ISSN: 0975-833X

Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 10, Issue, 04, pp.68808-68813, April, 2018

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

RESEARCH ARTICLE

EXPERIMENTAL RESEARCH ON MULTIPLE MEASUREMENTS BASED ON SECONDARY RESULT ANALYSIS

*** Dr. Moustakas Loukas**

Postdoc Researcher, University of the Aegean, Faculty of Humanities, Department Of Preschool Education & Educational Design

ARTICLE INFO

Article History:
Received 26th January, 2018
Received in revised form
08th February, 2018
Accepted 09th March, 2018
Published online 30th April, 2018

Key words:
Experimental method,
Statistical analysis,
Secondary analysis of data.

ABSTRACT

Experimental research in education is applied when the researcher is asked to conduct an educational experiment and at the same time to check its success. Typically, experimental patterns refer to the comparison of measurements either between different research groups or the measurements of the same research group at different times, but in many cases the research patterns are applied in combination, comparing multiple measurements of different research groups at different stages of the experiment. In the case of multiple measurements, the existing statistical criteria can actually inform the researcher about the statistical significance of the overall progress of the experiment, but they do not provide data on how specific groups of participants behaved in the experiment. When applying the experimental multivariate research method, it is possible with the use of statistical analysis to find out whether the group involved responded positively by improving its average performance or on the contrary if its performance has deteriorated. But there could be no knowledge as to whether this average success or failure is due to the participants as a whole or to specific subgroups. The purpose of the present study is to propose an innovative method of secondary analysis of the multiple measurements of experimental educational research, which enables the researcher to draw conclusions not only about the overall course of the experiment, but also about the participants' performance i.e. the degree to which individual subgroups of participants, for example participants of different potential, contributed to the success or failure of it . This proposal is developed in two stages, using a chart to identify current trends and a secondary analysis of data with statistical criteria in order to control the significance of these trends, thus providing the researcher with all the necessary data regarding the outcome of the experiment and in particular regarding the behavior and the degree to which all subgroups involved in the experimental process responded. With the implementation of this proposed secondary analysis method it is possible to detect if a mobilization of potential participants as in the case of marginalized groups occurred or did not in the experimental groups involved, which in the typical educational reality would remain indifferent, inert. The present proposal provides researchers of educational research with the potential that is not offered by the application of classical statistical criteria, which are limited to providing information on the overall outcome of the experiment. This proposed method concerns all teachers who apply an experimental approach and wish to further analyze the results of their measurements in depth for example the contribution of each subgroup of the participants to the final outcome of the experiment.

Citation: Moustakas Loukas, 2018. "Experimental research on multiple measurements based on secondary result analysis", *International Journal of Current Research*, 10, (04), 68808-68813.

INTRODUCTION

Experimental Research is applied to education when the researcher is asked to conduct an educational experiment and check its success at the same time. Typically, experimental forms refer to the comparison of measurements either among diverse research groups or among the measurements of the same research group at different times, yet in many cases the research patterns are applied in combination i.e. by comparing multiple measurements of diverse research groups at different stages of the experiment.

Corresponding author: **Dr. Moustakas Loukas,**
Postdoc Researcher, University of the Aegean, Faculty of Humanities, Department Of Preschool Education & Educational Design.

In the case of multiple measurements, the existing statistical criteria have the potential to inform the researcher of the statistical significance regarding the overall progress of the experiment, but they do not provide data as to the way specific groups of participants behaved in the experiment, as is the case of "efficient" or "weak" students. This specific proposal develops in two stages, using a diagram on the one hand, to detect the existing trends and a secondary analysis of data on the other hand, to check the significance of these trends statistically, thus providing the researcher with all the necessary data as far as the outcome of the experiment and the progress of the experimental group are concerned and more specifically regarding the behavior and response of all the participating groups in the experimental process. The researcher, throughout this process often chooses to compare the performance of two or more groups, depending on the kind of experiment that is carried out, according to which one of the groups involved is usually the control group (CG) and acts as a point of reference and the other or others, which are the experimental groups (EG), are those to which equivalent experimental approaches are applied respectively. In other cases, attention is focused on one specific experimental group whose development and progress the researcher monitors at different stages of the experiment. But at times there may also be selected mixed types, where the performance of many groups is being monitored several times during the development of the experimental process (Patrick, 2016 , Roussos and Tsaousis, 2005). The statistical criteria available in bibliography which are used to check the progress of these experimental measurements are limited to only providing information on their statistical significance at the overall level, i.e. whether the differences in the measurements between the different experimental groups or the diachronic measurements of the same group are statistically significant. In other words, they may provide information on the overall progress of the experiment and not on the differentiations of the subgroups individually. (Roussos and Tsaousis, 2005).

