EUROPEAN ACADEMIC RESEARCH
Vol. V, Issue 8/ November 2017

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

ISSN 2286-4822
www.euacademic.org

An Exploration on the Use, Misuse, and Abuse of Statistics in Educational Research

DYNAH D. SORIANO
JAN MARK S. MANALANSAN
IWYNE M. ABENIS
WYETH P. BERO
RAYMOND O. CARLOS
JENNALYN E. GARCIA
Don Honorio Ventura Technological State University
Bacolor, Pampanga, Philippines

Abstract:

This is a descriptive study which explored the use, misuse, and abuse of statistics in educational research. This study reviewed 21 research papers produced by the Master of Arts in Education of the Graduate School of a certain university dated from 2013 to 2015. Data were treated using frequency distribution, percentage, and rank order. As a result, The utilization of frequency distribution, weighted mean, and percentage as a basic statistical tool played a vital part in data analyses of the sample manuscripts. Basic statistical tools were the abused tools. The utilization of statistical treatments does not make inappropriateness for as long as it is aligned with the type of gathered data, research design, and the type of test to be considered if it is a test of relationship or difference. Statistical lapses were present in the majority of the sample manuscripts. There were still misconceptions of what appropriate statistical tools are to be used. Some of the conditions in choosing the right statistical tool to use were overlooked or unknowingly not satisfied. In some manuscripts, there is a lack of conformity of which of the alternative and null hypotheses must be tested to ensure the appropriate interpretation and justification of the

gathered data. At most, basic statistical tools are utilized to answer simple problems; on the other hand, as the problem becomes more complex, the use of higher statistical tools is more needed to come up with a more sophisticated solution. There is a need for the researchers to know the characteristics, appropriate use, and application of the statistical tools before utilizing them in data analysis in order to avoid misinterpretations of the gathered data.

Key words: Use, Misuse, Abuse, Statistics, Educational Research

INTRODUCTION

Everything in the educational system comes from research; from the beginning of seeking knowledge, regardless with its difficulty, up to imparting and applying of it. The development of education considers a complex and broad scope and cannot deny the needs of statistics. Statistics is the science of collecting, analyzing, and making inference from the gathered data; it is also one of the most useful and essential branches of mathematics that is not only studied theoretically by advanced mathematicians, but one that is used by researchers in different fields to organize, analyze, and summarize data. In educational research, the proper use, even its misuse, and abuse may contribute a significant effect on the development of educational system.

Knowledge in statistics is significant for researchers to attain the answers to what they are seeking. It will be easier for them to conduct the research if they have enough knowledge in statistics, especially when it comes to the data analysis. The importance of having proper knowledge in statistics, statistical tools and techniques, as well as the performance dealing with analyzing and interpreting data in different studies will be reflected to the quality produce in educational system, educational administration, educational technology, social

science, psychology of education, comparative education, curriculum construction and textbooks, educational measurement and test development, guidance and counseling, as well as in teaching behaviors.

Different statistical methods and analyses are often used to communicate research findings and results as well as to support the constructed hypotheses and gain credibility to the research methodologies and final conclusions of the study. This is essential for the researchers and the beneficiaries of the research; apart from being well informed, they can evaluate the credibility and usefulness of the information presented to them by the research work they had; and through this, they can be critical in making appropriate decisions.

Some people and organizations are aware of the contributions of statistics, especially in improving their educational research; however, there are some who are unaware of using statistics incorrectly in their research studies. Some organizations use statistics and conduct researches with limited knowledge.

Exploring a little of everything can lead to something that might benefit everyone. However, the question is that, how does an exploration on something be beneficial if there is a lack on the basic knowledge and understanding on the studies which an organization wishes to pursue?

As a student, teacher, or administrator, consider the frequency of hearing "evidenced-based practice" or "according to research". It seems that every new idea in education is research-based where education is a crucial factor in the growth and development of one's country. Educators need to be consumers (and producers) of research (Creswell, 2002). It plays a vital role in the development of human capital and is linked with an individual's well-being and opportunities for better living (Battle & Lewis, 2002).

