals with the aim to identify whether their personal background influences their interpretations. Lima (1998) analyzed the interpretation of data developed by designers and mathematics teachers. The author concluded that their interpretations were different in the way they read the data; however, both groups were alike in their use of statistical and mathematical knowledge during their work of interpreting. Monteiro (2002) investigated the processes of interpreting graphs in print media in a group of business people with different academic backgrounds and a group of economists. He did not identify differences in relation to the strategies of problem solving, although the economists tended to include more estimates in their interpretations. In these studies of interpreting data, the way readers use their mathematical and statistical knowledge is complex and not the consequence of any single aspect such as their academic background.

McKnight (1990) explores the interpretation of graphs that would be encountered either in popular science media (e.g., *Scientific American*) or in academic texts and monographs. Some graphs were related to propositions that were patently false (e.g., “storks bring babies”) or related to propositions that would seem more likely to be true (e.g., “population will increase faster in developing countries than in developed countries”). Seven participants from academic backgrounds (professors and graduate students) took part in the study, answering multiple-choice and open-ended questions. The results indicated that only those tasks that involved the observation of facts in the graphs seemed to be “unproblematic” (McKnight, 1990, p. 183). A preliminary catalogue of errors observed emphasizes the effect of the reader being distracted by extraneous knowledge (McKnight, Kallman, & Fischer, 1990, p. 14):

Translation from the ‘clean’ world of abstract mathematics to the ‘messy’ world of everyday reality – in which all of our knowledge has links to other knowledge as well as links to personal beliefs and emotional reactions – introduces yet another complexity. Sometimes that other knowledge – or what one thinks is other relevant, linked knowledge – or those beliefs and affective reactions interrupt the more cognitive, information processing tasks of interpreting the graph.

McKnight and her colleagues recognized that affectivity, expressed by readers when they are interpreting media graphs, is a relevant aspect that needs further investigation.

The term mobilization (Monteiro & Ainley, 2007) is related to the opportunity of re-using or re-sourcing (Adler, 2000) previous knowledge and experiences. Mobilization seems to be a process in which readers explore the data, confronting it with their own perspective and their previous experiences relating to the data. However, this process of mobilization in interpreting media graphs does not happen ‘naturally’. In order to recall and use their previous knowledge and experiences to interpret a graph, readers need to establish a certain level of engagement in the task. It leads to the articulation and recontextualization of the knowledge and experiences mobilized, comparing them to the data. The reader also needs to balance these different elements. Therefore, there is no direct application of knowledge and experiences in the process of interpretation.

This complex process of mobilizing and balancing different elements during the interpretation of statistical data is called *critical sense*. Monteiro (2005) develops the idea of *critical sense* during the interpretative process of statistical graphs. He distinguishes four types of previous knowledge and experiences.

1. *Mathematical knowledge*: knowledge related to statistical and mathematical notions and concepts associated with graphical representation, such as numerical and quantitative relationships, and notions of measurement and geometry.
2. *Contextual reference*: contextualization of the data displayed in the graph making references, which are related to their formal knowledge in different areas (e.g., geography, health care, sociology, etc.) and their opinion (e.g., personal view about governmental contraception policies or traffic safety).
3. *Affective exhibition*: explicit expression of aspects of personal emotions and feelings related to skepticism, sarcasm, anger or hope in relation to the data interpreted.
4. *Personal exemplification*: exemplification and justification of answers based on personal factual experiences related to the data displayed.

As discussed by Monteiro (2005), affective elements are involved in the interpretation of statistical data. These aspects need more investigation in order to obtain more evidence about their importance. In our research design we also analyze the concept of affective elements based on the theoretical investigations of Philipp (2007). Goldin (2002) expressed concern that research in mathematics education has primarily focused on cognition and far less on affect. He ascribes this situation, in part, to the popular myth that mathematics is a purely intellectual endeavor in which emotion is all but absent. It seems that this perspective is also commonplace in people’s opinions about statistics. Goldin (2002, p. 60) emphasizes that, when people use mathematics, the affective system is not merely auxiliary to cognition, it is central.

In the following section we report on our empirical investigation that explored the affective expressions of undergraduate students while interpreting statistical data in research tasks that use real data.

## 2. INVESTIGATING AFFECTIVE ASPECTS IN STUDENTS' INTERPRETATIONS OF STATISTICAL DATA

In an attempt to fuel the debate on these issues, we developed a study with final-year students from two university courses that teach future professionals who can contribute to the development of statistical literacy: statistics and pedagogy. In the Brazilian higher education system, each university course has a curriculum that is related to the specific goals for a given professional area. In general, the curriculum has compulsory and optional modules, which are called disciplines.

