Jordan Journal of Educational Sciences Vol. 15, No. 2, pp 251-257

Using SPSS as Instruction Assisted Program in Improving Postgraduate Students' Comprehension to Statistical Concepts

Yahya Nassar*

Received Date: 6/5/2018 **Accepted Date: 27/1/2019**

Abstract*: The main aim of this study was investigating the effects of using Statistical Package of Social Sciences software (SPSS) as computer-assisted statistics instruction in improving post-graduate students' comprehension of statistical concepts. Fifty postgraduate students were assigned randomly into three groups. In the first group, both the theoretical and SPSS sessions were presented simultaneously, whereas, in the second group, the theoretical component was introduced prior to the use of SPSS. The last group was taught using the traditional method, in which students only received theoretical lecturing about specific statistical content. A valid and reliable achievement test was used to measure students' understanding of statistics, before and after applying the mentioned methods. The results showed that the second technique was the best method, that improves postgraduate students' understanding of statistical concepts regardless of their gender. Finally, Jacobson-Truax method (JT) was used to estimate the clinical significance of the results.*

*(**Keywords**: SPSS, Instruction Assisted Software, Statistical Comprehension, the Clinical Significance).*

استخدام الحزمة الإحصائية للعلوم الإجتماعية (SPSS) كبرنامج تعليمي مساعد في تحسين استيعاب طلبة الدراسات العليا للمفاهيم الإحصائية

يحيى نصار، جامعة الامارات العربية، الامارات

ملخص: هدفت الدراسة الحالية لفحص أثر إستخدام الحزمة الإحصائية للعلوم الإجتماعية(SPSS) -كبرنامج إحصائي- تعليمي مساعد في تحسين إستيعاب طلبة الدراسات العليا للمفاهيم الإحصائية. ولتحقيق ذلك، تم توزيع (50) طالبا وطالبة من طلبة الدرسات العليا بصورة عشوائية إلى ثلاث مجموعات. في المجموعة الأولى، تم تقديم البرنامج وشرح المادة النظرية بصورة متزامنة، بينما في المجموعة الثانية، تم إستخدام البرنامج بعد تقديم شرح مفصل وعميق للمادة النظرية. أما المجموعة الثالثة، فقد تم فيها إستخدام طريقة شرح المادة النظرية بدون إستخدام البرنامج الإحصائي. وقد تم قياس مستوى إستيعاب الطلبة لبعض المفاهيم الإحصائية من خلال إختبار إحصائي تحصيلي-تم التحقق من صدقة وثباته- قبل وبعد تطبيق الطرق المذكورة. أشارت النتائج إلى أن الطريقة الثانية كانت هي الأفضل من حيث تحسين مستوى إستيعاب الطلبة للمفاهيم الإحصائية بغض النظر عن جنسهم. وأخيرا تم إستخدام طريقة جاكبسون-تروكاس Jacobson-Truax (JT) لتقدير الدلالة الإكلينيكية للنتائج.

(الكلمات المفتاحية: الحزمة الإحصائية للعلوم الإجتماعية (SPSS) ، برنامج تعليمي مساعد، إستيعاب المفاهيم الإحصائية، الدلالة الإكلينيكية)

Introduction: Many students might believe that statistics is not an easy topic. Studies have shown that while students repeatedly see statistics as the worst course taken at the college level (Hogg, 1991), many of them reveal negative attitudes and lack of excitement for the topic (Borresen, 1990; Wise, 1985). Consequently, many students might prefer to shift to qualitative research methods. And this is not because this type of research is more suitable to achieve the goals of their research. One of the reasons of such shifting could be due to students' lack of comprehension of the basic concepts of statistics, which further hinders their ability to reach the sufficient level of mastering statistical skills. However, some researchers might use qualitative methods as they believe that qualitative research can help researchers to access the thoughts and feelings of research participants, which can allow the development of an understanding of the meaning that people ascribe to their experiences (Sutton & Austin, 2015).

Many studies in behavioral sciences statistics literature address the need for modernizing graduate statistics programs (Garfield, 1995; Lehoczky, 1995; Tucker, 1995). Academia, industry, and government need graduates who can effectively communicate results and conclusions of research projects (Green & Carney, 1997).

