The Role of Sequential Statistical Approach To Upgrade Statistical Thinking And Reasoning of Agricultural Extension, Education and Development Graduate

Sahar Dehyouri

Department of Agriclture, Islamshahr Branch, Islamic Azad University, Islamshahr, Iran

Abstract: Lack of fit scientific methods along with the ignorance of the proper processes of science has somewhat diminished the importance of research. This study points out the dire need to examine "standards" which are based on the statistical method in graduates. Preliminary survey was investigated on PhD dissertations and M.Sc. theses and their authors (N= 315) in agricultural extension, education and development in Iran. In this study were used questionnaire. In the end, graduates could improve statistical reasoning when consideration and using from sequential statistics and obtain holistic view from using statistical test to exploit of mix statistical test.

KeyWords: sequential statistical analysis approach; Statistical thinking; statistical reasoning; post graduate research; agricultural extension, education and development

1.Introduction

Science requires objective evidence before making decisions about the truth or falsehood of a theory. Answering scientific questions demands unbiased observation and testing. Also to obtain appropriate data, knowing and dominating on the superior plan or approach is necessary, so using mix and sequential proper statistical tests to realize accurate outcomes help researchers. Flyvbjerg [1] argued many researches were driven by a continuing belief that the social and political world could be measured through objective, empirically testable and law-like data indicators. Fincher [2] stressed that “research on the substantive issues is handicapped by higher education's lack of status and recognition as an academic discipline and or professional specialty”, the issue of understanding research as a disciplined inquiry is still stealth, especially in higher education. Likewise, Dijkum [3] indicated that analysis of the practice of social research shows there is no easy answer to the question of how the knowledge of the natural sciences can be used to further understanding in the social sciences. It is useful to know that the inseparable section of scientific inquiry is statistical analysis and one of the important sections of statistics is using tests to analyze data. To obtain appropriate data, knowing and dominating on the superior plan or approach is necessary, so using mix and sequential proper statistical tests to realize accurate outcomes help researchers. In this regard, the main problem in conducting M.Sc. theses, Ph.D. dissertations or, even, onward academic or institutional research is that researchers either do not realize their crucial role in choosing, applying appropriate statistical tests and the process of choosing and applying the right test for the right situation (variable combination), or they are not fully aware of the coherence of required consecutive statistical procedures as a core ‘Statistical Mix’. Consequently, choosing and applying the statistical tests seems to become a mere copy-and-paste process for which no initial arguments and intellectual reasoning are offered [4]. In the other hand we need to study that are followed the challenges of combining complex statistics with individual stories, particularly in relation to the ongoing iteration between these different data sets, and issues of validity and reliability[5].

To eliminate this feeling needs dominate on with improving statistical knowledge and understanding the analysis process. So three key concepts are important in understanding the analysis process. These concepts are system, relationship, and model.

1. A *System* is a group of interrelated, interacting or interdependent elements forming a complex whole. A process or operation may be a system. In this standard the term system refers to the system, process, operation, or other subject being analyzed.

(DOI: dx.doi.org/14.9831/1444-8939.2014/2-4/MAGNT.23)

2. A *Relationship* is a statement about the similarities, differences, or interactions of two or more quantities or measurements called variables. Much of the work of operations research lies in identifying the proper variables and true relationships for use in solving a particular problem or evaluating alternatives.

3. A *Model* is a useful representation of the relationships that define a system or situation under study. It may be a set of mathematical equations, a computer program, and a hand played game, a written scenario, an experiment, or other type of representation ranging from verbal statements to physical objects [6]. So there is need the sequential roadmap based on statistical competency.

But, as indicated by Windish and Diener-West [7] and Govindrajulue [8], there are a few, if any, references to the use of sequential statistics in the literature. Although choosing the right statistical test for a particular set of data appears to be an overwhelming task, to Wheater and Cook [9], particularly if such decisions are rendered after the data are collected, what is overwhelming really, is the sequences and placements (array) of statistical tests to understand their role and mission in the first place. Wheater and Cook [9] believe that the investigator is definitely responsible for the choice of statistical methods used. Therefore, the researcher must be able to use statistics effectively to organize, evaluate, and analyze the data [10] and to apply the proper statistical tests. To ease the dilemma, it is helpful to identify the statistical test, as stated by Hoffman [11], which is a procedure for deciding whether an assertion (e.g., a hypothesis) about a quantitative feature of a population is true or false. There are a few cautionary steps to follow in selecting a statistical test in educational research; firstly, because of the high number of variables involved and secondly, because of the involvement of a considerable number of latent (unobserved) [12], hidden [13], and discrete variables.

