# Probabilistic reasoning under time pressure: an assessment in Italian, Spanish and English psychology undergraduates

**M Agus[1], P K Hitchcott[2], M P Penna[1], M Peró-Cebollero[3], J Guàrdia-Olmos[3]**
[1]Department of Pedagogy, Psychology, Philosophy, Faculty of Humanistic Studies, University of Cagliari, Via Is Mirrionis, 1, 09123 Cagliari, Italy
[2]Department of Psychology, Southampton Solent University, East Park Terrace, Southampton, SO14 0YN, UK
[3]Department 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**. Many studies have investigated the features of probabilistic reasoning developed in relation to different formats of problem presentation, showing that it is affected by various individual and contextual factors. Incomplete understanding of the identity and role of these factors may explain the inconsistent evidence concerning the effect of problem presentation format. Thus, superior performance has sometimes been observed for graphically, rather than verbally, presented problems. The present study was undertaken to address this issue. Psychology undergraduates without any statistical expertise (N = 173 in Italy; N = 118 in Spain; N = 55 in England) were administered statistical problems in two formats (verbal-numerical and graphical-pictorial) under a condition of time pressure. Students also completed additional measures indexing several potentially relevant individual dimensions (statistical ability, statistical anxiety, attitudes towards statistics and confidence). Interestingly, a facilitatory effect of graphical presentation was observed in the Italian and Spanish samples but not in the English one. Significantly, the individual dimensions predicting statistical performance also differed between the samples, highlighting a different role of confidence. Hence, these findings confirm previous observations concerning problem presentation format while simultaneously highlighting the importance of individual dimensions.

## 1. Introduction

Difficulties encountered by psychology undergraduates studying statistics have been documented [1,2]. The study of probabilistic reasoning provides an empirical basis for the development of effective interventions for improving statistical performance [3,4]. There is, however, ongoing debate around the factors that enhance vs impede the probabilistic reasoning is in the literature [5,6]. Some studies have highlighted that specific features of the problems (e.g., the use of graphical-pictorial illustrations) can enhance performance [5,7,8], while other studies have shown that mild time pressure also increases task accuracy [9,10]. Some progress has been made through the recognition that such factors may be either contextual (i.e., format of problem presentation, time pressure) or individual (i.e., abilities, anxiety, attitudes, confidence) [7,9] in nature. Previous evidence suggests a better understanding of the role of individual factors may resolve the effect of problem presentation format [9,10]. The present study was undertaken for this purpose.

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

## 2. Aim and method

This investigation assessed the accuracy of probabilistic reasoning under a time pressure condition. In similar previous studies we have assessed student performance with and without time pressure and observed that the presence of time pressure does not increase anxiety, but instead enhances performance in Italian undergraduates [9–11]. We matched the performances of the same student through the use of similar problems administered in two formats (verbal-numerical – N – and graphical- pictorial – G), attempting to detect the potential effect of graphical facilitation. Three samples (Italian, Spanish and English) were tested, in part, to establish the generality of the findings. All students were psychology undergraduates without prior statistical training. A key stimulus for the present study was the observation that the metacognitive dimension of confidence affected probabilistic reasoning performance [9]. Therefore, in order to confirm and extend this observation all students also completed measures assessing various individual characteristics (visuo-spatial and numerical abilities, attitudes towards statistics, statistical anxiety and confidence in the correctness of response). We hypothesized that (1) an effect of graphical facilitation would be observed in all three samples; and (2) that confidence would affect performance regardless of problem presentation format.

### *2.1. Participants, instrument and procedure*

We studied 346 undergraduates, enrolled in the first year of degree courses in Psychology, in three countries: Italy (University of "Sapienza", N = 173; female 60.1%), Spain (University of Barcelona, N = 118; female 72%), England (Solent University, N = 55; female 89.1%).

All participants were recruited via opportunity sampling. The mean age of the participants was in the Spanish sample 20.9 ± 6.4 years old (mean ± standard deviation), in the English sample m= 20.4 ± 5.4 years old, in the Italian sample m= 19.6 ± 1.6 years old. Participants were tested in a single session in large groups; all questionnaires were completed in pencil-and-paper format. All measures were translated into participants' native language; all responses were recorded in the participants' native language. We administered a specific instrument to inquire the probabilistic reasoning in two formats (N and G) [11,12], devised in the previous phases of our investigation. For each problem, participants were required to report their level of confidence in the accuracy of response. To evaluate participants' numerical and visuo-spatial abilities, we administered two scales of the Primary Mental Abilities Questionnaire (PMA) [13]. The Survey of Attitudes Towards Statistics (SATS-28) [14] and the Statistical Anxiety Scale (SAS) [15], were used to measure the participant attitudes and anxiety related to statistics; both these questionnaires demonstrated the cross-country validity between Italian and Spanish samples [16,17]. The descriptive statistics of the samples are showed in the Table 1.