Statistical packages such as computer-assisted statistics instruction can play a very critical role in improving students' comprehension of the different statistical concepts and procedures (Benson, 2013; Johnson & Evans, 2008). While various reasons may lead to the last conclusion, statistical packages can enable students to implement the different statistical procedures that they learned, as they provide students with an additional practice that can reinforce their understanding of the material (Gonzalez & Birch, 2000). Additionally, such computer-assisted statistics instruction enables post-graduate students to get more accurate results by means of the data analysis process. Also, they allow students to evaluate the process of teaching through which information is presented (Frederickons, Read, & Clifford, 2005).

* United Arab Emirates University, UAE.

Another important benefit of such computer-assisted statistics instruction is that students will be more able to focus on improving their skills in interpreting the results instead of spending all their time in the process of the manual computations of results. Finally, an additional advantage of the computer-based statistics instruction is its ability to provide students with effective feedback (Sosa, Berger, Saw, & Mary, 2011). These feedbacks are capacitated to confront misunderstandings and endorse mindful processing of information, which in turn facilitates the process of learning (Moreno & Mayer, 2005; Timmerman & Kruepke, 2006).

A meta-analyses study conducted by Sosa and others investigated the efficacy of computer-assisted statistics instruction. The main goal of that study was to examine a range of specific features that apparently influenced its effectiveness, such as the level of learner engagement, learner control, and the nature of feedback. In 45 experimental studies that included a control condition, computer-assisted statistics instruction provided a meaningful average performance advantage (d = 0.33). The authors' analyses showed three statistically significant findings. Larger effects were reported in studies in which treatment groups received more instructional time than control groups, in studies that recruited graduate students as participants, and in studies employing an embedded assessment. The results of a newly developed second order standardized mean effect size, revealed that additional characteristics may serve as meaningful moderators. The researchers suggested that using tight experimental control is needed to assess the importance of specific instructional features in computer-assisted statistics instruction (Sosa et al., 2011).

Through reviewing a significant number of studies in the fields of education and psychology, it can be concluded that the Statistical Package of Social Sciences (SPSS) is the most operated statistical package in the Arab countries (e.g. United Arab Emirate, Jordan, Saudi Arabia, and other countries). It is used either as an important data analysis tool in statistical courses or as an effective research tool. This might be because that package is easy to be learned by different beneficiaries, or maybe because it is a practical tool to be used to analyze quantitative data for most of the social and natural sciences.

With the increasing recognition of the importance of statistics and its wide application in various fields of theoretical and applied knowledge, most Arab Universities have mandated statistics to be taken as a compulsory subject, for most of the fields of social sciences as part of their post-graduate program requirement. The main goal of such curricula is to enable students to analyze the raw data by converting it into figures that can be easily handled and analyzed, using the appropriate statistical tool or tests (Qasem, Altrairy, & Govil, 2012). Although the literature suggests that Arab post-graduate students' comprehension of statistical concepts is not adequate, this could be improved by engaging students in more practical activities, e.g. using SPSS (Qasem et al., 2012).

In his research, Mobark (2015) investigated the impact of using SPSS on student's immediate and delayed achievement on statistics. In addition, his study attempted to explore the possible effect of SPSS on improving Arab University students' attitudes toward statistical courses. The sample consisted of (42) undergraduate students who enrolled in the course "statistical psychology" in the psychology department, at King Saud University in Saudi Arabia. The results of the study indicated that the means scores of students' post-achievement for the experimental groups were significantly greater than the mean post-achievement score of the control group. Furthermore, the results of Mobark's study showed a significant effect of using SPSS on improving students' attitudes towards the statistical courses. In a similar study, Aloraini (2012) attempted to investigate the effects of using multimedia as a teaching tool assistant, on improving students' academic achievement in the college of education at King Saud University. The results of study revealed that the performance of students who used multimedia was significantly better than the control group, in terms of their achievement level.

As can be observed through the previous studies, neither any of which attempted to examine the best order of using SPSS in improving graduate students' statistical knowledge and skills. Hence, the current study differs in terms of its attempt to explore the best way to maximize students benefits of using SPSS as a teaching tool in statistical courses. In other words, the main concern of the current study is to investigate whether using SPSS should be after providing graduate students with robust conceptual knowledge or both should be presented simultaneously.