This proposal is applied to the second experimental plan, i.e. when we want to examine diachronically the progress of the measurements of an experimental group , and involves the way in which we may continue the analysis of our research data to see not only the overall progress of the experiment but also the response of individual groups of participants. If the first research pattern is followed, then two equivalent groups can be formed, an experimental one (EG), where the educational innovation will be applied and a second control group (CG), where conventional coaching methods will be applied. In this case, the subject in question is to prove if and to what extent one group, (namely the experimental one), outrivals the other according to measurements. If the researcher follows the second pattern, they should create the experimental group (EG) to which the innovative method will be applied and then proceed with an initial measurement (Before) and at least a final (After) or as many required while the experiment is in progress.

At this stage, depending on the type of variables, the measurement scale and whether or not certain conditions are in order, which will be analyzed explicitly, it is required to use specific statistical criteria, which in any case will be there to inform that there is a very large probability (95%) that the two or even more measurements of the experimental group are statistically substantially different. In a case like this, if the last measurements are better, it is inferred that the experiment progresses positively and the coaching application is effective, while on the contrary, if they are worse, then the new coaching proposal appears not to have met the standards. However, the application of these statistical criteria fails to give information on a possible differentiation in performance among students of different levels. In other words, it is possible that the statistical criterion provides information on the overall success of the experimental application with that being restricted only among the good students or even more interestingly with that being detected only among the average ones. The implementation of an innovative methodological practice may initially be considered effective and through the proposed secondary analysis it may as well turn out that mainly "weak" students were particularly benefited from it , an interesting and useful deduction for the teacher (Moustakas, 2016).

CURRENT METHODOLOGY

Selection of statistical criteria for testing the significance of experimental data

The researcher in the context of experimental planning, after choosing which form to follow, should, along with the creation of the research tools, decide how to evaluate the results of the experiment held. It is possible to consider providing a questionnaire, filling in a criterion or an observation key, interviewing the participants, electronically recording and collecting data by using a specific device, or doing anything that is considered best suited to the specific case. This decision should require the creation of corresponding variables, which will "accommodate" experimental data for statistical analysis using specific criteria (Field, 2009, Howell, 2007). Irrespectively of the research planning, method and evaluation tools to be selected, the variables that can be used are of four kinds, quantitative on a ratio scale of measurement or interval scale of measurement, grading on a likert scale, ordinal scale of measurement and nominal. Because the determining factor for the selection of statistical criteria is the type of variables, the attention will also be focused on the analysis of the experimental data and on the two forms in this parameter (Andreadakis and Vamboukas, 2005, Yalamas, 2005).

Checking regularity

One thing to consider and check prior to initiating the analysis of experimental data, is whether their values follow the normal distribution, a basic prerequisite for selecting parametric or non-parametric statistical criteria, respectively. When checking is in effect, one may select the Shapiro-Wilk criterion for small experimental groups of up to fifty people or the Kolmogorov-Smirnov criterion for larger samples. In any case, in order for measurement values to follow the normal distribution, the statistical significance of the criterion should be greater than 5% ($p > .05$) (Field, 2009, Roussos and Tsaousis, 2005, Green and Salkind, 2003) (see Table 1).

Table 1. Presentation of statistical criteria of experimental research

Independent samples Between-subjects	Control variable	Parametric criterion	Non parametric criterion
1 Factor: 2 groups For example: Control group & Experimental group	Quantitative	Independent samples t-test	Mann-Whitney
	Likert scale		
	Nominal	-	x^2
1 Factor: >2 groups For example: Control group & 2 Experimental groups	Quantitative	One-way ANOVA	Kruskal-Wallis
	Likert scale		
	Nominal	-	x^2

Repeated measures Within-subject	Control variable	Parametric criterion	Non parametric criterion
2 Measures Before & after	Quantitative	Related samples t-test	Wilcoxon
	Likert scale		
	Nominal	-	McNemar
>2 Measures Beginning - during - end	Quantitative	One-way repeated measures ANOVA	Friedman
	Likert scale		
	Nominal	-	Cochran's Q

Table 1 above shows the statistical criteria used in one-factor analyses in both experimental patterns: a) independent samples different experimental groups (between- subjects) and b) repeated measures (within- subject) in relation to the kind of variables used (Moustakas, 2016).