Educational research is a cyclic process of steps that typically begins with identifying a research problem or issue of study. It involves reviewing the literature, specifying a purpose of the study, collecting and analyzing data, and forming an interpretation of information. This process culminates in a report, disseminated to audiences, and is evaluated and used in educational community (Creswell, 2002).

In dealing with the study of research, data analysis is a process of researching and a process of inspecting, clearing, transforming, and modeling data with the goal of discovering useful information, suggesting conclusion, and supporting decision-making. According to John Turkey (1961) cited by Patrick Riley (2014), "[Data analyses are the] procedures for analyzing data, techniques for interpreting the results of such procedures, ways of planning and gathering of data to make its analysis easier, more precise, or more accurate, and all the machinery and results of (mathematics) statistics which apply to analyzing data". Data analysis has multiple factors and approaches, encompassing diverse techniques under a variety of names in different business, science, and social science domains.

In analyzing data, statistics maybe used depending on its purpose. Proper statistical tools can lead to a fruitful analysis of data that produce a quality product research; but if it is not, quality research might not be obtained.

Research is a more systematic activity that is directed toward discovery and the development of an organized body of knowledge. Its systematic and objective analysis and recording of controlled observations may lead to the development of generalization, principles, or theories, resulting in prediction and possibly ultimate control of events.

Statistics is supposed to make something easier to understand, but when used in a misleading fashion, it can make anyone into believing something other than what the

data show. That is, a misuse of statistics occurs when a statistical argument asserts a falsehood. In some cases, the misuse may be accidental. In others, it is purposeful and for the benefit of the perpetrator. When the statistical tools, reasons, and techniques involved are false or misapplied, it constitutes a statistical fallacy. According to Charles Darwin in the "The Descent of Man" cited by Abelson (1995) and Porter (1995), "False facts are highly injurious in the progress of science, for they often long endure; but false views, if supported by some evidence, do little harm, as everyone takes a salutary pleasure in proving their falseness; and when it is done, one path towards error is closed and the road to truth is often at the same time opened."

Everyone should have some understanding of statistical concepts and their use in such fields as epidemiology and toxicology (Goldin, 2008).

Over the last 20 years, there had been increasing attention given to the teaching and learning aspects of statistics education (Garfield, 1993; Becker, 1996; Moore, 1997; Garfield, 1995; Garfield and Ben-Zvi, 2008; Garfield and Ben-Zvi, 2007). It is widely recognized that statistics is one of the most important quantitative subjects in a university curriculum (Watson, 1997). It is also acknowledged that teaching statistical courses is challenging because it serves students with varying backgrounds and abilities, many of whom have had negative experiences with statistics and mathematics (Garfield, 1995). Perhaps, the most critical is the fact that these courses affect life-long perceptions of and attitudes toward the value of statistics for many students, and hence, many future employees, employers, and citizens.

Statistics in Education Research over the last decade has emphasized the need for reform in the teaching of statistics with a growing body of research in this area. An increasing number of scientific publications devoted to this topic indicate

that statistics education is developing as a new and emerging discipline (Garfield and Ben-Zvi, 2008). However, research on the teaching and learning of statistics remains disconnected, fragmented, and difficult to access (Zieffler et al., 2008).

There is a growing recognition of the importance of literacy in the different aspects of people's lives and in the past few years, statistics educators have emphasized the place of statistical literacy in statistics education reforms (see e.g. Gal, 2002; Ben-Zviand Garfield, 2004). One of the implications of this was a movement to socially-based curriculum frameworks and towards applications-based approaches that teach students to think critically about social situations in which data are used and sometimes referred to as applying statistical literacy. The term "statistical literacy" has often been used in recent literatures. The most cited definition of this term is one by Wallman (1993) in her presidential address to the American Statistical Association: "Statistical Literacy is the ability to understand and critically evaluate statistical results that permeate our daily lives—coupled with the ability to appreciate the contributions that statistical thinking can make in public and private, professional and personal decisions."

Part of being statistically literate is the ability to discuss personal understandings of data, reactions to data, concerns over conclusions, and to communicate about statistical information (Gal, 2002).

The ability to communicate statistics evidence is identified as an important theme in the connection between statistics and the outside world and is a major problem for the profession. Greenfield (1993), addressing the question of communicating statistics, says that a change of culture is needed in how technical information is communicated to practitioners and the public in such a way that they can use and understand it. The best way to further statistical literacy in

the world is to educate both the consumers of statistics and those who present statistical facts to the public.

