

Classical item and test analysis with graphics: The ViSta-CITA program

RUBÉN DANIEL LEDESMA

Universidad Nacional de Mar del Plata, Mar del Plata, Argentina

AND

J. GABRIEL MOLINA

Universidad de Valencia, Valencia, Spain

Current advances in test development theory have mostly been influenced by item response theory. Notwithstanding this, classical test theory still plays a major part in the development of tests for applied educational and behavioral research. This article describes ViSta-CITA, a computer program that implements a set of classical item and test analysis methods that incorporate innovative graphics whose aim is to provide deeper insight into analysis results. Such an aim is achieved through the SpreadPlot, a graphical method designed to display multiple, simultaneous, interactive views of the analysis results. It behaves on a dynamic basis, so that users' changes (e.g., selecting a subset of items) are automatically updated in the graphical windows showing the analysis results. Moreover, ViSta-CITA is freely available, and its code is open to modifications or additions by the user. Features such as these constitute useful tools for research and teaching purposes related to test development.

Graphical techniques play an important role in exploring and modeling statistical data (Wainer & Velleman, 2001; Wilkinson, 1994, 2001; Yu & Behrens, 1995; Yu & Stockford, 2003). When they are well designed, graphics can support "a broad range of user tasks and abilities, are easy to learn, and provide powerful and flexible output formatting" (Colet & Aaronson, 1995, p. 257). This is especially true for modern dynamic and interactive graphics, which are widely implemented in software such as DataDesk or ViSta (Young, Valero-Mora, & Friendly, 2006).

The way in which dynamic and highly interactive graphics could be developed and structured in a complex visualization using the powerful XLispStat language and the ViSta Statistical System was discussed in a previous article (Ledesma, Molina, & Young, 2005). Along these lines, the present article presents a computer program, *ViSta-CITA* (classical item and test analysis with ViSta), oriented toward the analysis of some classical metric characteristics associated with test development. Thus, the program integrates a set of analysis methods rooted in classical test theory (CTT), a theoretical framework in test development that has served measurement specialists well for a long time. Even though modern test construction is more influenced by item response theory (IRT), CTT still plays an important part today due to its historical value and simplicity. Besides, it is common practice to teach CTT in graduate and undergraduate courses on measurement and research methods. Another important reason why CTT is widely applied in practice

is that it requires smaller samples to estimate the associated item parameters, as compared with IRT.

Needless to say, the purpose herein is not to provide a theoretical description of the CTT methods included in ViSta-CITA, since they are well known and have been documented in many textbooks (e.g., Crocker & Algina, 1986). Rather, it is to offer an overview of the program's main features and of how they work. Other, supplementary ViSta modules that can be useful in test development are also briefly covered in this article (e.g., parallel analysis, homogeneity analysis, and bootstrap procedures).

ViSta-CITA: A ViSta Plug-in for Classical Item and Test Analysis

ViSta-CITA was developed with the Lisp-Stat programming language and built as a plug-in for the ViSta statistical program (Young, 2006). ViSta, "the Visual Statistics System," is a free-distribution, open statistical program, originally set as a test bed for research and development in statistical visualization (Young & Smith, 1991). Ever since its origins in the early 1990s, ViSta has extended its field of application to offer a wide range of statistical methods (ANOVA, simple and multivariate regression, multidimensional scaling, and correspondence analysis, to name a few) and, more specifically, innovative graphical methods associated with the execution of these analyses, along with the visualization of their results (Molina, Ledesma, Valero, & Young, 2005).

R. D. Ledesma, rdledesma@gmail.com

Being an open system, ViSta's menu of features constantly increases as users interested in implementing new methods develop and integrate them into ViSta. In this respect, when adding new Lisp-Stat code to existing ViSta code, potential programmers can take advantage of the functions already available in this system, thereby avoiding unnecessary repetition. Moreover, since the creation of ViSta Version 6.4, the implementation of a program architecture based on plug-ins (Young, 1996) has favored the integration of new analysis methods into the system. This has allowed interested developers to create their own plug-in by just following a few specific rules when writing a code and then simply placing their program in the ViSta plug-ins folder so as to have it available in the ViSta environment (Ledesma et al., 2005; Young, 1996). ViSta-CITA is an example of a ViSta plug-in, and, in view of the fact that it is an open code development, users familiar with the Lisp-Stat programming language can extend or modify the ViSta-CITA code and either add new test analysis procedures or modify the existing ones to adapt them to specific requirements.

ViSta-CITA can be used to analyze some classical test properties (e.g., internal consistence, standard error of measurement, underlying test dimensionality), as well as some item properties (e.g., location/difficulty, discriminatory power). The program can also be used to compute several types of test scores, including raw scores, percentage scores, and absolute/normalized Z and T scores. However, the most distinctive feature of ViSta-CITA is the way in which some of this information is presented—that is, through the use of spreadplots. A spreadplot is a graphical method designed to reflect multiple, simultaneous views of the results of a statistical analysis (Young, Valero-Mora, Faldowski, & Bann, 2003). It consists of a container window including a linked list and plot windows. These windows show different aspects of the statistical analysis results in a dynamic way, so that changes by the user in one window result in automatic changes in the other windows linked to it.