Research Questions

The main question of the current study is: What are the effects of using Statistical Package of Social Sciences (SPSS) as computer-assisted statistics instruction in improving post-graduate students' comprehension of statistical concepts?

More specifically, the current study attempted to answer the following questions:

1- What is the best method of teaching to improve graduate students' comprehension of statistical concepts: presenting the theoretical and SPSS simultaneously, introducing the theoretical component prior of using SPSS, or using the traditional method in which students only received theoretical lecturing without using SPSS?

2- What are the clinical significances of using Statistical Package of Social Sciences (SPSS) as computer-assisted statistics instruction in improving post-graduate students' comprehension of statistical concepts?

Significance of the Study

The importance of the current study lies in exploring the best method of using statistical software, namely SPSS, to improve the Hashemite University graduate students' comprehension of statistical concepts. Such understanding will help them to use the different statistical procedures effectively, either to analyze their research data or to gain very essential academic skills that they might need to maximize their opportunity of employment and promotion. Further, enhancing the level of understanding of the statistical concepts among graduate students via using statistical software, will also contribute in achieving the goals of statistical courses learning outcomes. Finally, in general, using computer software in teaching statistical courses may participate in promoting graduate students' computer skills.

Study Objectives

The current research is focusing on using SPSS as a teaching tool with graduate students in the Hashemite University, a governmental university in Jordan. The researcher has noticed, throughout his experience working in several Arab Universities, that graduate students demonstrate difficulty in understanding statistical concepts. Consequently, the student's greatest challenge is their inability to use these statistical concepts in practical ways, to analyze the data of the different research topics, especially in the fields of psychology and education. Nevertheless, graduate students are also eager to get enough training to improve their skills in using SPSS, as they believe that such skills are very essential to analyze their research data and to increase their opportunity to secure positions at the university or other research centers.

Method

Participants

The sample of the current study was selected from graduate students in the college of education, at the Hashemite University. To achieve the aims of the current study, the researcher considered all graduate students who registered in both sections of the course "research design methods" as the sample of the study. That sample consisted of (50) graduate students, where the first group consisted of (15) students, (7) males and (8) females, the second group contained (17) students, (10) males and (7) females, and the third group composed of (18) students, (10) males and (8) females.

The researcher selected the three cases of t-test, one sample t-test, paired samples t-test, and independent samples t-test, to be the conceptual content, that will be used to examine the effects of using three different methods of teaching post-graduate students of the aforementioned statistical procedures. The basis of selecting t test to measure graduate students' statistical comprehension lies in its importance. As that statistical procedure could improve students' understating for the idea of the central limit theorem and, at the same time, the independent samples t-test can be considered as a special case of analysis of variance (ANOVA). Moreover, one more reason for such decision was that as some of the students who participated in the current study were unable to deal easily with advanced statistical procedures due to the insufficient mathematical background.

Instrument

The process of data collection commenced by applying a pre-test, which utilized a form with (20) multiple-choice items, for all students in all groups, in order to measure their pre-knowledge about the different methods of t-test. A small pilot study was conducted using (35)graduate students, who registered in the course "educational psychology" to examine the validity and the reliability evidences of the achievement test. More specifically, content and criterion validity evidences were used to check the validity.

Regarding the content validity, before using the research tool in the pilot study, (7) experts in the field of teaching research methodology had been asked to evaluate the extent to which there is an agreement between the exam and the predetermined table of specifications. That table had been used in the process of constructing that instrument. All the items obtained (80%) or more as a percentage of agreement among the different experts. However, all the modifications that have been required by the evaluations has been taken into consideration. Concerning the criterion validity evidences the sample of the pilot study has been asked to respond to another exam used in evaluating students' achievement in the same content. Yet that exam was constructed by another instructor in different University. Pearson correlation coefficient between students' scores in the research instrument and the parallel exam was (0.86) which can be considered as a satisfactory indication of the validity of the research tool.

To check the reliability, Cronbach's Alpha procedure was used to estimate the internal consistency of the applied achievement test. The value of Cronbach's Alpha coefficient was (0.88) which can be considered as high and statistically satisfactory indicator of the reliability of the test (Tuckman, 1988).

