The measurement of statistical reasoning in verbal-numerical and graphical forms: a pilot study

This content has been downloaded from IOPscience. Please scroll down to see the full text.
2013 J. Phys.: Conf. Ser. 459 012023
(http://iopscience.iop.org/1742-6596/459/1/012023)

View the table of contents for this issue, or go to the journal homepage for more

Download details:
IP Address: 182.254.226.161
This content was downloaded on 27/06/2016 at 05:17

Please note that terms and conditions apply.
The measurement of statistical reasoning in verbal-numerical and graphical forms: a pilot study

M Agus 1, M Peró-Cebollero2, J Guàrdia-Olmos2, M P Penna1

1Department of Pedagogy, Psychology, Philosophy, Faculty of Humanistic Studies, University of Cagliari, Via Is Mirrionis, 1, 09123 Cagliari, Italy

2Department of Methodology of the Behavioural Sciences, Faculty of Psychology, University of Barcelona, Pg. Vall d’Hebron, 171, 08035 Barcelona, Spain

E-mail: mirian.agus@unica.it

Abstract. Numerous subjects have trouble in understanding various conceptions connected to statistical problems. Research reports how students’ ability to solve problems (including statistical problems) can be influenced by exhibiting proofs. In this work we aim to contrive an original and easy instrument able to assess statistical reasoning on uncertainty and on association, regarding two different forms of proof presentation: pictorial–graphical and verbal–numerical. We have conceived eleven pairs of simple problems in the verbal–numerical and pictorial–graphical form and we have presented the proofs to 47 undergraduate students. The purpose of our work was to evaluate the goodness and reliability of these problems in the assessment of statistical reasoning. Each subject solved each pair of proofs in the verbal–numerical and in the pictorial–graphical form, in different problem presentation orders. Data analyses have highlighted that six out of the eleven pairs of problems appear to be useful and adequate to estimate statistical reasoning on uncertainty and that there is no effect due to the order of presentation in the verbal–numerical and pictorial–graphical form.

1. Introduction

In recent research we observe a strong interest in statistical reasoning and in judgment making in uncertainty situations [1,2]. Generally the concept of statistical reasoning is mentioned as the manner in which individuals reflect on statistical concepts and form conclusions on statistical data [3]. Garfield [2] suggests that six kinds of reasoning (on data, on representations of data, on association, on statistical measures, on samples, and on uncertainty) can be recognised in statistical reasoning. Each type of reasoning can be elicited by specific activities related to data comprehension, analysis, and communication. Specifically, the reasoning of association is important in statistics because this process is used to identify statistical relationships between two or more variables [2]. Reasoning on uncertainty is defined as the ability to apprehend and apply concepts of probability, casualness and chance, such as the capability to make conclusions with uncertainty, and with an awareness that not all results are obtained in the same way [3,1]. But many studies show that individuals make mistakes when they are reasoning on association and on probabilities [4]. Subjects appear not to be familiar with statistical association and probabilistic information [5,6]. For example, when individuals try to use uncertainty information, they tend to incorrectly estimate the probability of events; this occurs for both expert and non expert individuals [7]. Referring to these aspects, the use of several visual aids
could appear as an important method for clarifying and improving reasoning on uncertainty and on association [8,9]. Green [10] states that graphs could elicit an in-depth understanding of data complexity. On the other hand, Knauff and Johnson-Laird [11] affirm that visual aids could impede probabilistic reasoning in some problems, because they could overload the cognitive system with additional information.

2. Aim and method

This work represents the middle stage in our research which investigates the use of imagery in statistical reasoning, in order to detect the effect of graphical facilitation [12]. In this phase of the research we aim to explore statistical reasoning on uncertainty in association with appropriate proofs, in relation to two different forms of proof presentation: verbal–numerical and pictorial–graphical. We appoint a measure instrument of these articulated dimensions, which presents several pairs of the analogous proof; in this way we can compare, in the same individual, the statistical reasoning developed in comparable problems presented in two forms. We intend to analyse the reasoning in subjects without statistical skills, in order to evaluate the real contribution of the presentation method in the development of statistical reasoning. Then, we investigate reasoning on uncertainty and on association, with reference to basic mathematical abilities learnt at high school. Therefore we decided to exclude from our assessment those types of statistical reasoning which require a specific statistical knowledge.

