A proposal for the measurement of graphical statistics effectiveness: Does it enhance or interfere with statistical reasoning?

M Agus¹, M P Penna¹, M Peró-Cebollero², J Guàrdia-Olmos²

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 studies have examined students’ difficulties in understanding some notions related to statistical problems. Some authors observed that the presentation of distinct visual representations could increase statistical reasoning, supporting the principle of graphical facilitation. But other researchers disagree with this viewpoint, emphasising the impediments related to the use of illustrations that could overcharge the cognitive system with insignificant data. In this work we aim at comparing the probabilistic statistical reasoning regarding two different formats of problem presentations: graphical and verbal–numerical. We have conceived and presented five pairs of homologous simple problems in the verbal numerical and graphical format to 311 undergraduate Psychology students (n=156 in Italy and n=155 in Spain) without statistical expertise. The purpose of our work was to evaluate the effect of graphical facilitation in probabilistic statistical reasoning. Every undergraduate has solved each pair of problems in two formats in different problem presentation orders and sequences. Data analyses have highlighted that the effect of graphical facilitation is infrequent in psychology undergraduates. This effect is related to many factors (as knowledge, abilities, attitudes, and anxiety); moreover it might be considered the resultant of interaction between individual and task characteristics.

1. Introduction

Over the last decades many studies have been carried out to evaluate the influences on problem solution produced by the overall disposition of the matter. Probabilistic statistical problems are a particular type of enigmas, that persons are requested to solve both in academic contexts and in everyday life. Some authors affirm that the application of some visual aids could appear as a capital way to enlighten statistical reasoning [e.g. 1]. In contrast, Knauff and Johnson-Laird [2] sustain that illustrations could inhibit reasoning in these dilemmas because they could burden the cognitive system with supplementary evidences, beside the point.
In the light of bibliography, our study attempts at evaluating and comparing the performance in some pairs of analogous problems, presented in verbal-numerical and graphical format. By this, we investigate the existence of the effect of ‘graphical facilitation’ versus ‘graphical impediment’, considering also the influence of others important dimensions identified in literature (visuo-spatial ability, numerical skill, attitude towards statistics and statistical anxiety) [e.g. 3,4].

2. Aim and method
The aim of this investigation is to propose an experimental protocol in order to explore whether graphical representations can facilitate or impede probabilistic statistical reasoning in psychology undergraduates. Therefore we assessed and compared the student’s performance in problem solving with two homologous formats for each problem: verbal-numerical and graphical.

2.1. Participants, instrument and procedure
We studied 311 undergraduates in Psychology, enrolled in the first year of the degree course in three universities: University of Cagliari (Italy) (n=64), University of Trieste (Italy) (n=92) and University of Barcelona (Spain) (n=155). The mean age of the participants was 20.8 ± 5.9 years of age (mean ± standard deviation) and consisted of 85 males (27.3 %). Data collection was acquired from September to October 2013. The samples involved students who voluntarily participated in the research (accidental sampling). We submitted our questionnaires in one session, in paper form, to a large group in a lecture room; every group of undergraduates compiled the questionnaires respectively in their native language. By means of several pilot studies [4,5,6,7] we constructed a specific questionnaire in order to evaluate probabilistic statistical reasoning, comprising five pairs of homologous problems in two formats (verbal numerical – N - and graphical - G). The problems were presented in different orders (verbal-numerical and graphical -NG -vs. graphical and verbal-numerical -GN) and different sequences. We chose to inspect the reasoning in “inexperienced” students (without statistical skills) in order to assess the authentic impact of the format for the problem presentation. For each problem (in both formats), we offered a brief explanation and then four closed response options (when only one was correct).

With the purpose of estimating the numerical and visuo-spatial prerequisites for the problem solution, we applied two related scales of the Intermediate Form of Primary Mental Abilities (P.M.A.) [8,9]. Furthermore, we applied the Survey of Attitudes Towards Statistics (SATS-28) [10] and the Statistical Anxiety Scale (SAS) [11] to inquire the relationships among attitudes and emotional aspects and the problem solving (both these questionnaires demonstrated the cross-country validity between Italian and Spanish samples) [12,13].

In this work we evaluated the differences in problem solving between countries, the potential effect of the order of presentation between in N and G format. But, especially, we attempted to assess the potential ‘graphical facilitation’ effect. For these purposes we computed the Bivariate Chi Square, the Multiple Correspondence Analysis and the Hierarchical Cluster Analysis.