To evaluate the role of the study variables in predicting performance (in the N and G formats, respectively), as in our previous work [9], two Multivariate Linear Regressions were applied, independently for each format and for each sample. Subsequently, a mixed design Ancova was carried out (separately for each sample), using as repeated measures the performances in two formats (N and G) [18].

## 3. Results

We calculated the descriptive statistics and the linear Pearson's r correlations between the variables. Consequently, to explore the effect of the variables in probabilistic reasoning, two Multivariate Linear Regressions (by using the Forward Method) were applied, by means of the number of correct responses (firstly in the N format as the criterion, and then, subsequently, the number of correct responses in the G format). The predictors were students' abilities, attitudes, anxiety and confidence (Table 2). The most significant predictor of performance in the Italian and Spanish samples was the confidence about the correctness of one's own responses in both formats, differently from the English students, which show a significant effect of attitudes towards statistics and of visuo-spatial ability.

**Table 1 – Descriptive statistics in each sample**

| | **Spanish sample** | | **English sample** | | **Italian sample^** | |
|---|---|---|---|---|---|---|
| **N** | 118 | | 55 | | 173 | |
| **Percentage of women** | 72 | | 89.1 | | 60.1 | |
| **Age Mean** | 20.920 | | 20.420 | | 19.690 | |
| **Age Standard deviation** | 6.451 | | 5.415 | | 1.638 | |
| **Lowest age** | 17 | | 17 | | 18 | |
| **Highest age** | 54 | | 53 | | 36 | |
| **University affiliation** | Barcelona | | Southampton | | Rome | |
| | **Spanish sample** | | **British sample** | | **Italian sample ^** | |
| | **M** | **SD** | **M** | **SD** | **M** | **SD** |
| **PMA VISUO-SPATIAL SCALE** | 24.805 | 11.410 | 20.314 | 13.123 | 23.595 | 11.695 |
| **PMA NUMERICAL SCALE** | 15.871 | 5.740 | 12.907 | 5.864 | 18.500 | 6.173 |
| **SATS AFFECT** | 21.347 | 4.866 | 19.723 | 4.689 | 19.432 | 5.336 |
| **SATS COMPETENCE** | 26.898 | 5.055 | 25.723 | 5.261 | 25.845 | 5.894 |
| **SATS VALUE** | 46.550 | 8.718 | 43.425 | 8.334 | 44.214 | 9.167 |
| **SATS DIFFICULTY** | 25.474 | 5.447 | 26.425 | 5.999 | 25.880 | 6.164 |
| **SAS EXAMINATION** | 32.855 | 6.244 | 31.028 | 8.392 | 28.542 | 9.714 |
| **SAS INTERPRETATION** | 19.711 | 6.461 | 21.117 | 6.830 | 15.505 | 6.572 |
| **SAS HELP** | 18.771 | 8.086 | 17.882 | 7.642 | 16.226 | 8.074 |
| **FN CORRECT RESPONSES** | 1.593 | 1.248 | 1.355 | 1.131 | 2.277 | 1.472 |
| **FG CORRECT RESPONSES** | 2.196 | 1.240 | 1.409 | 1.106 | 2.572 | 1.435 |
| **FN CONFIDENCE** | 3.216 | .944 | 2.899 | .751 | 3.248 | .962 |
| **FG CONFIDENCE** | 3.542 | 1.000 | 3.004 | .835 | 3.530 | .893 |

^ Data presented in Agus et al. [9]

Then we applied a Mixed Design Ancova, wherein the repeated measures were the performances in two formats (N and G), and the covariates were identified among the significant predictors recognized in the former regressions.

In the Spanish sample significant effects of Format [F= 8.204, df= 1;111, p=.005, $\eta^2$= .069, Observed Power=.810] and of covariates were detected [Format*Sats_Affect, F= 5.560, df= 1;111, p=.020, $\eta^2$= .048, Observed Power=.647; Format*Confidence_G, F= 4.104, df= 1;111, p=.045, $\eta^2$= .036, Observed Power=.519]. Indeed, for Spanish undergraduates we observed a higher number of correct responses in G format (m=2.217, sd=1.240) than in the N format (m=1.617, sd=1.246), highlighting an effect of graphical facilitation (Table 3).