In this fast changing world, new discoveries are being revealed each day with the aid of revolutionaries in technology. The dissemination and passing of newly discovered information as well as the old information is now very easy and almost at hand of everyone in the world. New researches are being done from time to time in the whole wide world. The findings and results of these researches may minimally or greatly affect people's way of living and may bring change, positively or negatively.

This realization was instilled into the minds of the researchers to dig deeper on how statistics was used in some research papers—in the process of data gathering and more significantly to the data analysis where the conclusion depends on.

Some studies regarding the use, misuse, and abuse of statistics were already conducted—focusing on Medical Research, Information Security Research, and so on, e.g., "The Use, Misuse, and Abuse of Statistics in Information Security Research" by Ryan and Jefferson of The George Washington University. The difference of this study from these already conducted researches is that it will focus on the use, misuse, and abuse of statistics in Educational Research.

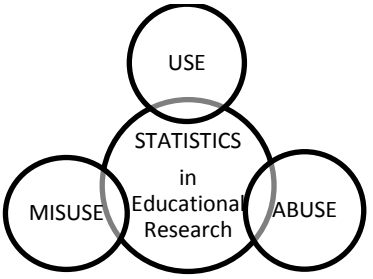

Figure 1.Use, Misuse, and Abuse of Statistics in Educational Research

Figure 1 shows the use, misuse, and abuse of statistics in educational research based on the research papers that the researchers considered for this study.

The use of statistics is inseparable with the educational research. Different statistical treatments are usually used to describe and analyze the present collected data in the pursuit of drawing conclusions. Mostly, the common statistical treatments are being abused for these are the simplest and most commonly known by the researchers. Thus, the abuse of statistics may create the misuse of statistics when these treatments are not used in full knowledge.

The researchers aimed to explore the use, misuse, and abuse of statistics in the educational research of a certain university.

Specifically, they sought answers to the following questions:

1. What are the uses of statistics in actual educational research under study?
2. What are the most commonly used, misused, and abused statistics in educational research?
3. What are the occurrences of the misuse of statistics in educational research?

The result may help statistics teachers to realize the importance of the subject that they are teaching. It will give them a clear view of their role as a statistics teacher that should teach the subject in exemplary ways. Statistics teachers will begin to have a deeper understanding on how important the statistics subject is to their students. They may start thinking innovative ways on how to teach the subject that may be easily understood by their students. They will also make their students aware on its proper use to avoid committing mistakes in using statistics in educational research. For research teachers this will give review on how statistical tools

are used on data analysis of the previous researches. Having enough knowledge on what appropriate statistical tool to be used in data analysis will help in diagnosing the accuracy of the statistical tool used and make them avoid the misuse and abuse of statistics in educational research. Hence, they easily advise their research students on the appropriate statistical tool to be used.The result of this study will give deeper understanding and awareness in committing the use, misuse, and abuse of statistics for their future research work. In such manner, they will gain information and overview on the different statistical tools. This will help them in easily selecting the appropriate statistical treatment based on the data and variables they have on their study.

The study was confined to the exploration on the use, misuse, and abuse of statistics in research papers of a certain university under the following educational areas: educational psychology, educational administration, teacher education and teaching behavior, and educational technology. Specifically, the researchers conducted an exploration on the following parts of the research papers: considered variables, type of gathered data, existing hypotheses, data analysis procedure and interpretation.

METHOD

This study is a descriptive study focused on the manuscripts recorded and produced as educational research by the Master of Arts in Education of the Graduate School of a certain university from the year 2013 to 2015. This encompassed the manuscripts within the following educational areas: Educational Administration, Educational Psychology, Educational Technology, and Teacher Education and Teaching Behavior. In line with this type of research, content analysis was employed. The manuscripts were coded from A to U.

Based on the records of the library, there are 55 manuscripts which were produced from 2013 to 2015. Twenty-one educational manuscripts from the target educational areas served as samples.

This study considered the guide proposed in the journal "Choosing the Right Procedure to Use" from the WINKS User's Guide (6th edition). In categorizing the manuscripts with respect to the educational area, the researchers used the "Scope and Functions of Educational Research" (Garden of Yokufundisa, 2014).