It is worth noting that, in practice, item/test analysis often involves advancing through an exploratory, iterative process intended to refine the metric properties of the test being built. For instance, it is customary to successively delete items with poor performance for purposes of checking the effect of such deletion on the psychometric properties of the revised test. It is precisely in this step-by-step process that the dynamic and interactive character of spreadplots contribute to improve the quality of the test/item analysis results, as well as user experience with this type of statistical software.

The ViSta-CITA program

The ViSta-CITA module has been integrated into the latest version of ViSta, which is available from the Web site www.uv.es/visualstats/Book/. The program can also be compiled into the previous 6.4 version of ViSta by running the setup program available at www.mdp.edu.ar/psicologia/vista/vista.htm. After installing this plug-in, a new menu entry (Item-Test Analysis) appears in ViSta's Analyze menu.

Preparing the data for analysis. ViSta-CITA requires that a data set be formatted as ViSta data files. This requirement can be accomplished following either of these two options: (1) Data may be entered through the ViSta data editor (New Datasheet menu item), or (2) text or Excel data files can be imported using the ViSta Import Data menu item. The data must be numeric and complete (no empty cells). Should values in the data set be missing, ViSta's Input Missing Data command can be used to achieve any of the processing options of ViSta for this type of data. The data set may contain additional variables other than items; however, before proceeding with the analysis, the user must select the subset of variables representing the items to be analyzed by using the Variable Selector command (Data menu).

Figure 1 depicts a partial view of ViSta with an open data file appropriated for Item-Test Analysis. These data are available at ViSta's Data folder (JSI-Data.lsp file). The file contains data gathered from a job satisfaction study conducted in a public health institution. The variables correspond to the items of a Job Satisfaction Inventory (JSI) and were rated on the basis of a 7-point scale ranging from *strong disagreement* to *strong agreement*. Total scores are computed by adding up the 16 item scores. The observations correspond to a sample of health professionals from a public hospital in Mar del Plata, Argentina.

Running the analysis. To proceed with the ViSta-CITA analysis, the Item-Test Analysis command must be selected from the ViSta Analyze menu. A dialog box presenting the analysis options (see Figure 2) appears. The first three options provide summary statistics for item and test scores, as well as classical item-test indexes. The fourth, fifth, and sixth options, in turn, relate to the internal consistency analysis of the questionnaire, and the focus of the seventh option is on the analysis of dichotomous items, since items in binary format usually require specialized analysis (e.g., discrimination index based on point-biserial correlation, p values, and inter-item Tetrachoric correlations matrix). Finally, the eighth option provides some exploratory factor analyses (EFA) methods.

The Item-Test Analysis in the ViSta WorkMap. The WorkMap, the upper left window in ViSta's Desktop (see Figure 1), maps the steps taken by the program user during a data analysis session. This map is built and displayed as the analysis progresses. At first, WorkMap has no content, but, as users advance through their data analysis, icons representing data analysis steps are added to the WorkMap. These icons are also connected to previous icons by lines displaying the flow of data analysis. For example, Figure 3 shows that the data analyst first opened a data set ("JSI"), then applied an analysis model ("Test"), and finally obtained the analysis results ("Test-Res"). This last action is represented by an icon with two smaller icons added, which represent the two output approximations of ViSta—that is, reports and spreadplots. The three icons below correspond to new data sets created from the analysis model applied to the original data.