3. Results
We evaluated the descriptive statistics in relation to the answers given to each pair of problems. In order to specifically evaluate the supposed ‘graphical facilitation’ effect, we recoded the answers in each pair of problems in a new variable, devising a typological index [14]. In fact, we attempted to identify more precisely the students that solved the problems only in the graphical format and not in the other ways of presentation. For these reasons we constructed a new categorical variable for each pair of problem and we distinguished between: subjects who didn’t give correct answers in correspondence to both presentation formats, subjects who gave correct answers only in the verbal-numerical format, subjects who gave correct answers only in the graphical format, subjects who gave
correct answers in both presentation formats. These new indices classified the answers in two formats in a single categorical variable for each pair of problems, individuating a profile of all potential arrays of answers in the problem presentation.

In relation to these new indices, we applied the Bivariate Chi Square to detect the potential effect of the country (Italy and Spain); this analysis didn’t show a significant effect ($p > .05$) then we decided to analyse all the students as a unique sample.

In Figure 1 we can observe the distribution frequencies for each index, in relation to all five pairs of problems; these indices are evaluated regarding the different orders of problems presentation (NG versus GN). Afterwards we computed a Bivariate Chi Square for each profile to distinguish the effect of the presentation order (NG versus GN) in the answers given in all five problems. The problem presentation order did not have an effect on the responses in relation to profile 1, profile 3 and profile 5. Indeed we found significant effects in relation to profile 2 ($\chi^2 = 11.166; df=3; p=.011$) and 4 ($\chi^2=18.348; df=3; p=.0001$). Especially in these two profiles we observed a higher number of subjects that did not solve in both forms in the GN order and in the NG order.

**Figure 1.** Frequencies for response profile in relation to each five pairs of problems

We decided to evaluate all the categorical indices related to five response profiles as extensive dimensions applying a Multiple Correspondence Analysis (MCA) [e.g. 15], using the software SPAD 5.5. We referred to the *Ecole française d'analyse des données* [e.g. 16,17]. This method allows exploring and resuming the categorical responses in relation to few dimensions, investigating the configuration of some categorical variables. It can be a powerful statistical tool because it reveals groups of variables and modalities in the dimensional spaces, without requiring fulfilling the stringent assumptions demanded by other statistical techniques largely used. This statistical method allows distinguishing among “active” and “supplementary” variables. The first ones define the space solution and the factor construction; the second ones do not affect the axis orientation, but complete and sustain the interpretation of dimensions identified by active variables.

Then in this technique we used as ‘active’ variables the categorical indices obtained in relation to each pair of probabilistic problems (five categorical variables associated to a four modalities each one,
as previously indicated). As ‘supplementary’ variables we used the categorical variables related to personal data (gender, age, country), to the order of problem presentation (Numerical Graphical NG / Graphical Numerical GN), to the levels of abilities (numerical and visuospatial), to the levels of attitudes towards statistics and to the levels of statistical anxiety.

In the MCA we chose to consider two factors. The first could be identified as the dimension of ‘probabilistic statistical reasoning’ (eigenvalue = 0.4514; % intertia= 15.05). It was characterised by the active modalities of good solution in both formats for all problems. Moreover, as supplementary variables, it was associated in the positive semi-axis with a high level of visuo-spatial abilities, high time solution in both formats, low levels of statistical anxiety and high positive attitudes towards statistics. In the negative semi-axis the active and supplementary modalities show the opposite trend (see Table 1).

| Table 1. MCA – Factor 1 |
|-------------------------|
| **Positive semi axis**   |
| active modalities       | V. test | Factorial coordinates | Contributions | Cos² |
| P2 SOLVE N & G          | 13.01   | 1.17                   | 17.2          | .54  |
| P4 SOLVE N & G          | 11.32   | .96                    | 12.6          | .41  |
| P5 SOLVE N & G          | 10.76   | 1.23                   | 13.2          | .37  |
| P1 SOLVE N & G          | 9.23    | .64                    | 7.2           | .27  |
| P3 SOLVE N & G          | 7.96    | .89                    | 7.2           | .20  |

| **Negative semi axis**  |
| active modalities       | V. test | Factorial coordinates | Contributions | Cos² |
| P4 NOT SOLVE N & G      | -10.39  | -.79                   | 9.9           | .35  |
| P2 NOT SOLVE N & G      | -9.97   | -.58                   | 7.2           | .32  |
| P5 NOT SOLVE N & G      | -9.88   | -.57                   | 7.1           | .31  |
| P3 NOT SOLVE N & G      | -8.45   | -.59                   | 6.1           | .23  |
| P1 NOT SOLVE N & G      | -7.50   | -1.28                  | 7.2           | .18  |

These results should be read as an evidence that lead us to suppose that the ‘graphical facilitation’ effect could subsist particularly when the subject know the structure of problem (previously presented and solved correctly in the verbal numerical format).