Frequency distribution, percentage, and ranking order were used to logically present data.

RESULTS

Uses of Statistics

Based on the sample manuscripts, Table 1 shows the findings of the researchers on the use of statistics.

Table 1. The use of statistical tools from the manuscripts.

Manuscript	Statistical tool	Use in the manuscript
A	frequency distribution, percentage, and weighted mean	To describe the factors that hinder the development of values of the children studying at DHVTSU and the suggestion of parents to reinforce the development of values of their children studying at DHVTSU by their teachers
B	frequency distribution, percentage, and mean	To describe the behavioral problems, causes, academic performance, and preventive measures
C	Frequency distribution, ANOVA, and Scheffe's test	To determine the proficiency of the pupils with regard to the different areas in mathematics
D	frequency and percentage	To determine the profile, academic performance, and multiple intelligences of pupil-respondents
E	chi-square	To test the significant difference between the assessment of the teachers and the pupils on the extent of the development of character traits by the latter, and

	weighted mean	the effectiveness of the teaching method and strategies used To describe the extent of development of character traits of the pupils
	frequency distribution	To describe the use of different evaluation tools or techniques
F	Frequency, correlation design	To know whether there was a relationship between the pupil, teacher, and school-related factors and pupils' academic performance
G	norms of interpretation, frequency, and weighted mean were used	To determine the physical and verbal abuse among intermediate pupils
H	Frequency, weighted mean	To determine the profile of the students; to determine the teaching efficacy of the respondents
I	Frequency	To describe educational leaders' decision-making styles, teachers' practices in the domains of the NCBTS, and the schools' performance
J	Frequency, weighted mean	To describe the managerial competencies and work performance
K	frequency distribution and percentage	To describe the support extended by the adapting agencies to school
L	Frequency, percentage, weighted mean	To describe the employability of LET passers, non-passers, and non-takers.
M	Frequency, percentage, weighted mean	To describe the respondents' profiles and their professional growth
N	Frequency, weighted mean	To describe the assessment of the respondents on the different factors of physical fitness
O	Frequency, weighted mean	To determine the assessment of the intermediate pupils and science teachers in the teaching practices used in teaching science
P	Frequency, weighted mean, rank order	To describe the assessment of the work values of elementary and secondary school teachers
Q	Frequency, weighted mean	To describe factors that influence and enhance work performance
R	Frequency, percentage, weighted mean, rank order	To describe the levels of job stresses, levels of symptoms, stress vulnerability, and coping mechanism of the lowland and of the upland teachers
S	Frequency, weighted mean	To describe the levels of instructional competencies
T	percentage, frequency distribution	To determine the ICT tools the

		researchers have used in their teaching
	Mean, weighted mean	To determine the degree and overall rating of effectiveness of ICT in the teaching of the teacher
U	Frequency, percentage, weighted mean	To describe the professional profile of the teachers

Commonly Used Statistical Tools

In this study, the researchers gathered all the statistical tools which were appropriately used by the sample manuscripts.

Table 2 shows the ranking of the used statistical tools of the sample manuscripts.

Table 2. The rank order from the most up to the least commonly used statistical tools.

Statistical Tool	Number of Manuscripts	Rank of Statistical Tool
Frequency Distribution	15	1
Weighted Mean	15	1
Percentage	9	3
Chi-Square	2	4
Mean	2	4
Rank Order	2	4
ANOVA	1	7
Correlation Design	1	7
Norms Of Interpretation	1	7
Scheffe's Test	1	7

Frequency distribution, weighted mean, and percentage were the most commonly used statistical tools based on the reviewed sample manuscripts. Among the sample manuscripts, five of them reflected the use of higher statistical tools namely: Chi-Square, ANOVA, Correlation Design, and Scheffe's Test. Mean, rank order, and norms of interpretation are the basic statistical tools which were used by five of the manuscripts.

Commonly Misused Statistical Tools

Table 3 presents the rank of the misused statistical tools from the most up to the least misused statistical tool.

Table 3. Commonly misused statistical tools and their ranks.