Conversely, in the English sample no format effect and no effect of covariates were identified; it is interesting to observe that in this sample the number of correct responses is insignificantly higher in the N format (m=1.545, sd=1.148), than in the G format (m=1.393, sd=1.116) (Table 3).

A significant difference in the means of the correct responses in the N and G formats was found in the Italian students, as yet observed in our previous research [9]. They showed a significant effect of Format [F=8.736, df= 1;154, p=.004, $\eta^2$= .054, Observed Power=.836] and of Format*Confidence_N [F=9.513, df= 1;154, p=.002, $\eta^2$= .058, Observed Power=.865]. Indeed, they registered a better performance in the G format (m =2.791, sd =1.252) than in the N format (m=2.474, sd=1.408) (Table 3).

In summary, under a time pressure condition graphical facilitation was observed in Italian and Spanish samples, but not in the English sample. However, the computed effect sizes call for a degree of caution in interpreting certain findings (Figure 1).

**Table 2 – Multivariate Linear Regressions (Forward Method). Criterion variable: number of correct responses in N format and in G format**

| | $R_2$ | SE | Variables | β | t | p |
|---|---|---|---|---|---|---|
| **Criterion variable: number of correct responses in N format** | | | | | | |
| **Spanish sample** | .178 | 1.142 | MEAN_CONFIDENCE_N<br>SATS_AFFECT | .258<br>.252 | 2.812<br>2.744 | .006**<br>.007** |
| **English sample** | .481 | .949 | PMA_VISUO-SPATIAL<br>SATS_COMPETENCE | .522<br>.440 | 3.073<br>2.592 | .007**<br>.018* |
| **Italian sample ^** | .315 | 1.243 | PMA_NUMERICAL<br>MEAN_CONFIDENCE_N | .184<br>.481 | 2.222<br>5.795 | .028*<br>.0001** |
| **Criterion variable: number of correct responses in G format** | | | | | | |
| **Spanish sample** | .129 | 1.157 | MEAN_CONFIDENCE_G | .359 | 4.089 | .0001** |
| **English sample** | .394 | .846 | SATS_DIFFICULTY | .627 | 3.780 | .001** |
| **Italian sample ^** | .218 | 1.119 | MEAN_CONFIDENCE_G | .467 | 5.464 | .0001** |

*Note:* ** p<.05;* *** p<.001* ^ Data presented in Agus et al. [9]

**Table 3 – Results of mixed design Ancova**

| Sample | Source | Wilks' Lambda | df | F | p | $\eta^2$ | N FORMAT M (SD) | G FORMAT M (SD) |
|---|---|---|---|---|---|---|---|---|
| **Spanish sample** | FORMAT | .931** | 1 | 8.204 | .005** | .069 | 1.617 (1.246) | 2.217 (1.240) |
| | FORMAT * SATS AFFECT | .952* | 1 | 5.560 | .020* | .048 | | |
| | FORMAT* CONFIDENCE N | .972 | 1 | 3.153 | .079 | .028 | | |
| | FORMAT*CONFIDENCE G | .964* | 1 | 4.104 | .045* | .036 | | |
| | ERRORS | | 111 | | | | | |
| **English sample** | FORMAT | 1.000 | 1 | .000 | .998 | .000 | 1.545 (1.148) | 1.393 (1.116) |
| | FORMAT * VISUO-SPATIAL PMA | .999 | 1 | .017 | .897 | .001 | | |
| | FORMAT * SATS COMPETENCE | 1.000 | 1 | .009 | .924 | .000 | | |
| | FORMAT * SATS DIFFICULTY | 1.000 | 1 | .001 | .982 | .000 | | |
| | ERRORS | | 29 | | | | | |
| **Italian sample ^** | FORMAT | .946** | 1 | 8.736 | .004** | .054 | 2.474 (1.408) | 2.791 (1.252) |
| | FORMAT * NUMERICAL PMA | .979 | 1 | 3.241 | .074 | .021 | | |
| | FORMAT* CONFIDENCE N | .942** | 1 | 9.513 | .002** | .058 | | |
| | FORMAT*CONFIDENCE G | .981 | 1 | 2.987 | .086 | .019 | | |
| | ERRORS | | 154 | | | | | |