The second factor could be recognised as the dimension of ‘graphical facilitation’ vs ‘graphical impediment’ (eigenvalue = 0.2764; % intertia= 9.21). Indeed it is characterised in the positive semi-axis by the active modalities for good solution only in the graphical format of all problems. As supplementary variables it is related with a low time for solution in the graphical format. The active and supplementary modalities display the reverse tendency in the negative semi-axis (see Table 2).

On the two factors resulting from the MCA, we applied the Hierarchical Cluster analysis to find undergraduate groups having analogous features, to group subjects, referring to the factors identified [e.g. 18]. This method allows us to classify the students’ groups based on their performance in all pairs of problems, in order to identify the meaningful taxonomies. The 5-membership clustering was deliberated to offer the best cluster solution. This solution individuated students’ groups that are good solvers in both formats of all problems (cluster 5, n=99) and bad solvers in both formats of all problems (cluster 2, n=133). But moreover three clusters were characterised by a minor number of
students that experiment the ‘graphical facilitation’ effect in all problems (cluster 4, n=29) and the ‘graphical impediment’ (cluster 1, n=41 and cluster 3, n=11). Especially cluster 3 is interesting because it consists in a small number of students that experiment a ‘very strong graphical impediment’. Generally, we could observe that the number of students that experience ‘graphical impediment’ is higher than the number of subjects that check out the ‘graphical facilitation’ effect.

Table 2. MCA – Factor 2

| Positive semi axis | V. test | Factorial coordinates | Contributions | Cos² |
|--------------------|--------|-----------------------|---------------|------|
| P5 NOT N, SOLVE G  | 9.54   | 1.69                  | 19.1          | .29  |
| P2 NOT N, SOLVE G  | 8.52   | 1.60                  | 15.4          | .23  |
| P4 NOT N, SOLVE G  | 8.13   | .88                   | 12.0          | .21  |
| P3 NOT N, SOLVE G  | 3.23   | .24                   | 1.6           | .03  |
| P1 NOT N, SOLVE G  | 2.65   | .16                   | .9            | .02  |

| Negative semi axis | V. test | Factorial coordinates | Contributions | Cos² |
|--------------------|--------|-----------------------|---------------|------|
| P3 SOLVE N, NOT G  | -8.56  | -2.43                 | 16.4          | .24  |
| P1 SOLVE N, NOT G  | -8.39  | -2.49                 | 15.8          | .23  |
| P4 SOLVE N, NOT G  | -6.77  | -1.02                 | 9.3           | .15  |
| P5 SOLVE N, NOT G  | -4.30  | -.46                  | 3.3           | .06  |

4. Discussion

In our work we aim at evaluating the effect of ‘graphical facilitation’ vs ‘graphical impediment’ in probabilistic statistical reasoning, in psychology undergraduates without statistical expertise. For this reason we devised an experimental protocol characterised by five pairs of analogous probabilistic problems in verbal numerical and graphical format. The presentation of pairs of problems and their data analyses highlighted that only few students could advantage of the ‘graphical facilitation’ effect in the solution of these homologous of problems on probabilistic statistical reasoning [19]. It appears infrequent for our psychology undergraduates to experiment the effect of ‘graphical facilitation’. Our results suggest the relevance of knowledge of the problem structure in facilitating the problem solving in both formats. Also the role of prerequisites, attitudes, and anxiety is shown, because bad solvers in both formats are characterised by high levels of statistical anxiety and low level of visuo-spatial ability. Moreover we observed that the good solvers in both formats had low levels of anxiety and high positive attitudes towards statistics. These facts are important because these students did not have a statistical education.