Watt and Berg [14] stressed that the choice of the correct statistical test depends on the definition of the variables, particularly upon their level of measurement. It also depends on the research design used and the nature of the hypotheses: are they comparative or related; is there more than one independent variable?

Since, students have problems to learn statistics [15]. It may be because of some wrong learned statistical concepts and applications. Malek Mohammadi [4] posed a model in the sequential statistical analysis approach (SSAA) to present a mixed and sequential method. In this model, he mixed three phases and divided each phases to several steps that researchers should use them to refine and improve research; they are as following:

According to Tabachnick and Fidell [16], since each statistical technique has specific assumptions, therefore, before applying any technique, or sometimes even before choosing a technique, it should be determined how the data fits some very basic assumptions underlying most of the multivariate statistics. Moreover, each statistical technique has some limitation along with its advantages.

Therefore no statistical test can replace the other under the very same circumstances. For instance, while Hill and Lewicki [17] identified multiple regression as "a seductive technique as plug in as many predictor variables as you can think of and usually at least a few of them will come out significant". Eventually, many difficulties tend to arise when there are more than five independent variables in a multiple regression equation.

One of the most frequent is the problem of two or more independent variables being highly correlated to one another. This is called multicollinearity. If a correlation coefficient matrix with all the independent variables indicates correlations of 0.75 or higher, then there may be a problem with multicollinearity. Primary investigations of M.Sc. theses and Ph.D. dissertations as supervisor, co-supervisor, external examiner, and researcher, revealed major challenges concerning selection and application of legitimate statistical techniques. This investigation showed that in majority of the cases, graduate and post graduate students are seduced (due to their lack of statistical literacy, reasoning and/or thinking) by widely applied statistic decision trees and/or statistical tables, and do not consider the authors' guidelines appropriately. Therefore, unintentionally, apply inappropriate tests.

What was found commonly neglected in graduate researches under this study is that all groups involved in their studies were taken as independent groups, while potentially, all or some of them have been dependent in their nature.

However, in some cases, as having one DV and two or more IVs with independent groups and with ordinal or interval nature of DV, suitable statistical tests are lacking.

Also, when there is one interval IV and one DV with interval and nominal nature in one case, and ordinal or interval nature in another case, correlation and nonparametric correlation are recommended, respectively, but this does not sound quite right because of very rare conditions that correlation may imply causation as explained by Huck [18].

In numerous cases of hypotheses testing when the "pvalue" was significant, then the researcher usually has not cared that some other tests would give a smaller (more significant) *p*-value. If the *p*-value is not significant, then the researcher usually considers whether there is a better test [19]. This is often true with graduate students whose hypotheses were mostly rejected; they usually try some other statistical tests to possibly change their results (Type I error). Likewise, when the study reaches a conclusion of "no statistically significant difference", it should not necessarily be concluded that the treatment was ineffective. Otherwise, a Type II error happens, as was the case in many thesis and dissertations. Consequently, the power of the statistical tests (the probability of rejecting null hypothesis when it is false) is questioned due to the fact that for a fixed Type I error rate ( ) the goal of constructing and testing a hypothesis is to maximize *Power* [20].

Regarding misapplication of statistical procedures, the following four consecutive phases were developed along with their components to build up SSAA. Each phase is composed of a few steps through which general and specific criteria for selecting and applying statistical tests are being discussed as follows.

A)Descriptive phase

- Variable mining and measurement

This entails listing of variables involved in the study and measuring them after a scrutiny of some general research notions as: research problem and research question [21,22]; the goal of the analysis [16]; nature of the data, research design, kind of research [13; 23; 24); kind of variables [9]; and variable mathematical nature (nominal, ordinal, interval, or ratio) (25; 7; 26); and finally, number of variables [16].

- Variable sorting out techniques

There are a few variable sorting out techniques to come up with the optimum IVs prior to involving all variables in the hypotheses-testing process. The following are some procedures that are already implemented by the author in different projects:

Reliability: Applying this technique makes it possible for a researcher to eliminate variables with low Cronbach's alpha [27]; Kudar Richardson [28], and recently, ordinal Theta coefficient [29].

Coefficient of variability (CV): Coefficient of variation [30] is recommended in this article for the purpose of consistency and accountability measurement as well as setting priority or even ranking variables. By applying CV, the researcher can select the most consistent variables with the lowest risk, and leave out the least consistent variables from the study.