The statistics course has disciplines that emphasize the theoretical and technical dimensions of statistics. When a student completes this course, he/she will be a statistician, a professional who can develop activities in different institutions and functions including those that may have repercussions on the development of statistical literacy in the community. The undergraduate students who complete the pedagogy course are called pedagogues. They can teach in early childhood education (nursery and kindergarten), and in the early years of fundamental school (primary school from Year 1 to 5). The pedagogues can also teach at Youth and Adult Education, which is a schooling modality for people who left school before finishing compulsory formal education, or for those who have never been to school at all. For all these groups, pedagogues need to teach statistical notions and concepts. However, the discipline of statistics for pedagogy is optional. Therefore, the students do not necessarily choose this option. However, they do study a very short unit included in a discipline for mathematics teaching methodology. Most disciplines in pedagogy cover the subjective aspects of school teaching and learning.

By choosing students from these two university courses with different course profiles, we expected that the way students interpret the statistical data would have been influenced by their course (Monteiro, 2002). For example, we anticipated that students, who had received only technical components linked to statistics during their undergraduate course, would develop different interpretative processes compared with students who had studied components related to the more subjective aspects. In order to achieve the aims of this empirical study, we developed a qualitative design based on standardized open-ended interviews comprised of tasks in which the students were asked to interpret statistical data presented in a variety of contexts.

### 2.1. PARTICIPANTS

The study was carried out with 11 final-year university students from two undergraduate courses at a public higher-education institution, the Federal University of Pernambuco (Brazil): 6 of them were studying pedagogy, and the other 5 were pursuing an undergraduate course in statistics. These courses are developed in academic semesters. A complete pedagogy course covers 10 semesters, while the undergraduate statistics course spans 7 semesters. We asked the coordinator of both courses for permission to access their students on a volunteer basis. For the pedagogy course, we approached the students in their classrooms. We described the aims of this study and its importance for statistics education. The interviewer gave sufficient information on each stage of the interview, including the average time it would take, and explained that it was not an assessment situation. Concerning the statistics students, the course coordinator gave us a list of 11 final-year level students. We contacted them by email, outlining the project in similar terms we used with the pedagogy students.

In order to ensure privacy and simplify the identification of participants' courses, we used a one-to-one anonymous code. The letter P followed by a number refers to pedagogy course students, and S for statistics course students. Table 1 shows information about gender and age of the participants: most of the P participants were female; in general, the S participants were younger than the P participants.

Table 1. Gender and age of the participants

	P1	P2	P3	P4	P5	P6	S1	S2	S3	S4	S5
Gender	F	F	F	F	F	M	M	F	M	M	F
Age (years)	40	23	30	30	33	26	25	21	27	23	26

## 2.2. INTERVIEWS: 3 PHASES

The individual standardized open-ended data collection was developed in June and July 2014. The use of interviews for data collection was targeted towards obtaining more detailed elements from the participant's interpretations. According to Kvale (1996) qualitative interviews are an important methodological instrument, which provides an opportunity to investigate the nuances and meanings of participants' interpretations. Lajoie, Jacobs, and Lavigne (1995) highlighted the need to give students the opportunity to verbalize how they came to an answer when they solved statistical problems. In addition, interviews can also record aspects of non-verbal communication that are an important aspect in participants' responses. The interviews were recorded via Camtasia Studio 8 software, which also recorded their operations on the computer. All participants had previously agreed to this procedure. The interviewer put the participants at ease before the interview, which was conducted in a quiet and reserved room at the Center of Education at the University of Pernambuco. The average time for the interviews was 65 minutes. The interviews were subdivided into three phases.

The *first* phase comprised eight questions concerning the background of the participants as well as their knowledge and experiences with statistics.

1. Have you ever studied a statistical discipline before going to college? If yes, which one and how was that experience?
2. Do you remember any statistics class prior to your undergraduate course? If yes, which and why?
3. How would you rate your performance in statistics? Expound that.
4. In your opinion, which factors influenced this performance?
5. Do you think you use statistics on a daily basis?
6. Would you like to learn more about statistics or do you think you know enough to perform? Develop your response.
7. Why statistics are important in your profession?
8. Do you know someone who has a hard time with statistics? If yes, what is the greatest difficulty for that that person, and why do you think they have it?

The *second* phase encompassed four tasks related to the interpretation of data from different media. These tasks will be presented in detail in the next subsection.

- Task 1. A short written report;
- Task 2. A line graph;
- Task 3. An excerpt from the news media;
- Task 4. A statistical data table.

During the *third* phase of the interview, the participants were asked about their impressions relating to the research tasks.

1. Which aspects have caught your attention in these tasks most? Why?
2. Which task has been the most interesting to interpret? Why?
3. Which are the least interesting tasks? Why?
4. What is the most familiar topic in these tasks? Why?

The transcriptions of the interviews of all three phases generated a protocol for each participant originally written in Brazilian Portuguese. The authors translated parts of the protocols into English. The methodological proposal and the procedures of the research were evaluated and approved by a committee composed of experienced researchers, two participants of the postgraduate program in Mathematics and Technological Education at the Federal University of Pernambuco and one external expert from Lisbon University. At the end of this investigation, participants received a file with the final report.