Our results might be also related to the subject’s learning style; particularly we refer to the ‘visual-spatial learners’ that are characterised by strong visual-spatial abilities and weaknesses in verbal information sequential processing [e.g. 20]. In fact, some authors recognise that the application of visual-spatial representation could be useful for the ‘visual-spatial learners’, but not for subjects with different learning styles [e.g. 20]. This aspect might explain both the effect of graphical facilitation [19] and the effect of graphical impediment [2]. Moreover our results support the theoretical concept of the individual-task combination [3] which emphasised that the probabilistic statistical reasoning could be considered consistent with a multifactorial relation among many features (the tasks
organisation in relation to the individual idiosyncrasy). The outcomes of this work could be useful regarding the intention to develop statistical teaching and to elaborate appropriate learning experiences to improve statistical reasoning. At this point, the assessment of subjects’ characteristics seems to be advantageous for the enhancement of statistical reasoning, in relation to the application of graphical methods. In fact, the effect of ‘graphical facilitation’ cannot be considered as neglected by interaction with this multiplicity of factors.

References
[1] Brase G L 2009 Pictorial representations in statistical reasoning Appl. Cogn. Psychol. 23 369-81. Doi:10.1002/acp.1460
[2] Knauff M and Johnson-Laird P N 2002 Visual imagery can impede reasoning Mem. Cognition 30 363–71. Doi:10.3758/BF03194937
[3] Zhu L and Gigerezen G 2006 Children can solve Bayesian problems: The role of representation in mental computation. Cognition 98 3 287–308. Doi: http://dx.doi.org/10.1016/j.cognition.2004.12.003
[4] Penna M P, Agus M, Peró-Cebollero M, Guàrdia-Olmos J and Pessa E 2014 The use of imagery in statistical reasoning by university undergraduate students: a preliminary study Qual. Quant. 48 1 173-187. Doi:10.1007/s11135-012-9757-5
[5] Agus M, Penna M P, Peró-Cebollero M, Guàrdia-Olmos J and Pessa E 2014 The application of graphical representations in estimation of probabilistic events J. Th. Res. Ed. 9 1 235-252. Doi: 10.6092/issn.1970-2221/4299
[6] Agus M, Peró-Cebollero M., Guàrdia-Olmos J and Penna MP 2013 The measurement of statistical reasoning in verbal-numerical and graphical forms: a pilot study. J.of Phys.Conf. Ser. 459 12023. Doi:10.1088/1742-6596/459/1/012023
[7] Agus M, Peró-Cebollero M, Penna M P and Guàrdia-Olmos J. 2014 Towards the development of problems comparing verbal-numerical and graphical formats in statistical reasoning. Qual Quant.1–19. Doi: 10.1007/s11135-014-0018-7
[8] Thurstone LL and Thurstone TG 1981a PMA: abilità mentali primarie: manuale di istruzioni K-1 - Batteria fattoriale delle abilità mentali primarie (Firenze: Organizzazioni Speciali)
[9] Thurstone LL and Thurstone TG 1981b TEA: tests de aptitudes escolares [niveles 1, 2 y 3]: manual (Madrid: Tea).
[10] Schau C, Stevens J, Dauphinee TL and Vecchio AD 1995 The Development and Validation of the Survey of Attitudes toward Statistics. Educ. Psychol. Meas. 55 5 868–875.
[11] Colet AV, Seva UL and Condon L 2008 Development and validation of the Statistical Anxiety Scale. Psicotherma 20 1 174–180.
[12] Carmona J, Primi C, Chiesi F. 2008 Testing for measurement invariance of the Survey of Attitudes Toward Statistics: A comparison of Italian and Spanish students. III European Congress of Methodology, Oviedo, Spain.
[13] Chiesi F, Primi C, Carmona J. 2011 Measuring Statistics Anxiety. J Psychoeduc Assess 29(6) 559–69.
[14] Marradi A 2007 Metodologia delle scienze sociali (Bologna: Il Mulino)
[15] Greenacre M 2010 Correspondence analysis in practice. (NW: Chapman and Hall/CRC).
[16] Benzécri J P, 1992 Correspondence analysis handbook. (CRC Press).
[17] Lebart L, Morineau A and Piron M. 1995 Statistique exploratoire multidimensionnelle. (Paris: Dunod).
[18] Di Franco G. 2006 Corrispondenze multiple e altre tecniche multivariate per variabili categoriali. (Milano: FrancoAngeli).
[19] Moro R and Bodanza G A 2010 El debate acerca del efecto facilitador en problemas de probabilidad condicional: ¿Un caso de experimentación crucial? Interdiscipl. 27 163-74
[20] Silverman L K. 1989 The Visual-Spatial Learner. Prev Sch Fail Altern Educ Child Youth 34(1) 15–20.