*Note:* M = mean; SD= standard deviation; **p < .01 ^Data presented in Agus et al. [9]

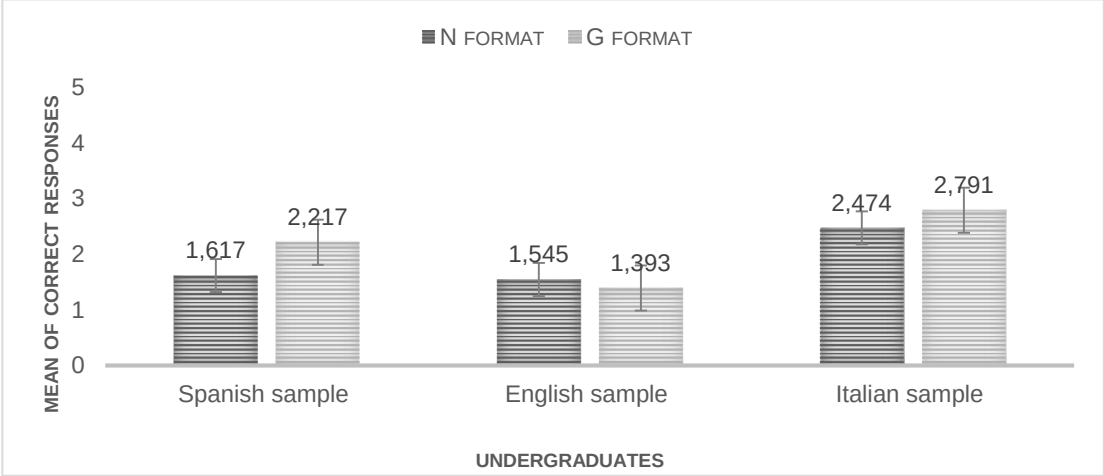



**Figure 1.** Means of correct responses in N and G formats in the samples.

## 4. Discussion

In broad terms the present data confirm and extend previous findings that have pointed to similarities between Italian and Spanish students [9]. Thus, in these two samples performance was affected by the confidence (related to the metacognitive dimension) in verbal-numerical and graphical formats and graphical facilitation effect was observed [9]. In contrast English undergraduates showed different characteristics in that their performances were affected by attitudes and abilities and not by the confidence [9]; moreover, they did not show graphical facilitation. These dissimilar outcomes might be related to the presence of effective differences among these populations, highlighted in the scores obtained in our assessments in the PMA and also in the PISA assessments [19] (i.e., the mean score of performance in mathematics is similar in Italy and Spain, but different in UK). In addition, differences in the pre-university teaching of probabilistic reasoning in these countries, including the manner in which probabilistic data are presented in text books (*e.g.,* the use of tree diagrams were used) could be a factor [20–24]

Despite the interesting pattern of findings, the results need to be approached with some degree of caution because of the low number of English students that participated in our research. Nevertheless, these preliminary findings suggest that the investigation of the probabilistic reasoning is a complex matter, having specific and different characteristics compared to the mathematical performance. Previous works highlighted that probabilistic reasoning is influenced by a wide array of interacting factors concerning the individual and the context. The inconsistent findings about the effect of graphical facilitation may be because the importance of some individual and contextual factors has been overlooked. In this investigation, the role of individual factors appears crucial; then, it seems reductive to talk about the effect of graphical facilitation in probabilistic reasoning related only to the presentation of the graphical devices, but it might be interesting to deepen the interactions among some individual factors (i.e., abilities, metacognition, anxiety and attitudes) [9,11]. Such evidences looks essential because our participants did not have any statistical education. Then, as consequence, in order to improve the probabilistic reasoning performance, it might be useful to apply the evaluation of individual specificities to support an adequate presentation of this type of problems.