Statistical tools	Frequency	Rank
Pearson's correlation coefficient	5	1
T-test	4	2
Chi-square	4	2
ANOVA	2	4
Mann-Whitney test	1	5
Spearman's correlation coefficient	1	5
Pearson Chi-Square Coefficient of Correlation	1	5

Commonly Abused Statistical Tools

These were the reflected abused statistical tools. Table 4 shows the rank order of the statistical tools regardless of appropriateness and correctness in usage from the 21 manuscripts from the most up to the least frequently used.

Table 4. The rank order of the actual statistical tools utilized in the sample manuscripts.

Statistical Tool	Number of Manuscripts	Rank
Frequency Distribution	21	1
Mean and Weighted Mean	16	2
Percentage	13	3
Rank Order,	5	4
Pearson's correlation coefficient	5	4
Chi Square	4	6
T-Test	4	6
Descriptive Rating	4	6
ANOVA,	3	9
Standard Deviation	3	9
Spearman's Correlation Coefficient	1	11
Pearson Chi-Square	1	11
Mann Whitney	1	11
Percentile Rank	1	11

Based on Table 4, the most frequently abused statistical tool from the reviewed manuscripts is frequency distribution. On the other hand, Spearman Rank Correlation Coefficient, Pearson Chi-Square Test, Mann Whitney, and Percentile Rank are the least frequently abused statistical tools.

Occurrences of the Misuse of Statistical Tools

Table 5 displays how the statistical tools were misused in some manuscripts.

Table 5. Comparison of the actual utilization in the manuscripts and the appropriate use of the statistical tools.

Manuscript		Variable	Type of data	Type of test [relationship or difference]	Statistical tool	Misuse
B	actual	Behavioral problems; academic performance	Categorical	Test of relationship	ANOVA	ANOVA is used to test the significant difference between groups of repeated measures where data are numerical and normal or not normal (repeated measures ANOVA and Friedman ANOVA respectively), and independent groups where in data are numerical and normal (one-way ANOVA)
	appropriate		Categorical	Test of relationship	Contingency coefficient	
D	actual	Multiple intelligence type and profile; academic performance of the pupil respondents	Categorical	Test of relationship	Chi-Square	Chi-Square is appropriate to use when testing for the significant difference of categorical data
	appropriate		Categorical	Test of relationship	Contingency coefficient	
F	actual	Pupils' academic performance; teachers' educational qualification; adequacy of facilities	Categorical	Test of relationship	Pearson's correlation coefficient	Pearson product moment of coefficient correlation is properly used when testing for the significant relationship of normal data
	appropriate		Categorical	Test of relationship	Contingency coefficient	
F	actual	Pupils' academic performance; school-related factors	Categorical	Test of relationship	Chi-square	Chi-Square is appropriate to use when testing for the significant difference of categorical data
	appropriate		Categorical	Test of relationship	Contingency coefficient	
G	actual	Assessment on causes of bullying	Categorical	Test of difference	ANOVA	ANOVA is appropriate to use with numerical data
	appropriate		Categorical	Test of difference	Chi-Square	
H	actual	Teachers' teaching efficacy; institutional performance	Categorical	Test of relationship	Pearson's correlation coefficient	Pearson product moment of coefficient correlation is properly used when testing for the significant relationship of normal data
	appropriate		Categorical	Test of relationship	Contingency coefficient	
	actual	Teachers' teaching efficacy; teachers' profile	Categorical	Test of relationship	Chi-square	Chi-Square is appropriate to use when testing for the significant difference of categorical data
	appropriate		Categorical	Test of relationship	Contingency coefficient	
	actual	Educational leaders' decision making styles; the schools' performance	Categorical	Test of relationship	Pearson Chi-Square Coefficient of Correlation	Pearson Chi-Square Coefficient of Correlation is just like chi-square which is appropriate to use when testing the significant difference and when data are categorical
	appropriate		Categorical	Test of relationship	Contingency coefficient	
	actual	Teachers' practices in the domains of the NCBTS; the schools' performance	Categorical	Test of relationship	Chi-square	Chi-Square is appropriate to use when testing for the significant difference of categorical data
	appropriate		Categorical	Test of relationship	Contingency coefficient	
J	actual	Management competencies; work performance of principals	Categorical	Test of relationship	Pearson coefficient correlation	Pearson coefficient correlation is used for normal data
	appropriate		Categorical	Test of relationship	Contingency Coefficient	
K	actual	Assessment	Categorical	Test of	T-test	T-test for independent sample is