Procedure

To achieve the aim of this study, students were randomly assigned to three groups. The researcher and the students in each group agreed on a suitable time for the lecture. The first group was exposed to t-test procedures using both theoretical perspectives and SPSS simultaneously. Whereas the second group was taught t-test procedures using a different approach of teaching. Here, students received a theoretical explanation about the different procedures of t-test, prior to receiving training on conducting the same procedures using SPSS. Only traditional method (lecturing method) were presented for the different procedures of t-tests with use of the third group. Moreover, it is essential to indicate that all groups had been taught by the same researcher.

The procedure that was used with the first experimental group can be described as a combination procedure, where the researcher combined between explaining the conceptual part of t-test procedures with the implementation of each procedure on SPSS. In other words, this group was exposed to both the theoretical part and a practical part simultaneously. The second experimental group was taught the same statistical procedure by the researcher by applying a different method of teaching. Here, the researcher tried to first provide students with an appropriate explanation for the same concepts and answered all students' questions about these statistical procedures. As the next step, the same statistical concepts were applied using SPSS, by trying to integrate students' conceptual background about t-test statistical procedures and the practical experience they acquired via SPSS training. The third group was taught the same statistical concepts using the traditional method of teaching, in which the researcher presented a clear explanation about the target content allowing for constructive discussion between the students and the researcher about these concepts at the same time. After 3 weeks of implementing the aforementioned methods for all groups, via three sessions, each session is 3 hours long, post-achievement test (same test used as pre-test) was applied to explore the extent to which students' statistical comprehension had been improved.

Results and Disscussion

To achieve the aims of the current study, SPSS was used to analyze the data using various descriptive and inferential statistical procedures, namely, means and standard deviations were computed. In addition, 2-way analysis of covariance was implemented to examine the level of significance among the subjects of the study according to the variables, gender (male, female), method of instruction (combination method, lecturing then applying SPSS, and lecturing method only), and/or to explore the interaction between these variables.

Prior of Utilizing 2-way ANCOVA to analyze the data of the current study, the assumptions of that statistical procedure were examined. The distribution of the first group didn't differ significantly from the normal distribution, However, the results of K-S procedure showed negative skewness in the second and third groups. Nevertheless, fortunately, according to Leven's results the violation of the normality assumption in the mentioned groups didn't prohibit fulfilling or achieving the homogeneity assumption as the error variance was equal across all groups. According to Stevens (1996) violating the normality assumption can be ignored if the homogeneity assumption is attained.

To capture the effects of the instruction method on students' scores on the post-achievement test, as an indicator, or measure of the extent to which students were able to understand the presented content, table 1 shows the means and standard deviations of the post-achievement scores for each group.

Table (1): Pre-test and post-test means and standard deviations of the subjects according to their gender and method of instruction.

Gender	Group	Pre-test Mean	Post-test Std. Deviation	Post-test Mean	Post-test Std. Deviation	N
Male	SPSS & theoretical	4.571	0.9759	11.8571	2.41030	7
	theoretical then SPSS	5.2000	1.5491	16.5000	1.43372	10
	Theoretical	5.6000	1.2649	9.9000	1.91195	10
	Total	5.1852	1.3311	12.8519	3.47191	27
Female	SPSS&theoretical	5.3750	1.5059	14.5000	3.33809	8
	theoretical then SPSS	5.7143	1.3801	16.7143	2.42997	7
	theoretical	5.2500	1.0351	11.0000	1.60357	8
	Total	5.4348	1.2730	13.9565	3.41758	23
Total	SPSS & theoretical	5.000	1.3093	13.2667	3.15021	15
	theoretical then SPSS	5.411	1.4602	16.5882	1.83912	17
	theoretical	5.444	1.1490	10.3889	1.81947	18
	Total	5.300	1.2975	13.3600	3.45679	50

According to the results that appear in Table1, the highest post-achievement mean score was obtained by the female postgraduate students, who were taught by the second method of instruction in which the theoretical content was presented prior to the training on SPSS application. The next highest mean in terms of the value of post-achievement scores, was achieved by male postgraduate students who were instructed using the same instructional method. In addition, results on Table 1 indicates that, males and females postgraduate mean scores who were taught using the first method of instruction, in which the instructor combined or mixed the theoretical part with utilization of SPSS at the same time, were greater than the means of males and females postgraduate students who were instructed only by the traditional or lecturing method.