## References

[1] Guàrdia-Olmos J, Freixa-Blanxart M, Peró-Cebollero M, Turbany J, Cosculluela A, Barrios M and Rifà X 2006 Factors related to the academic performance of students in the statistics course in psychology *Qual. Quant.* **40** 661–74

[2] Chiesi F and Primi C 2010 Cognitive and non-cognitive factors related to students' statistics achievement *Stat. Educ. Res. J.* **9** 6–26

[3] Sirota M, Kostovičová L and Vallée-Tourangeau F 2015 Now you Bayes, now you don't: effects of set-problem and frequency-format mental representations on statistical reasoning *Psychon. Bull. Rev.* 1–9
[4] Sirota M, Kostovičová L and Juanchich M 2014 The effect of iconicity of visual displays on statistical reasoning: evidence in favor of the null hypothesis. *Psychon. Bull. Rev.* **21** 961–8
[5] Brase G L and Hill W T 2015 Good fences make for good neighbors but bad science: a review of what improves Bayesian reasoning and why *Front. Psychol.* **6** 340
[6] Knauff M and Johnson-Laird P N 2002 Visual imagery can impede reasoning *Mem. Cognit.* **30** 363–71
[7] Mandel D R and Navarrete G 2015 Introduction to Improving Bayesian Reasoning: What Works and Why? *Front. Psychol.* **6**
[8] Garcia-Retamero R and Hoffrage U 2013 Visual representation of statistical information improves diagnostic inferences in doctors and their patients. *Soc. Sci. Med.* **83** 27–33
[9] Agus M, Peró-Cebollero M, Penna M P, Guàrdia Olmos J and Guàrdia-Olmos J 2015 Comparing Psychology Undergraduates' Performance in Probabilistic Reasoning under Verbal-Numerical and Graphical-Pictorial Problem Presentation Format: What is the Role of Individual and Contextual Dimensions? *Eurasia J. Math. Sci. Technol. Educ.* **11** 735–50
[10] Agus M, Penna M P, Peró-Cebollero M, Guàrdia-Olmos J and Pessa E 2016 Investigating the probabilistic reasoning in verbal–numerical and graphical–pictorial formats in relation to cognitive and non-cognitive dimensions: The proposal of a model *Pers. Individ. Dif.* **94** 44–53
[11] Agus M, Penna M P, Peró-Cebollero M and Guàrdia-Olmos J 2016 Assessing Probabilistic Reasoning in Verbal-Numerical and Graphical-Pictorial Formats: An Evaluation of the Psychometric Properties of an Instrument. *Eurasia J. Math. Sci. Technol. Educ.* **12** 2013–38
[12] Agus M, Peró-Cebollero M, Penna M P and Guàrdia-Olmos J 2015 Towards the development of problems comparing verbal-numerical and graphical formats in statistical reasoning *Qual. Quant.* **49** 691–709
[13] Thurstone L L 1948 *Primary Mental Abilities* (Chicago, IL: University Of Chicago Press)
[14] Dauphinee T L, Schau C and Stevens J J 1997 Survey of attitudes toward statistics: Factor structure and factorial invariance for women and men *Struct. Equ. Model. A Multidiscip. J.* **4** 129–41
[15] Vigil-Colet A, Lorenzo-Seva U and Condon L 2008 Development and validation of the statistical anxiety scale. *Psicothema* **20** 174–80
[16] Chiesi F, Primi C and Carmona J 2011 Measuring Statistics Anxiety. Cross-Country Validity of the Statistical Anxiety Scale (SAS) *J. Psychoeduc. Assess.* **29** 559–69
[17] Chiesi F and Primi C 2009 Assessing statistics attitudes among college students: Psychometric properties of the Italian version of the Survey of Attitudes toward Statistics (SATS) *Learn. Individ. Differ.* **19** 309–13
[18] Hair J F, Tatham R L, Anderson R E and Black W 2006 *Multivariate data analysis* vol 6 (Pearson Prentice Hall Upper Saddle River, NJ)
[19] Binder K, Krauss S and Bruckmaier G 2015 Effects of visualizing statistical information – An empirical study on tree diagrams and 2 x 2 tables *Front. Psychol.* **6**
[20] Phillips J L 1999 *How to think about statistics* (New York: W. H. Freeman)
[21] Tout D and Gal I 2015 Perspectives on numeracy: reflections from international assessments *ZDM* **47** 691–706
[22] Gal I 2005 Towards "probability literacy" for all citizens: Building blocks and instructional dilemmas *Exploring Probability in School* ed G A Jones (Springer) pp 39–63
[23] Agasisti T and Cordero-Ferrera J M 2013 Educational disparities across regions: A multilevel analysis for Italy and Spain *J. Policy Model.* **35** 1079–102
[24] Agasisti T and Pérez-Esparrells C 2010 Comparing efficiency in a cross-country perspective: the case of Italian and Spanish state universities *High. Educ.* **59** 85–103