		of the school heads and of the teachers on the implementation of adapt-a-school program	difference			used when testing the difference of two groups of independent samples with normal data
	appropriate		Categorical	Test of difference	Chi-Square	
O	Actual	Assessment of teaching practices of the pupils and teachers	categorical	Test of difference	T-test for independent sample	T-test for independent sample is used when testing the difference of two groups of independent samples with normal data
	appropriate		categorical	Test of difference	Chi-Square	
Q	actual	Factors influencing work performance; factors that enhanced work performance	Categorical	Test of difference	T-test; Mann-Whitney	The analysis of data was treated with T-test and Mann-Whitney. The two are appropriate if a researcher wants to compare two groups of independent samples. T-test, specifically independent group t-test is appropriate for normal data; while Mann-Whitney is appropriate for not normal data
	appropriate		Categorical	Test of difference	Chi-Square	
R	actual	Assessment on: level of job stress; Vulnerability and coping mechanism.	Categorical	Test of relationship	Pearson's correlation coefficient; Spearman's correlation coefficient	Pearson Coefficient of correlation and Spearman rank correlation coefficient are both used in analyzing the relationship between two variables with normal data. The first treatment is use for normal data while the second is use for not normal data.
	appropriate		Categorical	Test of relationship	Contingency Coefficient	
S	actual	Levels of instructional competencies	Categorical	Test of difference	T-test	Paired T-test is used when comparing two groups-samples paired, with normal data. Another t-test is when comparing two groups – independent samples, normal data.
	appropriate		Categorical	Test of difference	Chi-Square	
U	actual	Professional profiles of teachers; adequacy if materials for P.E.	Categorical	Test of difference	Pearson's correlation coefficient	Pearson product moment of coefficient correlation is used to test the significant relationship for normal data.
	appropriate		Categorical	Test of difference	Chi-Square	

Extension of the Commonly Misused Statistical Tools: Inappropriate Hypothesis Testing

Table 6 shows the additional misuse of statistical tools in terms of testing the hypothesis.

Table 6. Sample inappropriate hypotheses testing

Manuscript	Actual Hypothesis	Misuse
A	"Based on the computed p value of 3.172 which is smaller than the table value of 4.351 with df=5 at 0.05 level of significance, the alternative hypothesis if there is significant difference between the assessment of the elementary and secondary teachers on the dimension of achievement is rejected. This means that there is no similarity on the assessment of the two groups on this dimension."	As observed, Paper A rejected the alternative hypothesis rather than the null hypothesis. This analysis has already inappropriateness. Since the data requires rejecting the null hypothesis, therefore the interpretation "there is no similarity…" is correct given that there is no similarity means there is difference. With this observation, the researchers found that the finding was not parallel to the interpretation. Of the entire test, the author kept on rejecting or not rejecting the alternative hypotheses.
	In the summary of findings, the author stated: "There is no significant difference between the assessment of the elementary and secondary teachers along the following dimension, since the obtained p value are smaller than the table value at their given degree of freedom." "The alternative hypothesis if there is a significant difference between the assessment of the elementary and secondary teachers on the importance of the work values to their work is rejected based on the computed p value of 3.179, a value which is smaller than the table value of 4.351 with the df=5 at 0.05 level of significance."	The author rejected the alternative hypothesis, and consequently, did not reject the null hypothesis. This is inappropriate.
I	"There is a significant relationship between the educational leaders' decision making styles and the schools' performance."	The author rejected the alternative hypothesis instead of the null hypothesis.
R	"In testing the alternative hypothesis…"	This paper tested the alternative hypothesis instead of the null hypothesis.
	"…the null hypothesis…is accepted."	This did not reject the null hypothesis. "Do not reject" does not mean the hypothesis is "accepted".
	Other hypothesis of the same paper was this: "Is there a significant relationship between the assessment of teachers from the lowland schools and the teachers from the upland schools on the level of their job stress along with the stressors?"	The paper rejected the null hypothesis. It is expected that there is a significant relationship. But this made the alternative hypothesis its null hypothesis because that was what was reflected on the author's hypothesis. Therefore the author stated: "There is no significant relationship between the assessment of teachers from the lowland schools and the teachers from the upland schools on the level of their job stress along with the physiological symptoms". This was the interpretation: "This means that the way the lowland teachers varies their physiological symptoms with the way the upland teachers do."
S	"There is a significant difference between the respondents assessment on the teachers instructional competencies."	All of the treatment using T-test was done like this: testing the alternative rather than the null hypothesis.
U	"There was no significant difference between the assessments of the respondents on instructional materials as proven by the tabulated t value of 2.306, p value of .001 at 0.003 level of significance."	The author made a mistake in drawing his conclusion since p value is less than the level of significance, and then the null hypothesis must be rejected. But the alternative hypothesis was rejected instead.