As the results of Table 1 indicate that the best modality of instruction is the method in which the instructor uses or trains his/her students to apply SPSS, after providing them with sufficient theoretical explanation on the concepts related to the presented statistical procedure. In other words, the theoretical part plays a very essential role in enhancing post-graduate students' ability to understand statistical concepts, such that the level of comprehension of these concepts serves as the prerequisite to use SPSS effectively. Furthermore, descriptive statistics results indicate that, overall, the level of comprehension, which was measured by students' post-achievement scores, for Arab post-graduate students who were exposed to SPSS, was better than the performance of the traditional group, the group that was exposed to the lecturing method only. Previous studies obtained by reviewing the literature suggests similar findings (Mobark, 2015; Sosa et al., 2011). The last results demonstrate the importance of using SPSS as computer-assisted statistics instruction by improving Arab postgraduate students in their knowledge and application of essential statistical concepts, that is needed to enhance their quantitative research skills. As the researcher mentioned previously, using SPSS has several benefits that strengthen student's understanding of the material (Gonzalez & Birch, 2000). A further benefit of computer-based statistics is that it may provide students with effective feedback, e.g. via comparing their manual computations with SPSS results (Sosa, et al., 2011) which is extremely important, as it facilitate student's ability to learn statistical concepts (Moreno & Mayer, 2005; Timmerman & Kruepke, 2006).

In examining whether post-graduate students' scores on the post-achievement test differ significantly according to their gender, method of instruction, and/or the interaction between these variables, by holding the pre-achievement differences between the different group as constant or as a fixed value, 2-way analysis of covariance (2-way ANCOVA) was computed. Results of this inferential statistical procedure can be found on table 2.

Table (2): Results of 2-way ANCOVA, the effects of students' gender, the method of instruction, and the interaction on students' pos-test scores

Source	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Pre-achievement	28.992	1	28.992	6.607	.014	.133
Gender	15.106	1	15.106	3.442	.070	.074
Group	322.046	2	161.023	36.694	.000	.631
gender * group	10.259	2	5.129	1.169	.320	.052
Error	188.694	43	4.388			
Corrected Total	585.520	49				

Based on Table 2, results indicated that graduate students' post-test scores did not differ significantly ($P>0.05$) according to students' gender. Such a result may have justified via the fact that students, relatively, had been distributed equally to the different natural and humanities programs regardless of their gender. In other words, students' level of comprehension to the different statistical concepts could be affected by other factors other than their gender.

In addition, according to the exposed results in Table 2, students' post-test scores differ significantly ($P\leq 0.05$) according to the used method of instruction controlling the effect of the pretest scores. These results agree with most of the findings of the related literature (Qasem, Altrairy & Govil, 2012; Mobark,2015; Moreno & Mayer, 2005; Timmerman & Kruepke, 2006). In other words, the results of the current study confirm the critical role of using SPSS as computer-assisted statistics instruction on improving post-graduate students' ability to understand statistical concepts.

As shown in Table 2, the magnitude of partial Eta squared is 0.63, which according to the effect size index is indicating to very large effect size (Stevens, 1996). This last result is very essential as it shows that the effect of the method of instruction is both statistically and practically significant. According to the adjusted mean outcomes, associated with ANCOVA results revealed in Table 3, the best technique to stimulate students' ability to comprehend statistical concepts is by focusing on establishing a robust theoretical background as an initial step. The next step is to use SPSS to provide students with the opportunity to apply what they

have learned into practice, in a practical and effective manner. In fact, the last result can be justified easily, as any statistical package assumes that the user already has an adequate statistical background that will enable him/her to use the different statistical procedures, and consequently be able to interpret results associated with each of the statistical technique. The next method in terms of improving graduate students' statistical comprehension, based on the weight of the adjusted means, was the method in which both the theoretical background and the SPSS were presented concurrently. The previous method has a positive advantage because it tries to achieve an association between the theoretical and the practical components of the process of students' learning. The last method that could be considered as the least effective way of teaching statistics for post-graduate students among the three methods, is to restrict students' learning only on the theoretical method. In general, it can be said that SPSS can improve graduate students' level of understanding statistical concepts, especially if it followed very robust theoretical background.