### 2.3. TASKS: PHASE 2

In this section we present the four tasks, which were used during Phase 2 of the interview. The main reason for choosing the materials, which comprised the tasks, was that the topics were somewhat controversial. We expected that this type of data and its context would induce participants to make comments related to technical aspects as well as to their affective reaction to the data. The researcher read the specific questions for each task. The material associated with each task was displayed on a laptop. The materials concerning data, which formed the research tasks, were presented to the participants as slides on a screen. The computer was located in front of the participants, and they used the mouse to point out elements of data representation and to pass onto the next task when prompted by the researcher. The tasks were presented in the same order for all participants. The students could spend as much time as they deemed necessary for a given task. For each task specific questions were asked. We tried to frame the issues so that the participants could discuss the data, giving students room to express their knowledge as well as their affective reaction about the data. We tried to organize the tasks so as to approximate a reading context (Gal 2002). Therefore, we asked that students only develop interpretations of the data presented.

**Task 1.** The core material consisted of a short written report on a survey concerning access to mammography in 2012–2013, published in a national Brazilian magazine.

Cancer – access to mammography, a screening test for breast cancer, has become more readily available in the public health system. The VIGITEL data reveals that 77.4% of 50 to 69 year old women have taken the exam in the past two years. In 2007, the rate was 71.1%. The highest rate of participation, however, is in the capitals of the most developed states in the country: Curitiba (90%), Distrito Federal (87%) and Belo Horizonte (86%). In the North and Northeast regions, the discrepancy is striking compared to the rest of the country. In João Pessoa, only 61% of women were scanned, in Recife 64%, Boa Vista 67% and in Manaus 68%. (VEJA magazine).

1. What can you conclude from the results presented in this news item?
2. Do you have any questions about this data? If yes, which one(s)?
3. If I say “VIGITEL” means “Risk and protective factors surveillance for chronic diseases telephone survey” would that change anything in your considerations? Why?
4. If you could predict what this data will be like in 10 years, in 2023, how do you think it would evolve? Why?



**Task 2.** The questions below relate to a line graph (Figure 1, originally published in Waiselfisz, 2011) that shows the percentages of deaths among the Brazilian youth population caused by traffic accidents, as a function of mode of transport, between 1998 and 2008.

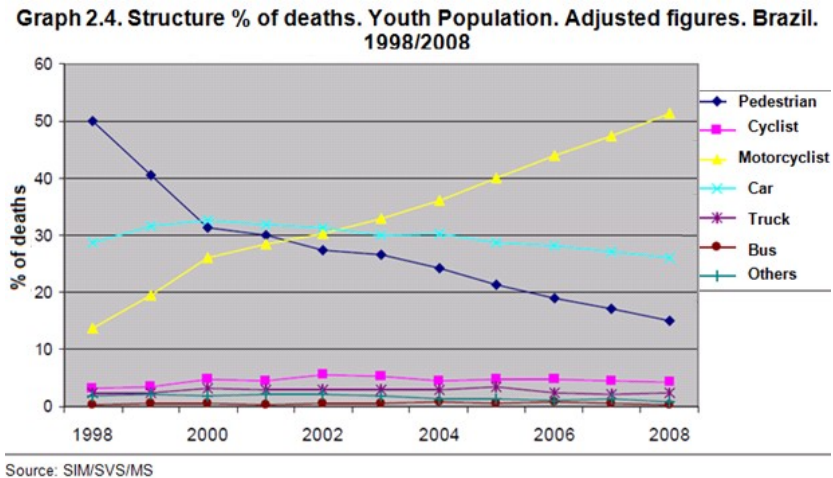


Figure 1. Structure of Brazilian youth mortality from 1998 to 2008 – adjusted figures

1. What can you conclude from the results presented in this graph?
2. If you could ask the graph's designer a question, would you do so? If yes, which one(s)?
3. What would you say? Which elements would you emphasize?
4. How do you think this data evolved between 2008 and 2013? Why do you think that?

**Task 3.** Originally used by Watson (2005), this task is an excerpt from the news media concerning the use of handguns by high-school students:

A survey demonstrates that about 6 out of 10 United States high-school students say they could get a handgun if they wanted one, a third of them within the hour. The poll of 2,508 junior and senior high-school students in Chicago also found that 15% had actually carried a handgun within the past 30 days, and that 4% took one to school.

1. Which conclusions can you draw from this study?
2. What do you think about the data collected in this survey?
3. Do you have any questions or comments about this research? Which one(s)?

**Task 4.** Answer the following questions related to the data table concerning the domestic travel market in Brazil 2010–2014 (Figure 2, published by the Brazilian government prior to the 2014 Soccer World Cup, held in Brazil).

1. Which conclusions can you draw from this table?
2. Do you have any questions relating to the table? If yes, which ones?
3. How can the number of licensed vehicles be an indirect indicator?
4. How would you comment the variation in the number of arrivals in the Regular Public Road Transport Collective?
5. Can it be stated that any of these indirect indicators would have a greater influence on the number of domestic trips taken?