DISCUSSION

Ahmed (2011) stated that "Statistics … helps in planning. We cannot imagine planning without the use of Statistics. Statistics is used to drawing valid inference." The researchers found out that there were 10 statistical tools which were used in the manuscripts mainly to present the gathered data and describe the variables according to the presented data. The authors mostly used basic statistical tools like frequency distribution, mean and weighted mean, percentages, and ranking order properly; and less of the higher statistical tools were used with correctness like the ANOVA, Scheffe's test, and Chi-square test.

Based on the gathered data, most of the misused statistical tools are higher statistical tools which require higher knowledge about statistics and reasonable conditions to be considered. One of the categories of misuse of statistics is the misleading visualization of results (de Smith, 2015). The flow of the result discussion depends on the knowledge of the author of the manuscript. Consequently, if the author is not knowledgeable enough to undertake a specific research and uses statistical tools without considering the "must-be-satisfied" conditions, the way to the misuse of statistical tool and the misinterpretation of data is wide open.

The researchers found out from the reviewed manuscripts that frequency distribution is the abused statistical tool. Frequency distribution is one of the simplest and basic statistical tools which the researchers can utilize. Mean and weighted mean, and percentage are the second and third abused statistical tools, respectively.

Of the reviewed manuscripts, Pearson's correlation coefficient is the most commonly misused statistical tool. Once incorrect procedures become common, it can be hard to stop them from spreading like a genetic mutation (Altman, 2002). It

was very evident that the manuscripts did not reflect the testing of normality, which, in the first place, must not be done given that the gathered data were categorical. Ironically, most of the manuscripts used statistical tools intended for numerical type of data.

All of the manuscripts reflected the utilization of frequency distribution. This tool was the most abused statistical tool, followed by mean and weighted mean, and percentage. It can be interpreted that the abused statistical tools are mostly basic statistical tools. Contrary, higher statistical tools are the least abused statistical tool: Pearson's correlation coefficient was utilized in five manuscripts; Chi-square and T-test were utilized in four manuscripts; ANOVA was utilized in three manuscripts; and Spearman's correlation coefficient, Pearson chi-square, and Mann Whitney were utilized each in one manuscript.

Only a person who has an expert knowledge of statistical method can handle statistical data (Ahmed, 2011). In addition to this, the simplest and basic statistical tools were abused not just because they are less complex compared to other statistical tools, but because they are the tools that are significant in creating foundations to use higher statistical tools.

All of the manuscripts had categorical type of data. Ironically, the utilized statistical tools were intended for the numerical type of data, e.g., instead of using contingency coefficient, authors used ANOVA, Pearson's correlation coefficient, Spearman's correlation coefficient, and the like. Other manuscripts tested the significant relationship using tests for the significant difference, e.g., using Chi-square and ANOVA (tests of difference) instead of the Contingency Coefficient (test of relationship).

Some manuscripts reflected misconception in testing hypothesis: they tested the alternative hypothesis instead of the null hypothesis. In testing the hypothesis, statistical tools test

the null hypothesis, not the alternative hypothesis. No matter what kind of hypothesis is reflected on the research paper, the test to reject or not to reject a hypothesis is always in terms of the null hypothesis.

At most, basic statistical tools are utilized to answer simple problems; on the other hand, as the problem becomes more complex, the use of higher statistical tools is more needed to come up with a more sophisticated solution.