Correlation matrix: Variables with statistically significant and higher correlation coefficient may be more legitimate and subject to further investigations in the research process. Therefore, variables with low or no significant correlation coefficient in the matrix can be eliminated.

B) Analytical phase

- Variable refinery

To isolate the sensitive cases and exclude them from the main study, personal characteristics of the respondents are being tested against each one of the dependent variables (DV) [31]. By applying this technique, the researcher can extract as few appropriate independent variables as he/she should, due to the limited capacity of inferential statistical techniques (that is, regression, path analysis and structural equation modeling), to enhance research reliability and create favorable environment to applying appropriate statistical tests.

- Variable reduction

Following the above technique and developing from what Thompson [32] and, Tabachnick and Fidell [16] explained about Exploratory Factor Analysis (EFA), this technique is implemented in SSAA as a converter to group numerous single variables into few "supervariables" or "factor/s", and explicit relevant discrete or continuous latent variables in the study on which the subject differs.

- ***Latent variable measurement***

Multiple regression analysis is highlighted in SSAA due to its capability to identify and measure latent variable/s in the study through a mathematical model. Of course, IVs (predictors) involved in predicting latent DVs (indicator/s) can be latent construct (factor) as the outcome of exploratory factor analysis, or simple variables. Either one of these should be specified prior to regression analysis. Notably, the scale or mathematical nature [25], of criterion variable is worthy of consideration in choosing the right regression model.

That is, when criterion variable is nominal dichotomous; Logistic Regression (LR), when it is ordinal (discrete); Ordinal Logistic Regression (OLR) [33,34], and when it is quantitative (continuous); Ordinary Regression (OR) suits the model.

Although it is stated that multiple regression is a seductive technique: "plug in" as many predictor variables as you can think of and usually at least a few of them will come out significant [35], but, to Palmer (MND), it is possible that the independent variables could obscure each other's effects. To prevent this situation, SSAA is considering multiple regression (in either forms), as another converter technique with dual simultaneous role to be applied after EFA. The first role deals with the limitation of regression analysis in embedding numerous variables in the equation. In this case, super-variables (the explicated latent variables through EFA) will be entered in the equation to measure variation of a latent predictor that could not be measured directly before. And, MRA, in its' second role eliminates those variables with no significant impact on the predictor variable. What remains will be utilized next in the SSAA inferential phase.

C)Inferential phase

As indicated by Bruin [36], to enable one infer from his/her population data, procedures that use significance tests must be employed. Rationales behind inferential phase to help the applicants choose appropriate statistical tests are as follows.

- ***Variables, data, and groups***

Variables (independent/dependent) [37,16]; or exogenous/ endogenous [38]; matched or paired data [26]; Kind of samples being compared (independent/dependent) and; number of groups being compared (one, two, or more than two).

also Morphological and syntactical analysis proved to be the most suitable technique for extracting entities, attributes and relationships from RS(55).

Hypotheses testing (choosing the legitimate statistical technique)

A hypothesis is a statement that describes or explains a relationship between or among variables [39]. Also, a statistical hypothesis test is defined by Lehmann and Romano [40] as a method of making statistical decisions using experimental data. If there is no hypothesis, then there is no statistical test. Pvalue [41]; effect size [42,43,39]; sample size [44]; central limit theorem [45];

number of independent hypotheses or multiple comparisons [13; 23]; paired or unpaired [13]; parametric/nonparametric [13; 19; 24; 45] are detected as major criteria for testing a hypothesis and considered in SSAA to choosing eligible statistical test.

- **Structure or model phase**

To Bartholomew [46], a model is:

1) An abstraction of the real world in which the relevant relations between the real elements are replaced by similar relations between mathematical entities.

2) A set of assumptions about the relationship between the parts of the system. Its adequacy is judged by the success with which it can predict the effects of changes in the system.

*- **Structural Equation Modeling (SEM)***

Haenlein and Kaplan [47], referring to Gefen et al. [48], named regression analysis as a first-generation technique, which analyzes only one layer of relationships among multiple independent and dependent variables. At the same time, they recommended SEM as a secondgeneration technique that allows simultaneous modeling of relationships among multiple independent and dependent constructs. Raykov and Markoulides [49] observed that SEM enables researchers to readily develop, estimate, and test complex multivariable models as well as to study both direct and indirect effects of variables involved in a given model. The combination of direct and indirect effects makes up the total effect of an explanatory variable on a dependent variable.