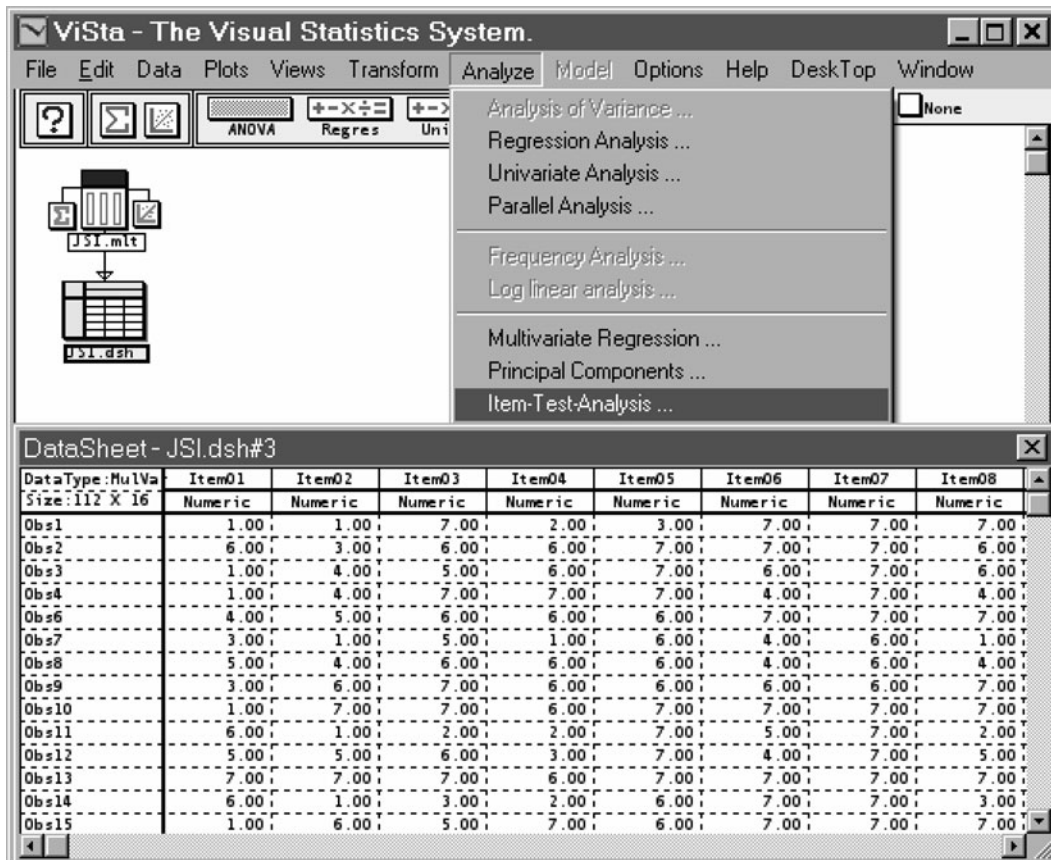


Figure 1. The Analyze menu extended over the ViSta Desktop window.

Obtaining reports and visualizations. When the user selects any of the analysis options, ViSta produces a specific analysis model for the data file just analyzed, which is represented as a new icon in ViSta's Desktop Workmap. The program user can examine an analysis model by utilizing one of the two strategies supported by ViSta: either a *report* or a *visualization* of the corresponding model. Hence, eight text-based reports and eight visualizations (SpreadPlots) were developed for the eight Item-Test Analysis options available in ViSta-CITA. By default, ViSta displays the analysis results with a report, a SpreadPlot, or both, depending on the preferences set by the program user through the command Options (Preferences menu).

Just as in many other traditional programs, a ViSta report consists of a list with the analysis results in text format. Users of SPSS, Statistica, or SAS may begin, rather, by analyzing the text-based report provided by ViSta, which, in fact, is very similar to that available for commercial statistical software. Such a report can be obtained by selecting the Report Model item from ViSta's Model menu. Figure 4 provides an example of a ViSta-CITA report—more specifically, the report for the reliability analysis based on Cronbach's alpha model, a type of analysis commonly used to evaluate, refine, and set the final version of a test. This report includes Cronbach's alpha index and the

corresponding 95% confidence interval; the alpha-based estimates of the standard error of measurement and the standard error of estimation; summary statistics based on the relationship between the items and the total test scores; and some descriptive statistics of each item if the item were separated from the test.

The visualization of each of the eight analysis options in ViSta-CITA (see Figure 2) is possible with a SpreadPlot obtained when the Visualize Model item is selected from ViSta's Model menu. Figure 5 shows the SpreadPlot corresponding to the second analysis option in Figure 2—that is, the SpreadPlot oriented to the provision of descriptive information on the test scores. Figure 5 displays some of the main features characterizing any ViSta SpreadPlot. The following information is furnished in the windows contained in this specific SpreadPlot.

1. A list of variables (upper left window) with the names of all the variables (items in this case) in the data set. In certain cases, this list can work as a control panel to select subsets of items and compute the analysis once again. If the user selects some of the items in the list and then clicks the "Update for the selected items" button in the "Update analysis & graphics" window, the rest of the SpreadPlot windows is updated, revealing the information corresponding to the selected items. By default, the SpreadPlot is created for all the items in the data set.

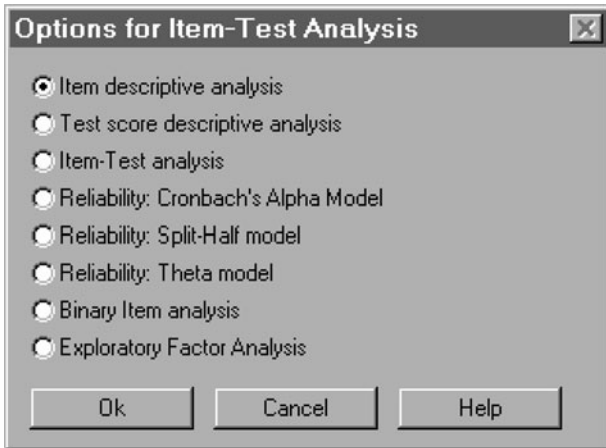


Figure 2. Item-Test Analysis options in ViSta-CITA.