Table (3): The adjusted means and standard errors for the three groups.

Group	Mean	Std. Error
theoretical and SPSS	13.380	.548
theoretical then SPSS	16.510	.518
theoretical	10.373	.498

Moreover, results on table 2 show no significant effect (P>0.05) for the interaction between students' gender and method of instruction on students' post-achievement scores. In other words, the effectiveness of the different teaching methods didn't affect by students' gender.

The clinical significance refers to the extent to which the treatment was effective in terms of changing the situation of the target subjects, at an individual level, from one case to another (Campbell, 2005). Basically, clinical significance differs from statistical and practical significance due to its main concern which is the individual rather than the group(s) (Jacobson, Follette, & Revenstorf, 1984). Many procedures could be utilized to estimate the clinical significance of a study. The most known methods to estimate clinical significance of psychological and educational studies can be identified as Jacobson-Truax method (JT), GulliksenLord-Novick method (GLN), Edwards-Nunnally method (EN), Hageman-Arrindell method (HA), and Hierarchical Linear Method (HLM) (Bauer, Lambert, & Nielsen, 2004). In the current study, Jacobson-Truax method (JT) was used to estimate the clinical significance of the results. Two reasons can be attributed to this decision. Firstly, because this method is the most widely used procedure in psychological and educational studies; and secondly, because this method as indicated by different studies, provides an average estimate of the effect of treatment rather than overestimating or underestimating the effect of the treatment (Bauer et al., 2004). According to this method, two values should be computed, namely, they are the cut-off scores and the reliability change index: RCI (Nassar, 2017).

Consequently, subjects of the study can be categorized into four groups: Recovered group, those whose post-test score exceeds the predetermined cut-off score and RCI score is positive; Improved group, those who were still in need of treatment, operationally based on JT method they are those individuals whose post-test score did not exceed the predetermined cut-off score although their RCI score was positive; the Unchanged group, where there was no difference between subjects pre-scores and post scores, operationally according to the JT method these individual's post-test scores did not exceed the predetermined cut-off score, however, their RCI score was zero; and finally, the Deteriorated group, whose pre-scores were better than their post scores, operationally according to JT method they are those subjects whose post-test score did not exceed the predetermined cut-off score and whose RCI score was negative (Nassar, 2017).

Based on JT method, the cut-off score and RCI in the current study were computed as follows (Bauer et al., 2004):

The cut-off scores =The mean of the pre-scores + 2 × the standard deviation of the pre-scores.

Thus, the cut-off score of the current research was 10; the RCI is the difference between the post-score and the pre-score, at individual level, divided by the standard error for the difference scores, computed using the data of the entire sample. Thus, according to the JT method, of the subjects in the first group where the lecturing method and the SPSS were used simultaneously, (86%) could be considered as recovered, whereas, (14%) can be evaluated as improved. Furthermore, all the subjects in the second group (100%), in which the lecturing method was used prior to using SPSS as computer-assisted statistics instruction, can be evaluated as recovered. However, only (44%) of the third group, in which only the lecturing method was used, can be evaluated as recovered; (50%) of them can be considered as improved, and (6%) as unchanged. The clinical significance results confirm that using the SPSS after providing Arab postgraduate students with a strong theoretical background is the best way to improve their level of understanding of statistical concepts.

To conclude, students, in general, do not like statistical courses (Green and Carney, 1997). While this phenomenon could be due to different reasons, one striking rationale could be identified as the method of instruction that is used in the process of teaching statistical courses. The basic assumption of the current study was that using statistical packages as Computer-

Assisted Statistics Instruction, namely SPSS, in teaching statistical concepts and/ or procedures, may play an essential role in improving students' ability to understand these very concepts. In addition, one of the basic questions of the current study was concerning the best way of introducing SPSS for students in an effective and constructive manner. In that, three methods of instruction were used. Based on the results of the current study, using SPSS after providing students with sufficient and rich explanation about the target statistical procedure(s) is the best method that can be used to improve students' level of comprehension in these statistical concepts. Additionally, the effectiveness of the different proposed methods was not found to vary based on students' gender. Yet, in the case of using the best method (solid theoretical background before using the SPSS), these differences vanished. In other words, among the three methods introduced in the present study, this method (SPSS after theory) can be considered as the best mechanism to teach statistical concept, regardless of students' gender.