Experimental pattern of independent samples Equivalence check

Depending on the research choices and the type of variables that will be required, the relevant statistical criteria will be used to control the significance of the changes in the measurement values among the members of the different groups. The main concern in each case is to check and verify the equivalence of the groups especially at the beginning of the experiment. There should not be found very large differences in measurement values and these differences are not to be statistically significant (p> .05). Otherwise, we conclude that the experimental groups are not equivalent, so the researcher should redesign and redistribute their members in order to achieve their equivalence, which is not always feasible or easy (Field, 2009, Roussos and Tsauousis, 2005). Equivalence checks may also be carried out during the experiment but are definitely required after its completion, to determine whether or not the experimental group (EG) is superior to the control group(CG). In this case the researcher expects to spot the greatest possible differences in the values of the measurements which in turn are bound to be rendered statistically significant by the corresponding inductive statistical criteria (Roussos and Tsaousis, 2005).

Selection of statistical criteria

In the case that we have only two groups in the experimental planning, one control group and one experimental or two experimental ones , then if our dependent variable - measurement variable is : a) categorical with two (2) categories then, if the conditions are met, we are to choose the statistical criterion x2, otherwise we will use the corrected value of Fisher's Exact Test, b) categorical with more than two categories then, if the conditions are met, we are to choose the statistical criterion x2, otherwise we will have to merge categories ,(c) quantitative or qualitative on a likert scale, then, if the conditions are met, we will choose the t-test statistical criterion for independent samples, otherwise the non-parametric Mann-Whitney criterion instead, and d) an ordinal scale (an option less preferable) for which we are to choose the non-parametric criterion of Mann-Whitney (Field, 2009, Andreadakis and Vamboukas, 2005, Roussos and Tsaousis, 2005, Sprent, 1993). In the case where the experimental planning includes more than two (2) experimental groups then, if our dependent variable - measurement variable is: a) categorical with two (2) categories then, if the conditions are met, we will select the statistical criterion $\chi 2$ otherwise, because we cannot merge categories in experimental groups we will need to perform the test x2 criterion per pair- groups and as a result wherever it is applicable we can use the corrected value of Fisher's Exact Test, b) categorical with more than two categories then, if the conditions are met, we will choose the statistical criterion x2, otherwise we will be forced to merge categories or perform the criterion x2 per pair – groups , c) quantitative or qualitative on a likert scale then, if the conditions are met, we will choose the statistic criterion One Way ANOVA, otherwise the nonparametric Kruskal-Wallis criterion and d) an ordinal scale of measurement, nominal scale of measurement , then we will choose the non-parametric criterion Kruskal-Wallis (Field, 2009, Andreadakis and Vamvoukas, 2005 Roussos and Tsaousis, 2005, Sprent, 1993).

Experimental Multiple Measurement Scheme

In the case where we have only two measurements of the experimental planning, one (BEFORE) and one (AFTER), then if the dependent variable - measurement variable is :a) categorical then we are to choose the statistical criterion McNemar, b) quantitative or qualitative on a Likert scale then, if the conditions are met, we will select the statistical criterion paired t-test for related samples, otherwise the non parametric Wilcoxon test, and c) an ordinal scale, then we will choose the non-parametric criterion Wilcoxon (Field, 2009, Andreadakis and Vamboukas, 2005, Roussos and Tsaousis, 2005, Sprent, 1993). In the case that the experimental planning contains more than two (2) measurements, then if our dependent variable - a measurement variable is a) categorical then we will choose the Cochran's Q statistical criterion, b) quantitative or qualitative on a likert scale then, if the conditions are met, we will choose the one-way Repeated measures ANOVA statistic criterion, otherwise the non-parametric Friedman criterion and c) the ordinal scale then we will choose the non-parametric criterion Friedman (Field, 2009, Andreadakis and Vamboukas, 2005, Roussos and Tsaousis, 2005, Sprent, 1993).