There is a need for the researchers to know the characteristics, appropriate use, and application of the statistical tools before utilizing them in data analysis in order to avoid misinterpretations of the gathered data: know the nature of the research design to be followed, set parameters of variables, know the type of data to be gathered, know the type of test to be implemented, and decide for the appropriate statistical tool satisfying the included conditions.

When there is numerical type of data, the test for normality must be applied. This is to satisfy the conditions in choosing the appropriate tool to be used. When there are categorical data, use statistical treatments intended for categorical data. When testing the hypothesis, always test the null hypothesis no matter what kind of hypothesis is reflected on the research study.

REFERENCES

1 Abelson, R. P. (1995). Statistics as Principled Argument. Lawrence Erlbaum Associates

2 Baldasane and Mittel, (1994) The Use, misuse, and abuse of Statistics Significance, http://www.quirks.com/articles/91994/19941101.aspx?searchID=29151901

3 Ben-Zvi and Garfield, (2004). How Students Learn Statistics Revisited.
4 Ben-Zvi, D. and Garfield, J. (2004).The Challenge of Developing Statistical Literacy, Reasoning, and Thinking, Springer.
5 Ben-Zvi, D. Using Wiki to Promote Collaborative Learning in Statistics Education, Technology Innovation in Statistics Education, 1(1)
6 Charles Teddie and Fen Yu, Journal of Mixed methods research 2007
7 Checklist for Evaluating a Research Report, Provided by Dr. Blevins (http://jan.ucc.nau.edu/pe/exs514web/How2Evalarticles.htm)
8 Dr M J de Smith (2015). Statistical Analysis Handbook (http://www.statsref.com/HTML/index.html?misuse_and_abuse_of_Statistics.html)
9 Gal, (2002). Adults' Statistical Literacy: Meanings, Components, Responsibilities. With Discussion, International Statistical Review, 70(1), 1-51.
10 Garfield, (1993).Teaching and Learning Aspects of Statistics Education. Retrieved from; http;//www.amstat.org/publication/jse/v20n2/tishkovskaya.pdf
11 Garfield, (1995). How Students Learn Statistics, International Statistical Review, 19(1), 44.
12 Goldin, (2008).Sppining Heads and spinning News; The Use, misuse, and abuse of Statistics in the media. http;//www.maa.org/meetings/calendar-evnts/spinning-heads-and-spinning-news-the-use-misuse-and-abuse-of-Statistics-in-the-media
13 Julie J.C.H. Ryan, D.Sc. and Theresa I. Jefferson, D.Sc. (George Washington University) "The Use, Misuse, and Abuse of Statistics in Information Security Research"

14 Moore, D. S. (1997). New Pedagogy and New Content; The Case of Statistics, International Statistical Review, 65(2), 123-165.
15 Porter, T. (1995). Trust in Numbers: The Pursuit of Objectivity in Science and Public Life. Princeton, N.J: Princeton University Press.
16 Review of Research on Teaching and Learning Statistics, International Statistical Review, 75(3), 372-396, doi: 10.1111
17 Scope and Functions of Educational Research, Garden of Yokufundisa (September 2014) (http://prasanthvenpakal.blogspot.com/2014/09/scope-and-functions-of-educational.html)
18 Study.com/academy/lesson/stratified-random-sample-example-definition-quiz.html
19 The handbook of research for educational communications and technology (www.aect.org/edtech/ed1/41/41-01.html)
20 Turkey, J. The Annals of Mathematical Statistics(http://projecteuclid.org/euclid.aoms)
21 Wallman, K. K. (1993). Enhancing Statistical Literacy: Enriching our Society, journal of theAmericanStatistics Education.
22 Watson, J. M. (1997). AssessingStatistical Thinking Using the Media, In TheAssessment Challenge in Statistics Education.
23 WINKS User's Guide (6th edition) (http://www.texasoft.com/WINKSManual6_Orig.pdf)
24 Zieffler, Garfield, Alt, Dupuis, Holleque, and Chang, (2008). What Does Research Suggest about the Teaching and Learning of Introductory Statistics at the College Level? A Review of the Literature, Journal of Statistics Education, 16(2)