Garson [50] believes that “SEM grows out of and serves purposes similar to multiple regression, but in a more powerful way, which takes into account the modeling of interactions, nonlinearities, correlated independents, measurement error, correlated error terms, multiple latent independents each measured by multiple indicators, and one or more latent dependents, also each with multiple indicators. SEM may be used as a more powerful alternative to multiple regression, path, factor and time series analyses, as well as analysis of covariance". This technique combines factor analysis, canonical correlation, and multiple regressions to evaluate whether the model provides reasonable fit to the data and the contribution of each of the IVs to the DVs [16]. While Garson [50] views SEM as a confirmatory rather than an exploratory procedure, Raykov and Markoulides [49] consider four types of SEM: path analysis model, confirmatory factor analysis model (CFA), structural regression model, and latent change model. Having the capacity of testing modeling hypotheses, SEM is installed in SSAA to develop structures or models considering the following applications of path analysis and CFA.

Path analysis (PA): To Streiner [38], path analysis is an extension of multiple regression therefore, it goes beyond regression to allow the analysis of more complicated models. Although, despite its previous name of “causal modeling,” Streiner does not believe in path analysis as to establish causality or even to determine whether a specific model is correct; rather, it can only determine whether the data are consistent with the model. However, it is extremely powerful for examining complex models and for comparing different models to determine which one best fits the data.

Path analysis basically examines the direct relationships through the postulation of some theoretical relationships between variables and then tests to see if the direction of these relationships is substantiated by the data.

Confirmatory factor analysis CFA: CFA is usually employed to examine patterns of interrelationships among several latent constructs. According to Raykov and Markoulides [49], “no specific directional relationships are assumed between constructs, only that they are potentially correlated to one another”. The starting point of CFA is a very demanding one, requiring that the complete details of a proposed model be specified before it is fitted to the data.” The latter statement by Raykov and Maroulides was more clearly explained by Stapleton [51] when he described CFA as “a theory-testing model as opposed to a theory-generating method like EFC.

4 Conclusion

This paper is intended to be a concise guide for choosing a statistical test with regard to notions extracted from SSAA and statistical literacy, reasoning and thinking for education assessment and for interpreting and analyzing educational studies without relying on mathematical theories. To provide a framework for understanding statistical concepts and to illustrate the decision-making process needed to choose a statistical test, we’ve presented an educational intervention detailing the hypothesis testing, data analysis, and interpretation of the results. Big ideas are shown to embed as SSAA process, with three phases becoming meaningful through the statistical competency. Meanwhile, they could improve knowledge content about statistic when apply this roadmap scientifically, also consideration and using from sequential statistics by agricultural extension and education students, can give them holistic view from using statistical test to exploit of mix and together statistical test. Whereas student can see statistical test in to the system that conduct to superior realize of relationship between level and phases and suppose them as group of interrelated, interacting or interdependent elements that forming a complex whole.

Finding synthetic test enable student to refine data and variables and extract appropriate and correct data and variable to extract pure result and ultimately pure knowledge.
researchers and graduate students who are after accurate application of statistical methods, and secondly, this will lower their stat phobia by leading them towards a 50-strategic-sequential-statistic-roadmap for choosing and applying appropriate statistical tests, interpreting their findings, and implementing scientific analysis more realistically in a creative research enterprise.

References:

1. Flyvbjerg, B. (2001) Making social science matter: Why social inquiry fails and how it can succeed again. Cambridge, UK: Cambridge University Press.
2. Fincher,C.(1991). The Possibilities and Actualities of Disciplined Inquiry. Research in Higher Education .32(6), 625-650
3. Dijkum, C.(2001). A methodology for conducting interdisciplinary social research. European Journal of Operational research 128(2): 290-299
4. Malek Mohammadi,I.(2009). Sequential Statistical Analysis Approach (SSAA) towards Contingency Framework Purification in Bevavioral Research and Practice. Presenred at First International Conference on Educational Research and Practice. University Putra Malaysia
5.Kington,A. Sammons, P. Day,C. Regan,E.(2011).Stories and Statistics: Describing a Mixed Methods Study of Effective Classroom Practice . Journal of Mixed Methods Research 1558689810396092, first published on February 21, 2011 as doi:10.1177/1558689810396092
6.Wild, C. J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (with discussion).International Statistical Review, 67(3), 223-265.
7. Windish D, Diener-West M (2006). A Clinician-educator's Roadmap to Choosing and Interpreting Statistical Tests. J. Gen. Int. Med., 21 (9, 8), 1009-1009. DOI 10.1111/j.1525-1497.2006.00390.x
8. Govindrajulue Z (2004). Sequential Statistics. World Scientific Publishing Inc. NJ, USA. p. 1
9. Wheater P, Cook P (2000). Using Statistics to Understand the Environment. Routledge, NY, USA.
10.Whitney (2005). Statistics: A Tool for Social Research. 7th Ed. Thomson Learning, Inc. Thomson Wadsworth. Belmont, CA 94002-3098, USA.
11.Hoffman H (2004). Internet Glossary of Statistical Terms in: Statistics Explained . www.animatedsoftware.com/statglos/statglos.htm)
12.Vermunt J, Magidson J (2003). Statistical Innovation. Encyclopedia of Social Science Research Methods, Sage Publications.
13.Motulsky H (1999). Analyzing Data with GraphPad Prism. GraphPad Software Inc., San Diego, CA.
14.Watt J, Berg S (2002). Research Methods for Communication Science. Amazon.com.
15.Meletiou, M., Lee, C.(2002).The Role of Technology in the Introductory Statistics Classroom: Reality and Potential. Proceedings of the International Conference on Mathematics/Science Education and Technology. San Antonio, Texas.
16.Tabachnick B, Fidell L (2007). Using Multivariate Statistics. 5th Ed. Pearson. NJ. USA. ISBN-13: 9780205459384
17.Lewicki, P. (2007). STATISTICS Methods and Applications. StatSoft, Tulsa, OK. USA.
18. Huck S (2009). Statistical Misconceptions. Taylor and Francis Group, NY, USA. pp. 43.
19.Dixon P (2003). Choosing Statistical Methods. Iowa State University, USA. From http://www.bcb.iastate.edu/faculty/Dixon_CV.pdf. Accessed in March 2009.
20.Anderson-Cook, M, Dorai-Raj S (2003). Making the Concepts of Power and Sample Size Relevant and Accessible to Students in Introductory Statistics Courses using Applets. J. Statistics Educ., [Online], 11(3).
21.Bruin J (2006). Newtest: Command to Compute New Test. UCLA: Academic Technology Services, Statistical Consulting Group. http://www.ats.ucla.edu/stat/stata/ado/analysis/.
22.Marion R (2004). The Whole World of Deduction Research Skills for New Scientists. The University of Texas Medical Branch, Texas, USA.
23.Wadsworth (2005). Choosing the Correct Statistical Test. Wadsworth Cengage Learning. Florence,KY.http://www.wadsworth.com/psychology_d/templates/student_resourc es/workshops/stat_workshp/chose_stat/chose_stat_01.html
24.Dinove I (2008). Choose the Right Test. From http://www..stat.ucla.edu Accessed in May 2009.
25.Healey J (2005). A Tool for Social Research. 7th ed. Wasdworth, Thomson Learning Academic Resource Center, CA, USA.