The projection of indicators related to the domestic travel market for the period 2011-2014

Direct indicator	Unit	2010	2011	2012	2013	2014
Domestics trips taken	Unt	187.406.328	197.863.601	210.867.197	225.771.290	243.408.543
Indirect indicators	Unit	2010	2011	2012	2013	2014
Arrivals on domestic flights	Unit	59.697.092	63.055.053	67.232.450	72.022.763	77.694.555
Car rental for tourists	Unit	179.759	201.199	225.197	252.056	282.120
Average price of domestic regular flight ticket (YIELD)	R\$	325	338	351	365	387
Arrivals on regular public road transport	Unit	59.496.017	57.871.324	56.290.997	54.753.825	53.258.630
Licensed vehicles - cars	Unit	2.801.347	3.089.647	3.407.618	3.758.314	4.145.100

Sources: MTur e FGV.

Figure 2. Projections of indicators in the Brazilian domestic travel market for 2011–2014

## 2.4. DATA ANALYSIS

The interviews were transcribed verbatim. The content analysis of these protocols was based on an interpretative approach using the WebQDA software (Costa, Breda, Pinho, Bakas, & Durão, 2015). The first procedure was to develop a floating appraisal of the whole protocol and marking the significant parts. In this reading, we tried to identify similarities and differences in the ways participants interpreted the data presented in the research tasks as well as in its other spoken occurrences referring to the initial part (participants' background) and the final part of the interviews (their thoughts about the tasks). The second procedure was to categorize the interview fragments. Based on a review of literature that investigated the elements involved in the interpretation of statistical data (Gal, 2002; Monteiro, 2005), we used three main categories:

- (i) Affective expressions,
- (ii) Mathematical and statistical knowledge, and
- (iii) Data reading.

The *affective-expressions* category is related to excerpts from the discourse in which the participants conveyed some kind of emotion, attitude or belief – conform to Philipp's (2007) analysis of the concept 'affect' – regarding the statistical data presented in the tasks. During the categorization of the interview extracts, six subcategories were identified linked to affective expressions related to:

- (a) Experiences with statistics;
- (b) The data of the tasks;
- (c) Previous personal life experience;
- (d) Opinions on a contextual aspect;
- (e) Concise explanations;
- (f) The interaction between researcher and participant.

The *mathematical and statistical knowledge* category is related to interpretations in which the participants provided explicit elements of their knowledge about the presented data. Finally, the *data-reading* category is associated with responses in which they only repeat the data displayed in the task materials. After categorizing all protocols using WebQDA, we explored the relationships between these categories. This was crucial in order to gain an overview as well as to focus on the characteristics of specific categories.

### 3. RESULTS

The analysis of all interviews (Phases 1–3) generated 513 excerpts, which were classified according to the analytical categories (i)–(iii). Table 2 gives an overview of the classified interview extracts related to the type of course and main category.

*Table 2. Structure of all interview extracts (N = 513, all phases) related to analytical category and course*

Categories/Course	Pedagogy		Statistics		Total	
(i) Affective expressions	230	(44.8%)	125	(24.4%)	355	(69.2%)
(ii) Statistical and mathematical knowledge	65	(12.7%)	55	(10.7%)	120	(23.4%)
(iii) Data reading	26	(5.0%)	12	(2.3%)	38	(7.3%)

*Affective expressions* were the most recurrent category representing 69.2% of the classified extracts. Pedagogy students account for 64.7% (N = 230) of speech fragments categorized as affective expressions, and the statistics student group provided 35.2% (N = 125) of their responses to that category. In the following sections we present the analysis of these categories and discuss the aspects associated with each main category (i)–(iii).

#### 3.1. AFFECTIVE EXPRESSIONS

As we were developing the analysis of the excerpts classified as affective expressions, we also identified a classification with subcategories (a)–(f). Table 3 presents these subcategories and the frequency of extracts classified for each type of course.

*Table 3. Structure of affective expressions (N = 355) related to subcategory and course*

Subcategories\Course	Pedagogy		Statistics		Total	
All affective expressions	230	(64.7%)	125	(35.2%)	355	(100.0%)
(a) Experiences with statistics	51	(14.4%)	41	(11.5%)	92	(25.9%)
(b) Data of tasks	55	(15.5%)	33	(9.3%)	88	(24.8%)
(c) Previous personal life experience	32	(9.0%)	14	(3.9%)	46	(12.9%)
(d) Opinion with contextual reference	41	(11.5%)	22	(6.2%)	63	(17.7%)
(e) Concise explanation	25	(7.0%)	13	(3.7%)	38	(10.7%)
(f) Researcher/participant interaction	26	(7.3%)	2	(0.6%)	28	(7.9%)

**(a) Experiences in statistics** In this category, the classified excerpts referred to participants' previous experiences and/or to their attitudes regarding statistics. We categorized those excerpts from the interviews in which students spoke about their feelings concerning statistics/mathematics, and exposing what they thought about them, what they liked or disliked in statistics/mathematics, and their reasons. This was the most frequent subcategory, representing 25.9% of all extracts classified as affective expressions. For exemplification, we will furnish extracts of interviews which are categorized as *affective expressions related to experiences with statistics*. The parts in

*italics* are the classified excerpts. The following excerpts come from the second phase of interviews with P1 and S3 while they were beginning to respond to Task 2.