In this pilot study we apply specific statistical analyses in order to detect the adequacy of proofs in this investigation. Moreover, in our work we aim at evaluating the potential effect of the order of presentation between proofs in verbal–numerical and pictorial–graphical form.

2.1 Participants, instrument and procedure

Participants consisted of 47 undergraduates at the University of Cagliari (Italy), attending the first year of the degree course in Psychology. The average age of the participants was 22.2 years (SD = 5.7). The group consisted of 32 females (68.1%) and 15 males (31.9%). Data collection took place in December 2012. The sample involved students who voluntarily participated in the research and the sampling was non probabilistic. We submitted for resolution eleven pairs of analogous items in paper format, to a large group in a classroom.

Each item was given to the same student to be solved using two modalities: verbal–numerical (form A) and pictorial–graphical (form B). Each problem necessitated an open response, where the subjects were required to explain the reasoning applied in the resolution. We analysed the students’ answers as given to each problem and then in each pair of problems and we coded the correct and the incorrect answers (1/0). In the Appendix we show some exemplificative pairs of items in two dimensions.

Double forms of problems were presented in two separate sessions (one week apart), in different orders of subgroup students. We analysed the students’ open answers and then we coded the responses as correct or incorrect. We computed the descriptive statistics the Bivariate Chi Square, the McNemar test, the Discrimination index, the Difficulty index, and the Alpha Reliability.

3. Results

In Figure 1 we show the percentages of correct responses in the two forms in each pair of items of two dimensions: uncertainty (pairs 1,2,3,4,5,6,7) and association (pairs 8,9,10,11). By this figure we can observe that our participants frequently had a better performance in the resolution of problems in the verbal–numerical form of each pair of proofs.

Next we applied a Bivariate Chi Square for each problem, in order to detect the effect of the order of presentation (verbal–numerical / pictorial–graphical; pictorial–graphical / verbal–numerical) in the correct/incorrect answers of each item (see Table 1). The sequence of the order of proof presentation did not have an effect on the number of correct responses; indeed we did not find significant effects by the application of these inferential statistics.
Table 1. Bivariate Chi Square: responses correct/incorrect by order of problem presentation

| Proof | Pearson Chi | Proof | Pearson Chi |
|-------|-------------|-------|-------------|
|       | Sqr. (df=1) |       | Sqr. (df=1) | p     |       | p     |
|       |             |       |             |       |       |       |
| **Reasoning on uncertainty** | | | | | | |
| FA_1  | 1.561       | FB_1  | 0.774       | 0.212 | 0.379 |
| FA_2  | 2.623       | FB_2  | 0.000       | 0.105 | 1.000 |
| FA_3  | 0.000       | FB_3  | 0.000       | 1.000 | 0.999 |
| FA_4  | 0.038       | FB_4  | 0.094       | 0.846 | 0.759 |
| FA_5  | 0.083       | FB_5  | 0.127       | 0.774 | 0.721 |
| FA_6  | 0.000       | FB_6  | 0.000       | 1.000 | 1.000 |
| FA_7  | 0.000       | FB_7  | 0.003       | 1.000 | 0.953 |
| **Reasoning on association** | | | | | | |
| FA_8  | 0.112       | FB_8  | 0.000       | 0.738 | 1.000 |
| FA_9  | 0.047       | FB_9  | 1.281       | 0.828 | 0.258 |
| FA_10 | 0.064       | FB_10 | 0.094       | 0.800 | 0.759 |
| FA_11 | 3.194       | FB_11 | 0.127       | 0.074 | 0.721 |