Introducing the innovative method of secondary analysis of experimental data

Purpose of the proposed secondary analysis method

The proposed secondary analysis method can be applied to experimental patterns of multiple measurements, when we perform diachronic measurements on the members of the same group during the experiment in order to see the change in the values of the measurements of individual groups of participants. In particular, with this method we draw safe conclusions regarding the progress of students of diverse potential throughout the experimental application. In this way, we can safely control the degree of mobilization among indifferent or below average participants through the application of experiential-group cooperative methods and teaching techniques.

METHODOLOGY

For a better understanding of the proposed method, the experimental model of two measurements will be used, the first of which is carried out at the beginning of the process (BEFORE) and the second one at the end (AFTER). We assume that the measurement values are either quantitative on a ratio scale or interval scale or qualitative on a grading scale and follow the normal distribution. Grade scores, criteria, other metrics, ratings, placements on a likert scale, etc could be typical examples of the afore mentioned measurement values.

With this data in the context of primary analysis, we should use the parametric statistical criterion Paired samples t-test, otherwise, when the values of the measurements do not primarily follow the normal distribution, then the non-parametric criterion Wilcoxon signed rank test should be used (Moustakas, 2016, Field, 2009, Andreadakis and Vamvoukas, 2005, Roussos and Tsaousis, 2005).

After completing the primary analysis then for the secondary analysis that follows we will

- Interpret the correlation index of the values f the two measurements, which is incorporated in the parametric criterion Paired samples t-test. A high value on the correlation index basically would mean that there is correlation in the values of the two measurements, i.e. that the low values remained respectively low in the second measurement as well and that the high values remained respectively high during the second measurement too. Conversely, a low value on the index, which is often welcome, suggests that there are variations, which means that either the low values have increased significantly in the second measurement or that the high ones have been reduced.
- Create a scatter diagram, which depicts the set values of the measurements of each participant, in order to exhibit the current trends, which are not only interpreted by the correlation index in the primary analysis.
- Categorize the initial measurements in categories, (for example, (14-16) moderate participants, (17-18) good ones and (19-20) excellent performers), especially in the case where both the correlation index and the diagram exhibited interesting variations, according to the data of the present study.
- Create a new variable, where the value variations of the two measurements for each participant will be inserted.
- Control the statistical significance of the above variations of the students' categories we created using the One-way ANOVA statistical criterion if the conditions are met or else the non-parametric criterion Kruskal-Wallis ,both of which will finally demonstrate, if the variations per category in the performance of the participants in the two measurements are statistically significant (Moustakas, 2016, Field, 2009, Andreadakis and Vamboukas, 2005, Roussos and Tsaousis, 2005).

Stage 1: Implementation of the Paired samples t-test and interpretation of the integrated correlation index

In the case that we have only two measurements on the experimental planning, one (BEFORE) and one (AFTER) and our dependent variable is quantitative or likert, the values of which follow the normal distribution, then in the context of primary research we will select the paired samples t-test (Muijs, 2011, Field, 2009, Roussos and Tsaousis, 2005). (see Table 2).

Repeated measures Within-subject	Control variable	Parametric criterion	Non parametric criterion
2 Measures Before & after	Quantitative	Related samples t-test	Wilcoxon
	Likert scale		
	Nominal	-	McNemar
>2 Measures Beginning - during - end	Quantitative	One-way repeated measures ANOVA	Friedman
	Likert scale		
	Nominal	-	Cochran's Q

The table shows the statistical criteria used in one-factor analysis in repeated measures analysis. (within- subject) in relation to the kind of variables used. The results of the criterion analysis include three tables, one with the descriptive data, providing information about the averages of the values of the experimental measurements, a second one informing about the degree of correlation of the values of the two measurements and a third one with the probabilistic results, referring to the statistical significance of the descriptive results. From the analysis of the results in our example we can see that there is a clear improvement of the values of the second time measurement, the average of which improved from 16.35 to 18.45, t (19) = -4.702, p = .000. It has been stated earlier that the criterion at the primary analysis stage does not provide any particular information to the researcher except for the statistically significant variation of the two measurements and as a result this analysis might as well be complete at this point. But on the sight of the correlation index of the measurement values found in the second scoreboard of the criterion lies the stimulus for the secondary analysis as the correlation of prices appears to be particularly low (-21.9%), which might as well refer to some random value alteration, but might also indicate a specific and systematic differentiation. For example, in an experiment where this specific correlation index appears low, it might mean that the second scores have fluctuated at random, but it may also mean that "weak" participants improved their performance significantly in the second measurements after the experimental application of an innovative course (Muijs, 2011, Field, 2009, Roussos and Tsaousis, 2005, Leech, Barrett and Morgan, 2005). (see Table 3)