P1: *Oh dear! When it begins with these graphs! ...*

Intv: Don't you like graphs?

P1: *No, I don't like them much.*

S3: Look ... it is a thing that I also find very ... *cool in statistics. When we see the reality of our everyday life, which is reflected in a graph.* In graphs like this one. *So, I like graphs.*

**(b) The data of the tasks** In analyzing the tasks, students occasionally express some kind of emotion related to the data. All participants used this type of affective expression. For example, when P5 was interpreting the data related to the use of guns by students (Task 3), she could not separate the information displayed from her feelings about it.

P5: *My sister! The little one could get a gun, couldn't she?* I am not sure, that... Out of 2.508 students, 15% carry guns, I don't know, *I can only think about the social implications, I am not thinking about mathematics, no.*

S3 had a similar reaction; when he interpreted the data from Task 1 mixing his hopes and wishes for better statistics on the subject.

S3: It is because ... as a lot of people are talking about access to mammography, then we ... *what is desirable is an increase, isn't it? And we have faith (smiling) in the people responsible for that [the authorities], they also look forward to an increase ... the increase of this [access to mammography], don't they?*

According to Monteiro (2005), when confronted with the data, the reader also reveals personal expectations, feelings and ideas about the data and so maybe participants cannot only follow the mere analysis of the data but also need to express, somehow, what they think and believe concerning the data.

**(c) Previous personal life experience** This subcategory is related to those excerpts in which the participants establish a relationship between personal experiences with their interpretations of the data. Students use examples taken from their own lives to explain their views or to agree or disagree with what was exposed in the task. The following excerpts exemplify that, and how, participants brought their own personal experiences into their interpretation of the data presented in the tasks.

P3: [Interpreting Task 1] I felt sorry for Recife, Recife is the only place that really works, when *I brought my mother to take these tests ... so ... in the Cancer Hospital, I didn't see this accessibility anywhere else.* They always send these cases to that hospital and to get an appointment there is really difficult.

S3: [Interpreting Task 4] Look ... the first thing that came to mind ... Well ... that stood out, was exactly ... I have a personal relation to this question, because I spent those two years in the metro [working], right? And we live with that. And one thing that we try ... *I tried to work with the other staff members in the metro, in order to promote public transportation, right?*

An interesting incident happened immediately after the interview with S5. While S5 was interpreting the data on traffic accidents (Task 2), she repeated three times that she was sleepy and apologized for the way she was responding. She was evaluating her performance, recognizing that she could do better. At end of the interview, the participant asked the researcher if he was still recording. After she made sure that the recording was off, she said that her brother had died just two years before in a motorcycle accident. He was 18 years old at that time. She seemed to be very sensitive about the data on accidents. The researcher engaged her in a conversation to calm her down because of her emotional state. Albeit the participant reported this fact only at the end of the interview, it can be inferred that her feelings about the statistical data influenced the way she was interpreting the data.

**(d) Opinion on contextual reference** This category refers to those excerpts of the interviews in which the participants express opinions regarding the data and justify or explain their opinion referring to some aspect associated with a social context. As a social context we included references to politics, law, economics, or mentioning the facts that they see in television programs or media in general. In an interview situation, these types of responses are voiced with certain feelings or emotional reactions and are expressed by intonations or gestures. We identified excerpts related to this subcategory in all interviews, such as the following ones.

- Intv: But if you could predict how they will be in 10 years, would you say, you ... How do you think they would be? [Referring to figures about the accidents presented in the graph]
- S5: *Due to the modernization of the population, I believe it will increase, yes, the amount of people involved in prevention (...)*
- P3: It is also what is drawing attention because it would be a logical issue ... and for the cyclists what I think is that ... from now on ... well ... in few years, I don't know ... *in 10 years this will increase because it is gaining wider acceptance, ... [for example] there is a campaign "to use bicycles" (...)*

The analysis of these two excerpts suggests that their interpretation of the data on accidents (displayed in Task 2) was based on their perception of a social context. S5 seems to have a positive point of view concerning society. He seems to believe that "the modernization of society" could include the idea of "risk-prevention". On the other hand, although P3 approves of the social ad campaigns promoting the use of bicycles, she has a pessimist perspective and emphasizes a probable increase in the number of accidents.

**(e) Concise explanation** This sub-category includes those excerpts in which the participants express an opinion about the data without justifying or commenting on it. It is a short and clearly expressed opinion without verbalized justification or comments. It is likely that the participants made some inferences about the data even though they were not verbalized and thus not registered as interpretation during the interview. Thus, the participants gave short answers when they were interpreting. The following excerpts are examples of *concise explanations*.