2. A list of observations (lower left window) with labels for all the cases in the data set. This list allows the user to identify the selected cases in the graphical windows of the SpreadPlot and, also, to select and exclude some of them from the analysis. In the last case, the analysis is recomputed and the rest of the SpreadPlot windows are redrawn accordingly. By default, the SpreadPlot is created for all the cases in the data set.

3. A set of graphical windows with information of interest for the analysis model applied to the data. In the case of the SpreadPlot oriented to describe the subjects' scores in the test (see the example in Figure 5), a histogram, a box-dot-diamond plot, and a normal probability plot are included.

Figure 6 displays a more complex ViSta-CITA SpreadPlot, which aims at visualizing the results related to the reliability analysis based on Cronbach's alpha. Notice should be paid to the fact that, when the user selects a set of items

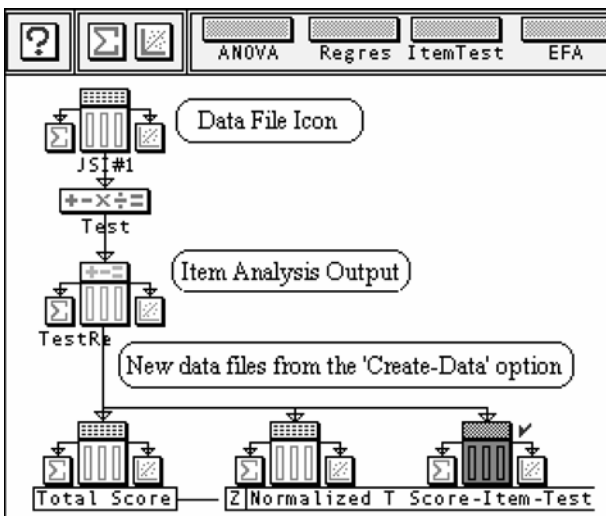


Figure 3. Item-Test Analysis session as represented in ViSta's WorkMap.

in the list of variables and then clicks the "Update for the selected items" button, the plots adapt themselves to the selected items, thereby enabling the program user to compare the reliability of different subsets of items. The horizontal line in the upper left window shows the alpha value for the test—in fact, for the test items selected in the list of variables—whereas the curved line shows the way in which this reliability index would increase as the test length increases n times, in agreement with the Spearman-Brown prophecy formula. The alpha if-item-deleted plot (lower right plot) illustrates how this coefficient is affected by each of the test items: The vertical axis represents the alpha value if the item were deleted from the test, and the horizontal black line the alpha value for all the items, which can be used as a reference to evaluate how each test item (represented as points) influences the reliability coefficient.

A third illustration of the SpreadPlots implemented in the ViSta-CITA application is provided in Figure 7. The purpose of this SpreadPlot is to visualize the split-half reliability analysis results. The main pieces in this SpreadPlot are a box-diamond-dot plot displaying the test score distributions of the two test halves (odd vs. even items) and a scatterplot showing the relationship between both halves (lower plot windows).

Creating data sets with the subjects' scores. ViSta-CITA allows users to create a new data set with the subjects' scores calculated on the basis of a number of approximations. This can be achieved through the Create Data menu item available in the Model menu of ViSta. Figure 8 illustrates the dialog box shown when this command is chosen. The options below are then available to create a data set.

1. *Total scores:* This creates a univariate data set containing the total test scores for all the subjects; total scores are computed as the sum of the item scores.

2. *Normalized Z scores:* This creates a univariate data set containing the total test scores expressed as normalized Z scores.

3. *Normalized T scores:* This creates a univariate data set containing the total test scores expressed as normalized T scores ($M = 50, SD = 10$).

4. *Total/Z/T scores if item deleted:* Each one of these options creates a multivariate data set with the total test scores if items are deleted one by one. This may be useful to compare test scores based on the different $k - 1$ subsets of items.

5. *Mean scores:* This creates a univariate data set containing the mean test scores (total test score divided by the number of items).

6. *Deviation scores:* This creates a univariate data set containing deviation scores from the mean test score.

7. *Estimated test scores:* This creates, by default, a univariate data set containing the estimated scores given the test mean, the standard deviation, and the alpha reliability coefficient. A ViSta-CITA user can modify these parameters by using the dialog box that appears when this option is selected.