Table 3. Primary analysis results with paired samples t-test

Pair Samples Statistics	Mean	Std. Deviation			
Pair 1 1st Quantitative measurement	16,35	1,565			
2nd Quantitative measurement	18,45	,945			
Paired Samples Correlation	**N**	**Correlation**	**Sig.**		
Pair 1 1st Quantitative measurement 2nd Quantitative measurement	20	-,219	,049		
Paired Differences	**Mean**	**Std. Deviation**	**t**	**df**	**Sig.**
Pair 1 1st Quantitative measurement 2nd Quantitative measurement	-2,100	1,997	-4,702	19	,049

Evaluation and interpretation of the correlation index of the values of the two measurements

Although the analysis of the results of this criterion could be completed at this point, we may further continue the process to see if the alterations in the second marking are random or refer to a specific level of students. First, we observe from the value index (Correlation) that there is no correlation to the two scores (-, 219 or -21.9%) whatsoever, indicating that there were value fluctuations during the second measurement. What we need to find out is if these variations are random or if they define some systematic trends (Muijs, 2011, Field, 2009, Roussos and Tsaousis, 2005, Leech, Barrett and Morgan, 2005). (See Table 4). The proposed secondary analysis will be performed by following these steps: a) by creating a scatter chart b) by categorizing the values of the first measurement c) by creating a new variable whose values are the values' alterations of the two measurements and d) by the analysis of mean difference using one-way ANOVA or Kruskal-Wallis of the categorized values (participants' levels) of the above alterations (Field, 2009, Roussos and Tsaousis, 2005).

Table 4. Correlation index of the two measurements of the Paired samples t-test

Paired Samples Correlation	N	Correlation	Sig.
Pair 1 1st Quantitative measurement 2nd Quantitative measurement	20	-,219	,045

Stage 2: Creating a scatter diagram

The first stage of the secondary analysis involves the creation of a scatter diagram, on the horizontal axis of which the variable of the 1st measurement is placed and on the vertical axis that of the 2nd. In the diagram, if there are systematic tendencies for differentiation, three high concentration value areas may occur: a) the first area, which extends along the diagonal and includes the most correlated values, i.e. those that did not differ significantly and remained stable in both measurements, b) the second area, located in the upper left corner of the diagram (green circular sector) and includes those values, which were low (below average participants) in the first measurement but improved significantly in the second measurement; and c) the third lower-right area (red circular sector), which includes the values of the individuals who scored high in the first measurement, but did not do so in the second one. In our example, we find that there are two systemic concentrations, of which the first one (a) includes non-differentiated scores and the second one (b), which is probably of particular interest to the researcher as it refers to participants who significantly improved the low grades of the first measurement. (See Figure1)

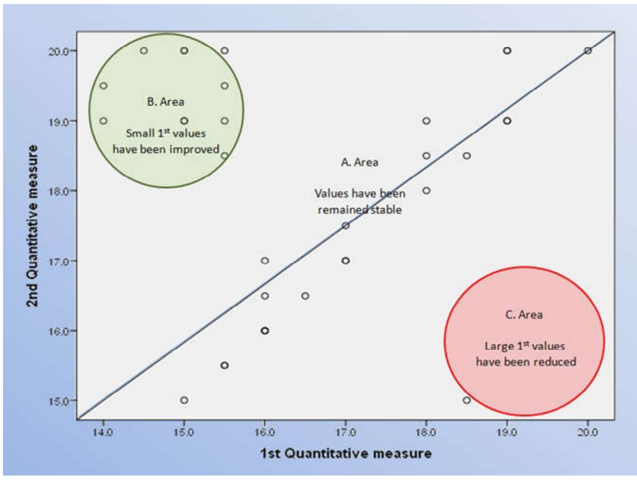

Figure 1. Scatter diagram of the values of the two measurements

Figure 1. Scatter diagram of the values of the two measurements: (a) the first area includes the most correlated values, i.e. those that remained stable in both measurements, b) the second area, located in the upper left corner of the diagram and includes those values, which improved in the second measurement; and (c) the third lower-right area with those values that have been reduced