- P6: [Responding to Task 4] *People will stop using more road transports, bus.*
- S4: [Commenting Task 2] *Perhaps the pedestrians were no longer pedestrians and began to be motorcyclists.*

Although P6 and S4 did not make any other explicit exposition about their conclusions, they verbalized a personal evaluation of the data presented.

**(f) Researcher / participant interaction** This subcategory includes the dialogues relating to the interviewer or to the study. Even though we sought to create a situation similar to a reading context (Gal, 2002), the research interview is an uncommon form of dialogue for the participants. In addition to that, once the interviewer asks for collaboration, it is possible that, in their willingness to collaborate with the researcher, the participants were striving to provide the best responses. In other words, this subcategory is related to this interaction between researcher and participant during the interview. For example, in the following excerpt from P2, she seemed to check her answers, but not in a way that she was asking for approval.

P2: Gosh ... I do not know if that's right, but if he grew up, where is it? 71 to 77, like, 6 percent in two years, in ten years ... oh dear! Was it that considerable? Six, twelve, eighteen, twenty-four, thirty [counting on fingers] I don't know ... Is that correct? It's wrong, isn't it?

In the extract above the sentences ended with questions. These formulations are common in colloquial Brazilian Portuguese and do not represent genuine questions, but a way to bring the interlocutors closer together during a conversation.

### 3.2. STATISTICAL AND MATHEMATICAL KNOWLEDGE

We expected students to submit answers in reference to their mathematical and statistical knowledge for the interpretation of the data. The total number of occurrences for this category was 120 excerpts (see Table 2). Each group of participants had a closer number of this type of comment:  $N = 65$  or 54.1% from pedagogy students and  $N = 55$  or 45.8% from statistics students. An example from this category is this excerpt from P4's interview.

P4: "Interviews"? It is a percentage ... so ... *if we were to calculate the number of high-school students throughout the United States; that is a low percentage, right? But concerning Chicago, I do not know how many students there are ... Maybe it was ... in only a few schools too, right?*

P4 paid attention to the sampling of the data. She says that it depends on the population, a sample can be considered as low or high proportion. The student seems unsure of her response because the news report does not provide any data regarding how the sample was taken. In S4's excerpt, he explicated how he came to a result of 60% in Task 3, concerning guns in schools.

S4: Ready. Because ... at least my mind is trained for it, whatever. I do not know if it is trained, well not in a robotic sense, but so ... because of the recurrence of these numbers, *what I would do is the following: 6 in 10, then it would be 6 divided by 10, I would multiply 6 by 10, 60 divided by 10, 100, right? This gives the proportion of 60%, right? Gives the percentage. Then, so ... 6 in 10 is 60%, right? Much more ... is more than half, right? It would be 50% plus one.*

The analysis of this excerpt suggests that the participants have a critical view regarding the data, questioning what was being presented. We also noticed that the students' responses from both groups – the pedagogy and the statistics group – are qualitatively similar, because both groups challenged points of the tasks.

### 3.3. DATA READING

The data-reading category includes the students' responses in which they repeat what is written in the task or the data displayed in the graph or table. We realize that the answers in this category are less expanded (7.3% see Table 2). Participants had no explanations or opinions, or did not even show explicit mathematical or statistical knowledge regarding the data. Examples from this category are presented in the following excerpts from the interviews.

- P1: From 2013 to 2007 increased slightly, *which is now 77% and in 2007 it was 71%*, then increased by 6%.
- S5: Those pedestrian accidents, the young population decreased over time 1998-2008 (...).

We also realize that the pedagogy students have a higher percentage for this type of response with  $N = 26$  or 68.4% (from 38 excerpts).

Based on the descriptive analysis of the occurrence of the three main categories (i)-(iii), we can conclude that *affective expressions* represent more than two thirds or 69.2% of all categorized expressions when interpreting statistical data. This underlines the importance of affective elements in the interpretation of data. On the other hand, there is a low frequency in the data reading category (7.3%), which seems to be a good thing, as it implies that the participants tried to analyze the data from different perspectives and did not just repeat the exposed data.

## 4. FINAL CONSIDERATIONS

This study explored the affective component in the interpretations of statistical data, which we call *affective expressions*. According to Philipp's theoretical frame (2007), we perceive *affective expression* as an exhibition of emotions, attitudes or beliefs during the process of interpreting data. The quantification of excerpts gave us an idea of the occurrence of *affective expressions* related to the different subcategories and the different background of the undergraduate students.

The research tasks presented data concerning controversial contemporary issues, and our intention was to present tasks that would encourage personal expressions during data interpretation. We expected that the different curricular background (pedagogy/statistics) would influence how the participants interpreted and expressed their affect in relation to the data. Regarding students' previous academic experiences in the curriculum of their degree course, we emphasize that the pedagogy students had participated in a university course involving aspects more related to affectivity, while the statistics students had attended various related disciplines of data analysis and had not passed courses involving subjects related to affectivity. However, we cannot say that these aspects had an influence on the data found in our research. Both groups of undergraduate students (pedagogy/statistics) expressed in their interpretations what we called (i) affective expressions, (ii) statistical and mathematical knowledge, and (iii) data reading.