References

Bauer, S., Lambert, M., & Nielsen, S. (2004). Clinical significance methods: A comparison of statistical techniques. *Journal of Personality Assessment, 82*(1), 60–70.

Benson, E. (2013). Graphing calculators in teaching statistical p-values to elementary **statistics** students. *Journal of Instructional Pedagogies, 12, 1-5*.

Borresen, C. (1990). Success in introductory statistics with small groups, *College Teaching*, 38(1), 26-28.

Campbell, T. (2005). An introduction to clinical significance: An alternative index of intervention effect for group experimental design. *Journal of Early Intervention*, 27(3), 210-227.

Frederickson, N., Reed, P., Clifford, V. (2005) . Evaluating web-supported learning versus lecture-based teaching: Quantitative and qualitative perspectives. *Higher Education*, 50(4), 645–664.

Garfield, J. (1995). Modernizing statistics Ph.D. Programs: Respondent. *The American Statistician*, 49(1), 18-20.

Gonzalez, G., & Birch, M. (2000). Evaluating the instructional efficacy of computer-mediated interactive media: Comparing three elementary statistics tutorial modules. *Journal of Educational Computing Research*, 22(4), 411–436.

Green, S., & Carney, J. (1997). Statistics: Can we get beyond terminal? *Paper presented at the Annual Meeting of the American Educational Research Association*, Chicago, IL.

Hogg, R. (1991), Statistical education: Improvements are badly needed, *The American Statistician*, 45(4), 342-343.

Jacobson, N., Follette, W. & Revenstorf, D. (1984). Toward a standard definition of clinically significant change. *Behavior Therapy, 17*(3), 308–311.

Johnson, H., & Evans, M. (2008). Illustrating sampling distribution of a statistic: Minitab revisited, *Australian Mathematics Teacher*, 64(1), 35-40.

Lehoczky, J. (1995). Modernizing statistics Ph.D. programs, *The American Statistician*, 49 (1), 12-17.

Mobark, W. (2015). The effects of using SPSS software in teaching statistics course on statistical achievement and attitudes toward it, *Gesten*, 51, 73-90.

Moreno, R., & Mayer, R. (2005). Role of guidance, reflection, and interactivity in an agent-based multimedia game. *Journal of Educational Psychology*, 97(1), 117–128.

Nassar, Y. (2017). Clinical significance of psychological and educational experimental studies: Analytical study. *Journal of Educational and Psychological Studies*, 11(2), 353-369.

Qasem, M., Altrairy, A., & Govil, P. (2012). Constructing a criterion-referenced test for measuring the statistical competencies of the postgraduate students in education colleges in Yemeni Universities, ERIC database. (ED538980).

Aloraini, S. (2012). The impact of using multimedia on students' academic achievement in the college of education at King Saud University. *Journal of King Saud University-Languages and Translation, 24*(2), 75-82.

Sosa, G., Berger, D., Saw, A., & Mary, J. (2011). Effectiveness of computer-assisted instruction in statistics**:** A meta-analysis, *Review of Educational Research*, 8(1), 97-128.

Stevens, J. (1996). *Applied multivariate statistics for the social scienc*es (3rd ed.). NJ: Lawrence Erlbaum.

Sutton, J., & Austin, Z. (2015). Qualitative research: Data collection, analysis, and manageme-nt. *Canadian Journal of Hospital Pharmacy*, 68(3), 226–231.

Timmerman, C., & Kruepke, K. (2006). Computer-assisted instruction, media richness, and college student performance. *Communication Education*, 55(1), 73–104.

Tucker, J. (1995). Foreword, *The American Statistician*, 49, 1, DOI: 10.1080/00031305.1995. 10476099

Tuckman, B. (1988). *Conducting educational research*. San Diego: Harcourt Brace Jovanovich.

Wise, S. (1985). The development and validation of a scale measuring attitudes toward statistics, *Educational and Psychological Measurement*, 45(2), 401- 405.