8. *Confidence interval for the observed scores:* This creates a bivariate data set with the upper and lower limits

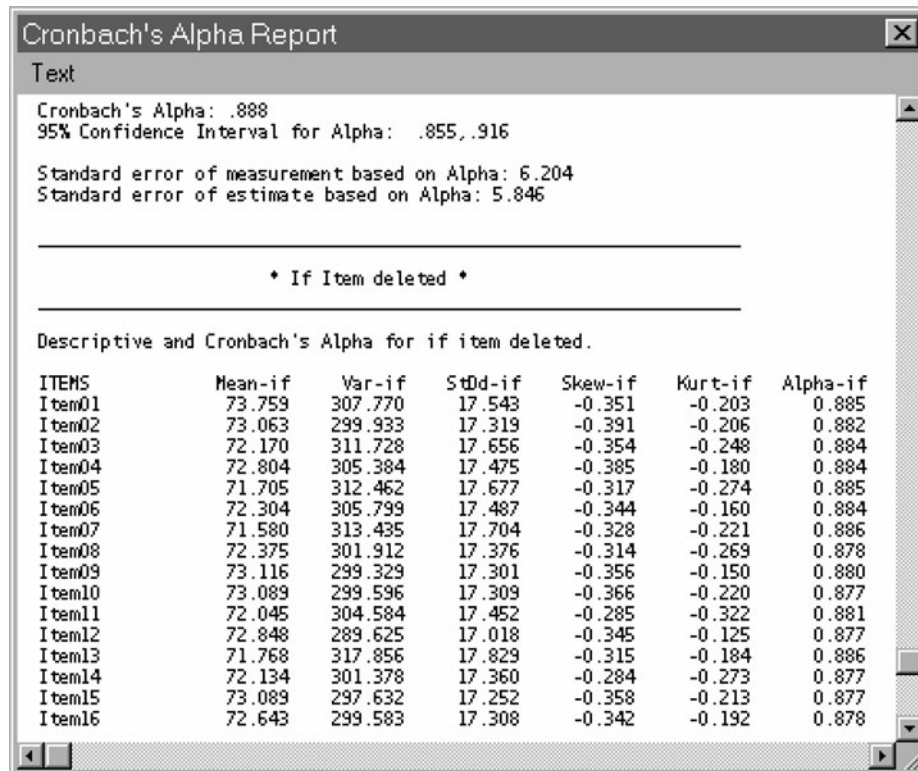


Figure 4. Example of a ViSta-CITA report for the Cronbach's alpha reliability analysis.

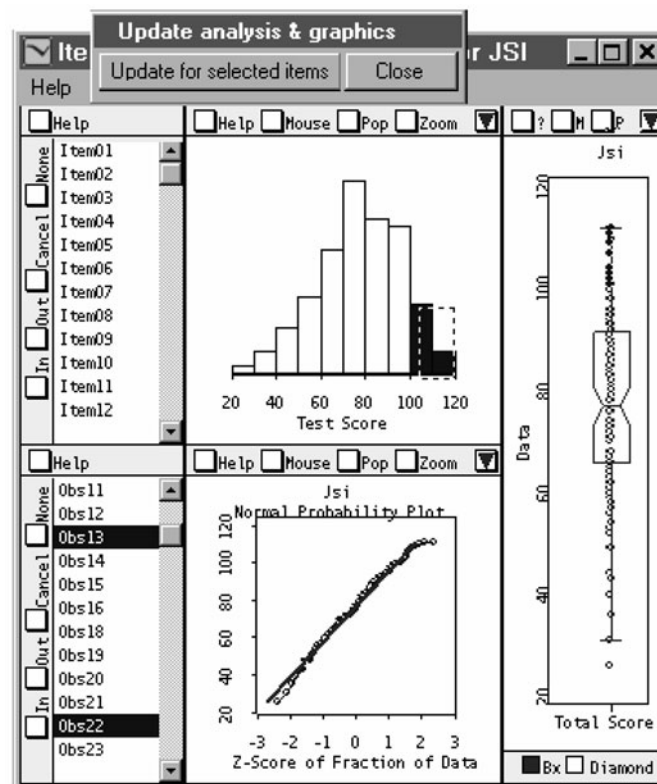


Figure 5. SpreadPlot for the descriptive analysis of the test scores.