One conclusion drawn from the quantitative data is that pedagogy students produced a higher number of excerpts classified as affective expressions than statistics students and this for all subcategories of affective expressions. When we compared the frequencies of response categories between both groups of students, we found that there was not much difference in the types of affective expressions. There is no specific type of affective expression that characterizes a group of students. In affective expressions linked to the task data, both statistics and pedagogy students were continually displaying their affectivity concerning the research tasks. Their previous personal experiences also surfaced frequently, especially during Phase 1 of the interview. The views with contextual references tended to occur when participants were trying to justify their view on a particular topic or task (Phase 2); for example, when trying to justify why an index would increase or decrease over the coming years. These results lead us to point out that affective expressions have taken center stage in students' responses, and that they emerged in all proposed tasks. These findings confirm and strengthen the theoretical debate on the interpretation of statistical data as a complex activity, which is associated with affective aspects.

The results of our research lead us to think that affect may also have implications in classroom practices. When addressing statistical contents, teachers can make choices that are influenced by affective components. For instance, choosing a theme may depend on the sense of familiarity one has with the subject matter or could be influenced by the teachers' own preferences. In this regard, this research can advance such questions, but cannot deliver the means to supply the answers. Future studies should be developed to understand what it means for instructors and those learning statistics, and, how we might harness or be mindful of affect when interpreting statistical findings.

There were also limitations of the empirical study, which need to be addressed in further investigations. A first limitation is the selection of participants. We selected students from undergraduate courses and they volunteered to take part within the chosen research design, which can generate a bias due to the relationship with the interviewer. A second limitation is the selection of research tasks, which were comprised of themes associated with controversial issues. A third limitation is the process of categorization of protocol excerpts and the limitations for a quantitative analysis.

In future studies, the discussion needs to dig deeper into expanding the research into other student cohorts, varying levels of statistical knowledge, and other countries. The research design should take into account the interpretation of neutral or less affective data and it would be interesting to compare interpretations of the same data where the context (emotive vs. non-emotive) is manipulated. To overcome the limitation related to the categorization, a more objective process is needed if we want to quantify research findings. A more qualitative research design is recommended if we want to develop a dynamic categorization in order to analyze the interrelations between the different types of affective expressions.

#### **ACKNOWLEDGEMENT**

We thank all our participants who agreed to contribute to this study and interpret statistical data, which could represent a sensitive matter for them. We also thank the reviewers of this paper who provided constructive support and challenging questions. The presented study was supported by the CAPES Foundation, Ministry of Education of Brazil.



## REFERENCES

- Adler, J. (2000). Conceptualising resources as a theme for mathematics teacher education. *Journal of Mathematics Teacher Education*, 3(3), 205–224.
- Ainley, J., Nardi, E., & Pratt, D. (1999). Constructing meaning for formal notation in active graphing. In I. Schwank (Ed.), *Proceedings of the First Conference of the European Society in Mathematics Education* (Vol. 1, pp. 189–200). Osnabrück, Germany: Forschungsinstitut für Mathematikdidaktik.  
[Online: [www.mathematik.uni-dortmund.de/~erme/index.php?slab=proceedings](http://www.mathematik.uni-dortmund.de/~erme/index.php?slab=proceedings)].
- Blanchette, I. (2006). The effect of emotion on interpretation and logic in a conditional reasoning task. *Memory & Cognition*, 34(5), 1112–1125.
- Blanchette, I. & Richards, A. (2003). The effect of emotion on conditional reasoning. In R. Alterman & D. Kirsh (Eds.), *Proceedings of the Twenty-Fifth Annual Conference of the Cognitive Science Society* (Vol. 1, pp.151–156). New Jersey: Lawrence Erlbaum.
- Blanchette, I., & Richards, A. (2010). The influence of affect on higher level cognition: A review of research on interpretation, judgement, decision making and reasoning. *Cognition & Emotion*, 24(4), 561–595.
- Chiesi, F., & Primi, C. (2010). Cognitive and non-cognitive factors related to students' statistics achievement. *Statistics Education Research Journal*, 9(1), 6–26.  
[Online: [iase-web.org/Publications.php?p=SERJ\\_issues](http://iase-web.org/Publications.php?p=SERJ_issues)]
- Costa, C., Breda, Z., Pinho, I., Bakas, F., & Durão, M. (2015). The use of the Web Qualitative Data Analysis (webQDA) software to maximize the potential of qualitative analysis. *Proceedings of the 4th Ibero-American Congress on Qualitative Research* (Vol. 3, pp. 70–75). Aracajú, Brazil: Tiradentes University.
- Damasio, A. R. (1994). *Descartes' error: Emotion, reason, and the human brain*. New York: Putnam.
- Estrada, A., Batanero, C., & Lancaster, S. (2011). Teachers' attitudes towards statistics. In C. Batanero, G. Burrill, & C. Reading (Eds.), *Teaching statistics in school mathematics – challenges for teaching and teacher education* (pp. 163–174). New York: Springer.
- Evans, J. (2000). *Adult's mathematical thinking and emotions: A study of numerate practices*. London: Routledge Falmer.
- Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities (with discussion and a rejoinder by the author). *International Statistical Review*, 70(1), 1–51.
- Goldin, G. A. (2002). Affect, meta-affect, and mathematical belief structures. In G. C. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 59–72). Dordrecht: Kluwer Academic Publishers.
- Gordon, S. (2004). Understanding students' experiences of statistics in a service course. *Statistics Education Research Journal*, 3(1), 40–59.  
[Online: [iase-web.org/Publications.php?p=SERJ\\_issues](http://iase-web.org/Publications.php?p=SERJ_issues)]
- Hannula, M. S. (2002). Attitude towards mathematics: Emotions, expectations and values. *Educational Studies in Mathematics*, 49(1), 25–46.
- Kvale, S. (1996). *Interviews: an introduction to qualitative research interviewing*. London: Sage.
- Lajoie, S., Jacobs, V. R., & Lavigne, N. C. (1995). Empowering children in the use of statistics. *Journal of Mathematical Behavior*, 14(4), 401–425.
- Lerman, S. (2006). Socio-cultural research in PME. In A. Gutiérrez & P. Boero (Eds.) (2006). *Handbook of research on the psychology of mathematics education: Past, present and future* (pp. 347–366). Rotterdam: Sense Publishers.
- Lima, L. (1998). *Interpretação de gráficos de quantidades veiculados pela mídia impressa: Um estudo exploratório*. (Unpublished master's thesis). Federal University of Pernambuco, Brazil.