In assessing the dimension of reasoning on uncertainty for the verbal–numerical form we obtained an Alpha Reliability = 0.556 (St. Alpha = 0.553) and for the pictorial–graphical form an Alpha Reliability = 0.656 (St. Alpha = 0.666). Then, in evaluating the dimension of reasoning on association for the verbal–numerical form we had an Alpha Reliability = 0.357 (St. Alpha = 0.320) and for the pictorial–graphical form an Alpha Reliability = 0.416 (St. Alpha = 0.432). We also calculated the indices of Difficulty and Discrimination in order to evaluate the goodness of each proof (see Table 2 and Table 3). The reliability of the two dimensions was low; particularly in the dimension of reasoning by association some statistics appear inept in order to pursue our aim to construct an evaluation instrument of statistical reasoning.
Table 2. Descriptive statistics and indices on the verbal–numerical form

| Proof | Number of correct responses | Difficulty index | Discrimination index | Alpha deleting each item in turn | Std. Alpha deleting each item in turn | r (item, total) |
|-------|-----------------------------|------------------|----------------------|----------------------------------|--------------------------------------|----------------|
| FA_1  | 26                          | 0.553            | 0.866                | 0.454                            | 0.453                                | 0.424          |
| FA_2  | 8                           | 0.170            | 0.333                | 0.333                            | 0.547                                | 0.197          |
| FA_3  | 26                          | 0.553            | 0.666                | 0.454                            | 0.547                                | 0.259          |
| FA_4  | 22                          | 0.468            | 0.733                | 0.549                            | 0.547                                | 0.435          |
| FA_5  | 11                          | 0.234            | 0.533                | 0.503                            | 0.501                                | 0.322          |
| FA_6  | 29                          | 0.617            | 0.400                | 0.600                            | 0.589                                | 0.066          |
| FA_7  | 36                          | 0.766            | 0.333                | 0.516                            | 0.512                                | 0.284          |
| FA_8  | 27                          | 0.570            | 0.733                | 0.179                            | 0.161                                | 0.277          |
| FA_9  | 31                          | 0.660            | 0.268                | 0.268                            | 0.218                                | 0.211          |
| FA_10 | 42                          | 0.894            | 0.200                | 0.431                            | 0.430                                | 0.013          |
| FA_11 | 19                          | 0.404            | 0.733                | 0.226                            | 0.195                                | 0.243          |

Table 3. Descriptive statistics and indices on the pictorial–graphical form

| Proof | Number of correct responses | Difficulty index | Discrimination index | Alpha deleting each item in turn | Std. Alpha deleting each item in turn | r (item, total) |
|-------|-----------------------------|------------------|----------------------|----------------------------------|--------------------------------------|----------------|
| FB_1  | 29                          | 0.617            | 0.866                | 0.579                            | 0.593                                | 0.490          |
| FB_2  | 6                           | 0.128            | 0.400                | 0.628                            | 0.640                                | 0.352          |
| FB_3  | 28                          | 0.596            | 0.400                | 0.668                            | 0.671                                | 0.220          |
| FB_4  | 7                           | 0.149            | 0.266                | 0.627                            | 0.637                                | 0.350          |
| FB_5  | 9                           | 0.191            | 0.533                | 0.581                            | 0.581                                | 0.512          |
| FB_6  | 23                          | 0.489            | 0.933                | 0.554                            | 0.578                                | 0.555          |
| FB_7  | 30                          | 0.638            | 0.466                | 0.684                            | 0.692                                | 0.161          |
| FB_8  | 27                          | 0.574            | 0.800                | 0.429                            | 0.408                                | 0.166          |
| FB_9  | 15                          | 0.319            | 0.600                | 0.379                            | 0.379                                | 0.203          |
| FB_10 | 40                          | 0.851            | 0.400                | 0.345                            | 0.382                                | 0.240          |
| FB_11 | 9                           | 0.191            | 0.533                | 0.240                            | 0.276                                | 0.342          |

4. Discussion
This research presents a pilot study in order to develop a measurement instrument to assess undergraduate students’ statistical reasoning on uncertainty and on association, in relation to methods of proof presentation. By the construction of paired items in two forms, we could compare the reasoning applied to problem resolution, with regard to the specific problem structure.

With evaluation of the statistical analysis, we could presume that the dimension of reasoning on association did not have good indices; therefore these pairs of items could be discarded by our questionnaire [13]. In relation to the reasoning on uncertainty, we could assume that one item (FB 4) did not have a good Discrimination Index, therefore it is not appropriate for this measurement [13].
We could also assume that the pair corresponding to this item could be discarded by the instrument, using the remaining six couples for the assessment of reasoning on uncertainty.