Stage 3: Categorization of initial measurements

The previous stage provides the researcher with quick information on trends developing regarding the differences in measurements- scores, but does not provide detailed information about them, nor is it able to support the statistic significance of these trends. For this reason, we have to categorize the first scores of the participants according to their performance in three or four categories, depending on the variance of the measurement values , and then check with the One-Way ANOVA criterion or criterion Kruskal-Wallis, if the differences of the two measurements per category are statistically significant (Field, 2009, Roussos and Tsaousis, 2005). For this reason, a new (Recoded) variable should be created, which in our case will involve the categorization of the first measurement scores into three categories (14-16,17-18 and 19-20) , because the range of the prices is small.

Stage 4: Creating a variable with the differences of the two measurements

In order to be able to move on to this stage, we need to create a new computed variable that will include the differences of the two measurements for each participant. With the creation of the two new variables in the 3rd and 4th stage we have actually created a new independent variable - the categorization - and a new dependent one - the differences in the measurements – so as to apply the appropriate statistical criterion for them at the next stage of the secondary analysis and see if the differences in their measurements are statistically significant in relation to the different types of the participants we have created.

Stage 5: Checking the statistical significance of the differences of the two measurements per participant type

From the analysis of the results, we find that the program had a more positive impact on participants with lower grades and in particular that the ones with an initial grade of less than 16 scored an average increase of their score by 3.5 points, those with an initial grade of 17-18 improved it barely by 0.63 points, while the score of the excellent ones remained unchanged F (2.17) = 18.374, p = .000. (Table 5)

Table 5. Secondary analysis results with the one way ANOVA criterion

		N	Mean	Std. Dev.	S.S.	df	F	P
Score	<16	10	3,70	1,567	51,825	2,17	18,374	,000
	17-18	8	,63	,518				
	19-20	2	,00	,000				
	Total	20	2,10	1,997				

Conclusion

This specific proposal, which refers to the secondary analysis of experimental multiple measurement projects, provides the researcher with useful information not only on the overall course of the program but also on the progress of the participants ,the type of information that would not be acquired from the primary analysis (Field, 2009, Andreadakis and Vamboukas, 2005, Roussos and Tsaousis, 2005). In the experimental application of innovative teaching methods, techniques and means, the modern teacher-researcher needs to know every aspect of the system they are handling and studying, and especially whether and to what extent these innovations can motivate students who appear to be inactive and indifferent to educational daily routine due to conventional teaching methods. In conclusion, with the secondary analysis of experimental research measurements, the teacher is given the opportunity to know not only if the research that is carried out is considered scientifically successful but also the extent to which this successful effort has been able to influence and change the interest and performance of the students in general.

Conflict of Interest Statement

The author declares that there is no conflict of interest.

Funding Statement

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

Andreadakis, N. and Vamvoukas, M. 2005. Guide to the designing and drafting of a written research paper. Athens: Atrapos.

Yalamas, B. 2005. Statistical Techniques and Applications in Education Sciences. Athens: Pataki Publications.

Moustakas, L. 2016. Methods of Social Research and Statistical Analysis with SPSS (University Notes), Department of Preschool Education and Educational Planning, Winter Semester. Rhodes.

Roussos, P. and Tsaoussis, G. 2011. Statistics in behavioral sciences using SPSS. Athens: Place.

Field, A. 2009. Discovering Statistics using SPSS (3rd ed.). UK: SAGE.

Howell, D. C. 2007. Statistical Methods for Psychology (6 th ed.). Belmont, CA: Thomson Wadsworth.

Green, S. B. and Salkind, N. J. 2003. Using SPSS for Windows and Macintosh: Analyzing and Understanding Data (3rd Ed.). Upper Saddle River, NJ: Prentice

Leech, N. L., Barrett, K. C. and Morgan, G. A. 2005. SPSS for Intermediate Statistics: Use and Interpretation (2nd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.

Muijs, D. 2011. Doing Quantitative Research in Education with SPSS. UK: SAGE.

Patrick, A. 2016. What are quantitative methods of research? What is its definition, when are these methods used and what are its characteristics? PSU-University Research Office. Philippines.

Sprent, P. 1993. Applied Nonparametric Statistical Methods. London: Chapman and Hall.

****