- McKnight, C. C. (1990). Critical evaluation of quantitative arguments. In G. Kulm (Ed.), *Assessing higher order thinking in mathematics* (pp. 169–185). Washington, DC: American Association for the Advancement of Science.
- McKnight, C. C., Kallman, C., & Fisher, M. (1990). Task analyses of critical evaluations of quantitative arguments: first steps in critical interpretation of graphically presented data. In *Proceedings of the Annual Meeting of the American Educational Research Association* (pp. 1–22). Boston: AERA.
- McLeod, D. B. (1992). Research on affect in mathematics education: A reconceptualization. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 575–596). New York: Macmillan.
- Monteiro, C. (2002). Investigating the interpretation of economics graphs. In A. Cockburn & E. Nardi (Eds.), *Proceedings of the 26<sup>th</sup> Conference of the International Group for the Psychology of Mathematics Education*, (Vol. 3, pp. 361–368). Norwich, UK: University of East Anglia.
- Monteiro, C. (2005). *Investigating critical sense in the interpretation of media graphs* (Unpublished doctoral dissertation). University of Warwick, Warwick, U.K. [Online: [wrap.warwick.ac.uk/73122/](http://wrap.warwick.ac.uk/73122/)]
- Monteiro, C. & Ainley, J. (2007). Investigating the interpretation of media graphs among students. *International Electronic Journal of Mathematics Education, Special Issue: Emerging Research in Statistics Education* 2(3), 187–207.
- Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 257–315). Charlotte, NC: NCTM.
- Piaget, J. (1974). *Biology and knowledge: An essay on the relations between organic regulations and cognitive processes*. Chicago: University of Chicago Press.
- Shaughnessy, J., Garfield, J., & Greer, B. (1996). Data handling. In A. Bishop, K. Clements, C. Keitel, & C. Laborde (Eds.), *International handbook on mathematics education* (pp. 205–237). Dordrecht: Kluwer Academic Publishers.
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge; MA: Harvard University Press.
- Wallon, H. (1942). *De l'acte à la pensée, essai de psychologie comparée*. Paris: Flammarion.
- 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. 277–294). Dordrecht, The Netherlands: Kluwer.
- Waiselfisz, J. J. (2011). *Mapa da Violência 2011. Os Jovens do Brasil. [Map of violence 2011. The young people of Brazil]*. São Paulo: Instituto Sangari; Brasília, DF: Ministério da Justiça. [Online: [www.sangari.com/mapadaviolencia/pdf2011/MapaViolencia2011.pdf](http://www.sangari.com/mapadaviolencia/pdf2011/MapaViolencia2011.pdf)]
- Wild, C. & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (with discussion and a response from the authors). *International Statistical Review*, 67(3), 223–265.
- Zieffler, A., Garfield, J., Delmas, R., Le, L., Isaak, R., Bjornsdottir, A., & Park, J. (2011). Publishing in SERJ: an analysis of papers from 2002–2009. *Statistics Education Research Journal*, 10(2), 5–26. [Online: [iase-web.org/Publications.php?p=SERJ\\_issues](http://iase-web.org/Publications.php?p=SERJ_issues)]

CARLOS MONTEIRO

Department EDUMATEC, Federal University of Pernambuco  
Avenida da Arquitetura s/n, Cidade Universitária, Recife PE, 50.740-550  
Brazil