26.Kaminsky J (2008). Qualitative and Quantitative Analysis. Kwantlen Polytechnic University. BNS Nursing Program. Greater Vancouver, British Columbia, Canada.
27.Ferrando P (2009). A General Factor-Analytic Procedure for Assessing Response Bias in Questionnaire Measures. Structural Equation Modeling: A Multidisciplinary J., 16(2):364–381. DOI: 10.1080/10705510902751374
28. Rudner LM, Schafer WD (2001). Reliability. ERIC Digest. Online http://www.ericdigests.org/2002-2/reliability.htm
29. Zumbo B, Gadermann A, Zeisser C (2007). Ordinal Versions of Coefficient Alpha and Theta for Likert Scale. J. Modern Appl. Stat. Methods, 6(1): 21-29.
30.Calvin D (2004). Choosing and Using Statistics, a Biologists Guide. Blackwell Publishing Company. MA. USA
31. Malakmohammadi I (2008). Variables Refinery Process to Ensure Research Unbiasedness (Validity) and Invariance (Reliability) in Agricultural Extension and Education. Am. J. Agric, Biolo. Sci., 3(1): 342-347.
32. Thompson B (2004). Exploratory and Confirmatory Factor Analysis: Understanding Concepts and Applications. Washington, D.C.: American Psychological Association.
33. Conne A (2006). Logistic Regression Models for Ordinal Response Variables. Sage Publications Inc. Thousand Oaks, Cl, USA., Pp. 10- 25
34.Hilbe J (2009) Logistic Regression Models. Chapman & Hall/CRC Press, Texts in Statistical Science. NY.USA. PP: 353-3 Hill, T. &
35. Statistica (2008). Multiple Regression. Electronic Textbook. Statsoft, Inc. http://www.statsoft.com/textbook/stmulreg.html
36. Bruin J (2006). Newtest: Command to Compute New Test. UCLA: Academic Technology Services, Statistical Consulting Group. http://www.ats.ucla.edu/stat/stata/ado/analysis/.
37. Hill T, Lewicki P (2007). Statistics Methods and Applications. StatSoft, Tusla, OK.
38. Streiner D (2005). Research Methods in Psychiatry. Finding Our Way: An Introduction to Path Analysis. The Canadian J. Psych., 50: 115–122
39. Graveter F, Forzano L (2008). Research Methods for the Behavioral Science. Third ed. PP:19-22. Thomson/Wadsworth, ISBN0495091456, 9780495091455.
40. Lehmann E, Romano J (2005). Testing Statistical Hypotheses (3rd ed.). New York: Springer. ISBN 0387988645.www.virtualcurriculum.com/N4120/LA9_Figure8.pdf. Accessed on March 2009. http://www.ats.ucla.edu/stat/sas/notes2/ Accessed in March, 2009.
41. Denis D (2003). Alternatives to Null Hypothesis Significance Testing. J. Theory. Sci., 4-1 The International Consortium for the Advancement of Academic Publication, Athabasca, Canada.
42. McCloskey D (2008). The Cult of Statistical Significance. Ann Arbor: University of Michigan Press. ISBN 0472050079.
43. NN (2007). Statistical Test: More Complicated Discriminate. PHY310: Lecture 14. http://nngroup.physics.sunysb.edu/~mcgrew/phy310/lectures/phy310-lecture-14-2007.pdfnngroup.physics.sunysb.edu/~mcgrew/phy310/lectures/phy3 10-lecture-14-2007.pdf
44. McDonald JH (2008). Handbook of Biological Statistics, Choosing a Statistical Test. Sparky House Publishing, Baltimore, Maryland. P.282-287.
45. Bartholomew J (1998). Social Process. London School of Economics and Political Science. Data and Sources of Collinearity. New York: John Wiley & Sons.
46. Haenlein M, Kaplan A (2004). A Beginner's Guide to Partial Least Squares *Analysis*. Lawrence Erlbaum Associates Inc. Understanding Stat., 3(4): 283-297.
47. DelMas, R., Garfield, J., & Chance, B. (1999). A model of classroom research in action: Developing simulation activities to improve students' statistical reasoning. Journal of Statistics Education, 7 (3)
48. Raykov T, Marcoulides G (2006). A First Course in Structural Equation Modeling. Lawrence Erlbaum Association Inc. Publishers. NJ, USA. pp. 117-118.
49. Garson. D. (2008). 'Structural Equation Modeling' from Statnotes: Topics in Multivariate Analysis. North Carolina State University, Retrieved May 25, 2009.http://faculty.chass.ncsu.edu/garson/pa765/statnote.htm.
50. Salkind N (2008). Statistics for People Who (Think They) Hate Statistics. Excel 2007 Edition. Sage Publication Inc. CA, USA. pp.326
51. Stapleton. C .(1997). Basic Concepts and Procedures of Confirmatory Factor Analysis. Paper presented at the Annual Meeting of the Southwest

Educational Research Association. Austin, TX, January 23-25.
52. Vermunt, J. D. and Vermetten, Y. J. (2004), "Patterns in Student Learning: Relationships Between Learning Strategies, Conceptions of Learning, and Learning Orientations," *Educational Pyschology Review*, 16(4), 359-384.
53. Garfield, J. (2003), "Assessing Statistical Reasoning," *Statistics Education Research Journal* [Online], 2(1), 22-38. http://www.stat.auckland.ac.nz/~iase/serj/SERJ2(1).pdf
54. Sorto, M. A. Prospective Middle School Teachers' Knowledge about Data Analysis and its Application to Teaching. Doctoral dissertation. Michigan State University. East Lansing. 2004
55. K. Zafar, B. Khan, A. Mubeen, A. Syed, A. Khan, *Z. Halim, S. Rehman. (2014). Autonomous Semantic Data Modeling Tool using Computational Linguistics .
BRIS Journal of Adv. S & T (ISSN. 0971-9563). Vol.2 (3):pp.8-14