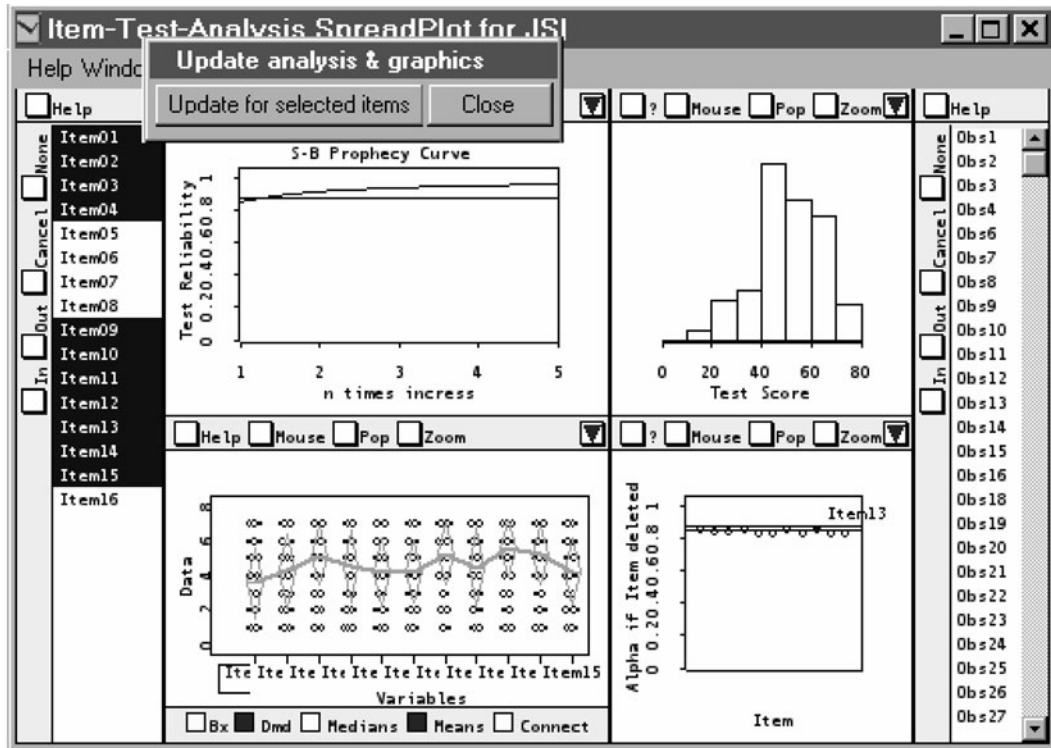


Figure 6. SpreadPlot for the Cronbach's alpha reliability analysis.

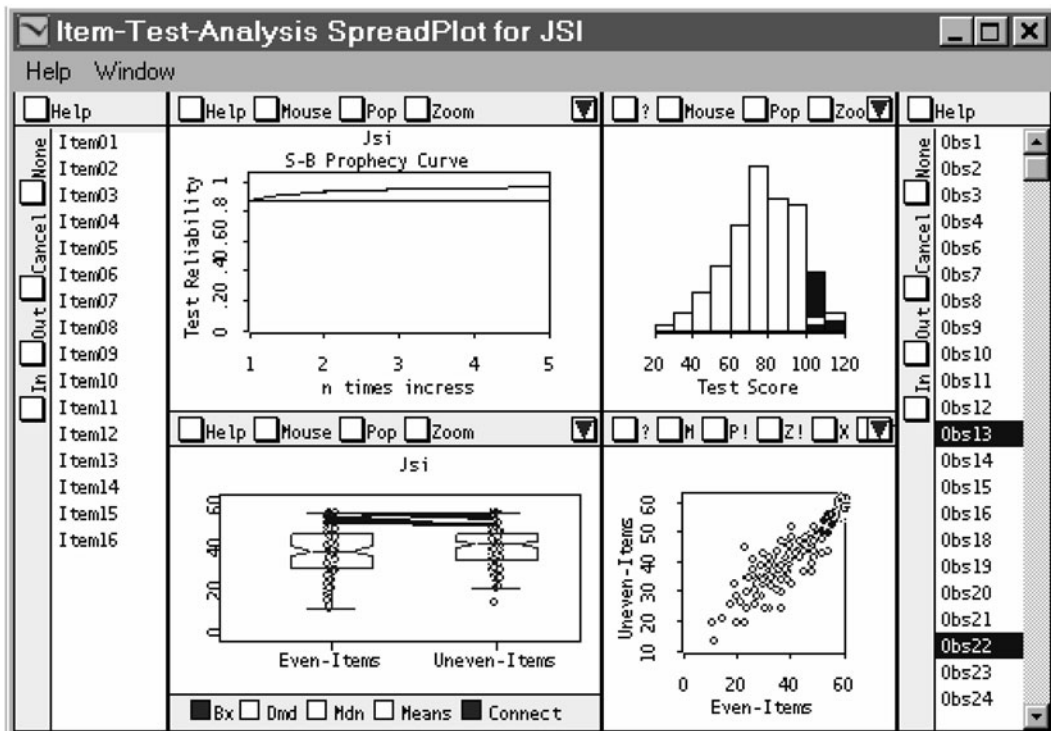


Figure 7. SpreadPlot for the split-half reliability analysis.