Moreover, in order to improve the agility and the promptness of this instrument, we could prepare a new version characterized by closed response options for these six pairs of items; then each problem could be followed by an open question, inquiring as to the reasoning implemented in the proof solution. In this way we could reduce the solution time of the proofs and simultaneously investigate the cognitive resolution process applied.

Furthermore we have highlighted that the proof disposition could not affect the problem resolution. Nevertheless, in order to avoid and control this contingent effect, it could be useful to balance the two forms of proofs by a casual disposition of problems in the random sub-samples of students.

The results of this work could be useful regarding the intention to develop statistical teaching and to project appropriate learning experiences in order to improve statistical reasoning.

5. Appendix
Some exemplificative pairs of items in verbal–numerical format and pictorial–graphical format

5.1. Reasoning on uncertainty - VERBAL-NUMERICAL FORMAT
A factory produces electronic games, but not all work well. Of every 100 game products: 20 may have an electrical problem; 80 can function correctly. The company has developed control systems to identify faulty games; however, these systems do not work properly. In reality, half of the games with electrical problems continue in the production line where they are considered as functional.

If you randomly extract a game that has been addressed to the shops and evaluated free of defects, what is the probability that it is defective?

5.2. Reasoning on uncertainty - PICTORIAL–GRAPHICAL FORMAT
A factory that produces personal computers has problems in the production process. Some of the computers are defective (problems with the video card). Such problems are not always identified by the quality control and consequently some defective computers are sent forward in the production line. The graphic below shows this process. What is the probability that a computer addressed to the shops, evaluated free of defects, is defective?

![Figure 2. Pictorial graphical format – Reasoning on uncertainty](image)

5.3. Reasoning on association - VERBAL–NUMERICAL FORMAT
X and Y are two variables in the relationship between them. The values of these variables are listed below:

| X  | 1 | 2 | 3 | 4 |
|----|---|---|---|---|
| Y  | 1 | 8 | 27| 64|

Taking into account this information, define the mathematical expression that describes the relationship between the two variables.
5.4. Reasoning on association - PICTORIAL–GRAPHICAL FORMAT

A and B are two variables in the relationship between them. On the basis of the graphical representation below, can you identify the type of mathematical relationship between the two variables?

![Graphical representation](image)

**Figure 3.** Pictorial graphical format – Reasoning on association

### References

[1] Garfield J. Assessing statistical reasoning *Stat. Educ. Res. J.* 2003; 2 22-38

[2] Lavigne NC, Lajoie SP. Statistical reasoning of middle school children engaged in survey inquiry *Contemp. Educ. Psychol.* 2007; 32 630-66

[3] Garfield J, Chance B. Assessment in statistics education: Issues and challenges *Math. Think. Learn.* 2000; 2 99-125

[4] Ibrekk H, Morgan MG. Graphical Communication of Uncertain Quantities to Nontechnical People *Risk Anal.* 1987; 7 519-29

[5] Hoffrage H, Gigerenzer G. Using natural frequencies to improve diagnostic inferences *Acad. Med.* 1998; 73 538-48

[6] Van Dijk E, Zeelenberg M. The discounting of ambiguous information in economic decision making *J. Behav. Decis. Making* 2003; 16 341-52

[7] Shafir E, Tversky A. Thinking through uncertainty: Non consequential reasoning and choice *Cognit. Psychol.* 1992; 24 449-74

[8] Moro R, Bodanza GA. El debate acerca del efecto facilitador en problemas de probabilidad condicional: ¿Un caso de experimentación crucial? *Interdiscipl.* 2010; 27 163-74

[9] Brase GL. Pictorial representations in statistical reasoning *Appl. Cogn. Psychol.* 2009; 23 369-81

[10] Green H. Grapharti: A new visual summary of data, *J. Stat. Ed.* 2007; 15 3

[11] Knauf YM, Johnson-Laird PN. Visual imagery can impede reasoning *Mem. Cognition.* 2002; 30 363–71

[12] Penna MP, Agus M, Peró-Cebollero M, Guàrdia-Olmos J, Pessa E. The use of imagery in statistical reasoning by university undergraduate students: a preliminary study *Qual. Quant.* 2012; 1-15, doi:10.1007/s11135-012-9757-5

[13] Reynolds CR, Livingston RB, and Willson V. T V *Measurement and Assessment in Education* Boston: Pearson; 2009