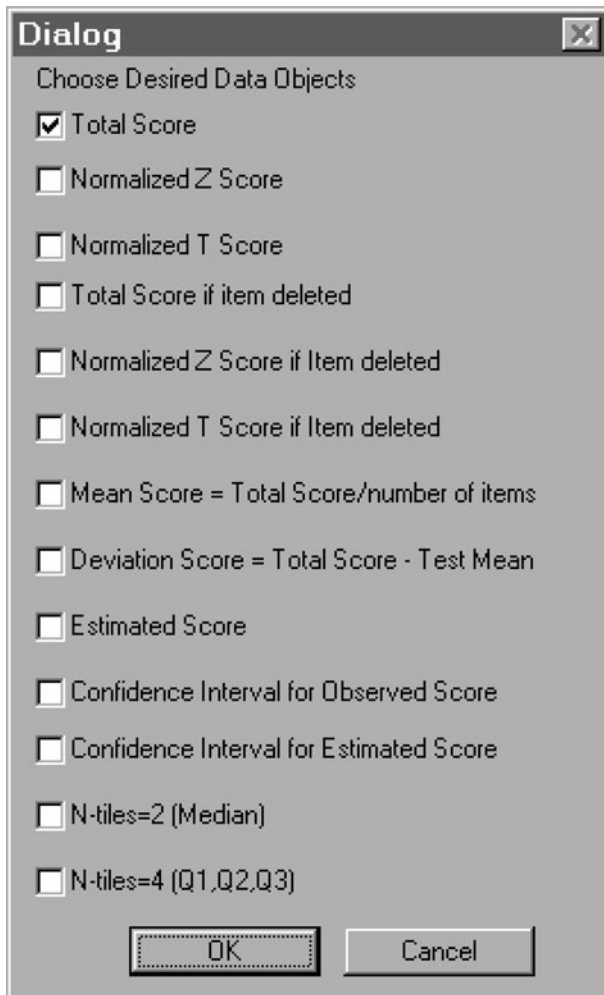


Figure 8. Dialog box associated with the Create-Data menu item of ViSta-CITA.

of the confidence interval ($\alpha = .05$, by default) for the total test scores.

9. *Confidence interval for the estimated scores*: This creates a bivariate data set with the upper and lower limits of the confidence interval ($\alpha = .05$, by default) for the estimated test scores, given the computed standard deviation and the value of the reliability coefficient. These parameters can be modified by using the dialog box appearing when this option is selected.

10. *N-tiles = 2*: This creates a new univariate data set with ones for the scores less than the median and twos for the scores equal to or greater than the median.

11. *N-tiles = 4*: Same as the previous option, although now taking the quartiles as the reference values to split the test scores.

Additional ViSta plug-ins with application to test development. ViSta-CITA integrates some additional procedures that can be of use in data analysis related to test development. This is the case of the following three plug-ins: (1) ViSta-PARAN (Ledesma & Valero-Mora,

2007), which supports the application of parallel analysis, a method that has been recommended to determine the number of factors to retain in the application of EFA (Yu, Osborn-Popp, DiGangi, & Jannasch-Pennell, 2007; Zwick & Velicer, 1986); (2) ViSta-Homals (Ledesma, Valero-Mora, & Young, 2002), which supports the application of homogeneity analysis (Gifi, 1990), an optimal scaling data reduction method that can prove useful in the dimensionality analysis of categorical item response data, a type of data not unusual in the psychological context; and (3) ViSta-Boot, ViSta-Coor-Boot, and ViSta-Alpha-CI (Ledesma, 2008), which implement some nonparametric bootstrap methods to estimate confidence intervals. Some of them can be applied directly to item analysis. ViSta-Alpha-CI, for instance, computes confidence intervals for Cronbach's alpha coefficient. Considering that its application is also supported through dynamic, interactive SpreadPlots, these bootstrap plug-ins could also be used to introduce students to the data resampling methods.

Discussion

When compared with other statistical programs incorporating CTT techniques (e.g., SPSS, Statistica, SAS), ViSta-CITA offers some advantages over them in the education arena. First, ViSta-CITA is free, noncommercial, and available online with no restrictions, so students and teachers can easily obtain it through an Internet connection. Second, the program incorporates advanced dynamic graphics (e.g., SpreadPlots) intended to provide a better understanding of the concepts associated with item and test analysis. Thus, each of the eight SpreadPlots implemented in the ViSta-CITA plug-in consists of a set of linked, interactive plots programmed to support a specific type of Item-Test Analysis.

The SpreadPlots provide greater dynamism to the analysis, thanks to the possibility of graphic interaction with the data and the statistical results. In effect, through the SpreadPlots, the analyst can establish a more fluid "conversation" with the data, favored by the interactive graphic representations involved. This interactive communication with the graphics allows students and novice analysts to better explore the data and to discern more clearly the significance of the statistical information and the theoretical concepts involved. In addition, ViSta-CITA is developed with the Lisp-Stat programming language and runs as a plug-in for ViSta (Young, 2006), an open software development environment geared to statistical visualization and data analysis. As a consequence, advanced users can extend or change the program code to better adapt it to their requirements.

A limitation of ViSta-CITA, though, is that it includes only analysis procedures based on CTT. It does not support certain analysis methods associated with more recent measurement approaches, such as the Rasch modeling and other models related to IRT. Indeed, this opens a new line of work and a platform for future research. However, as a counterweight to this shortcoming, ViSta-CITA is integrated into a statistical system, ViSta, that already includes a number of analysis methods that can also be

useful in test development, such as multidimensional scaling and optimal scaling, among others. A description of the application of these and other methods in ViSta can be found in Young et al. (2006). Planned developments of ViSta-CITA include EFA for ordinal item response data, given that, at present, the EFA option is capable of computing and analyzing Pearson's product-moment and Tetrachoric correlation matrices.

AUTHOR NOTE

Correspondence concerning this article should be addressed to R. D. Ledesma, CONICET, Universidad Nacional de Mar del Plata, Río Negro 3922, Mar del Plata (7600), Argentina (e-mail: rdledesma@gmail.com).

REFERENCES

- COLET, E., & AARONSON, D. (1995). Visualization of multivariate data: Human-factors considerations. *Behavior Research Methods, Instruments, & Computers*, *27*, 257-263.
- CROCKER, L., & ALGINA, J. (1986). *Introduction to classical and modern test theory*. Orlando, FL: Holt, Rinehart & Winston.
- GIFI, A. (1990). *Nonlinear multivariate analysis*. New York: Wiley.
- LEDESMA, R. D. (2008). Introducción al Bootstrap: Desarrollo de un ejemplo acompañado de software de aplicación [Introduction to the Bootstrap: An example with computer software application]. *Tutorials in Quantitative Methods for Psychology*, *4*, 51-60.
- LEDESMA, R. D., MOLINA, J. G., & YOUNG, F. W. (2005). Enhancing dynamic graphical analysis with the Lisp-Stat language and the ViSta statistical program. *Behavior Research Methods*, *37*, 684-690.
- LEDESMA, R. D., & VALERO-MORA, P. [M.]. (2007). Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research & Evaluation*, *12*(Art. No. 2). Retrieved April 18, 2009, from <http://pareonline.net/getvn.asp?v=12&n=2>.
- LEDESMA, R. D., VALERO-MORA, P. M. W., & YOUNG, F. (2002). Análisis de homogeneidad en ViSta "The Visual Statistics System" [Homogeneity analysis in ViSta "The Visual Statistics System"]. *Metodología de las Ciencias del Comportamiento*, *4*, 139-149.
- MOLINA, J. G., LEDESMA, R. [D.], VALERO, P. M., & YOUNG, F. W. (2005). A video tour through ViSta 6.4, a visual statistical system based on Lisp-Stat. *Journal of Statistical Software*, *13*(8), 1-10.
- WAINER, H., & VELLEMAN, P. (2001). Statistical graphics: Mapping the pathways of science. *Annual Review of Psychology*, *52*, 305-335.
- WILKINSON, L. (1994). Less is more: Two- and three-dimensional graphics for data display. *Behavior Research Methods, Instruments, & Computers*, *26*, 172-176.
- WILKINSON, L. (2001). Presentation graphics. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 6369-6379). Amsterdam: Elsevier.
- YOUNG, F. W. (1996). *ViSta: Developing statistical objects*. The Visual Statistic project (Research Memorandum 1996-1). Chapel Hill: University of North Carolina, Thurstone Psychometric Lab.
- YOUNG, F. W. (2006). ViSta, the Visual Statistic System [Software]. Available from www.visualstats.org.
- YOUNG, F. W., & SMITH, J. (1991). Towards a structured data analysis environment. A cognition-based design. In A. Buja & P. A. Tukey (Eds.), *Computing and graphics in statistics* (pp. 253-279). New York: Springer.
- YOUNG, F. W., VALERO-MORA, P. [M.], FALDOWSKI, R. A., & BANN, C. M. (2003). Gossip: The architecture of spreadplots. *Journal of Computational & Graphical Statistics*, *12*, 80-100.
- YOUNG, F. W., VALERO-MORA, P. M., & FRIENDLY, M. (2006). *Visual statistics: Seeing data with dynamic interactive graphics*. Hoboken, NJ: Wiley.
- YU, C. H., & BEHRENS, J. T. (1995). Applications of multivariate visualization to behavioral sciences. *Behavior Research Methods, Instruments, & Computers*, *27*, 264-271.
- YU, C. H., OSBORN-POPP, S., DIGANGI, S., & JANNASCH-PENNEL, A. (2007). Assessing unidimensionality: A comparison of Rasch modeling, parallel analysis, and TETRAD. *Practical Assessment, Research & Evaluation*, *12*(Art. No. 14). Retrieved April 19, 2009, from <http://pareonline.net/pdf/v12n14.pdf>.
- YU, C. H., & STOCKFORD, S. (2003). Evaluating spatial- and temporal-oriented multi-dimensional visualization techniques for research and instruction. *Practical Assessment, Research & Evaluation*, *8*(Art. No. 17). Retrieved April 19, 2009, from <http://pareonline.net/getvn.asp?v=8&n=17>.
- ZWICK, W. R., & VELICER, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 432-442.

(Manuscript received January 22, 2009;
revision accepted for publication May 30, 2009.)