Statistical Fallacy: A Menace to the Field of Science

Kalu Emmanuel Ogbonnaya[*], Benard Chibuike Okechi[**], Benedict Chimezie Nwankwo[***]

[*] Department of Ind. Maths and Applied Statistics, Ebonyi State University, Abakaliki
[**] Department of Psychology, University of Nigeria, Nsukka.
[***] Department of Psychology and Sociological Studies, Ebonyi State University, Abakaliki

DOI: 10.29322/IJSRP.9.06.2019.p9048
http://dx.doi.org/10.29322/IJSRP.9.06.2019.p9048

Abstract- Statistical fallacy has been a menace in the field of sciences. This is mostly contributed by the misconception of analysts and thereby led to the distrust in statistics. This research investigated the conception of students from selected departments on statistical concepts as it relates statistical fallacy. Students in Statistics, Economics, Psychology, and Banking/Finance department were randomly sampled with a sample size of 36, 43, 41 and 38 respectively. A Statistical test was conducted to obtain their conception score about statistical concepts. A null hypothesis which states that there will be no significant difference between the students' conception of statistical concepts with respect to their departments was tested using One-way Analysis of Variance (ANOVA). The hypothesis was rejected based on $F(3,157) = 23.471$ and p-value < 0.001 which suggests statistical significant difference in their conceptions. It was then found out that Statistics students have a higher understanding about statistical concepts followed by Economics, Psychology and Banking/Finance students with a low understanding of Statistical concepts using a post-hoc (Tukey HSD) test and mean chart.

Index Terms- Statistics, Fallacy, Menace, Sciences, Misuse, Misconception, fallacies, statistical methods.

I. Introduction

The science of statistics is broadly divided into descriptive and inferential statistics. In research study, statistics is needed from the beginning to the end of the study. Statistics is the systematic collection and analysis of numerical data, which enables the researcher to investigate relationships among phenomena and develop accurate and reasonable inference (McGraw−Taylor, 2007; Annapurna, 2017). The use of statistics in almost all the fields of study can never be over emphasized. Since statistics deals with data collection and data analysis which is mostly used in making inference, extreme carefulness needs to be applied when working with data because, the usefulness of data in any field of study depends greatly on the method of collection, analysis and interpretation. The conclusion drawn will be accurate if the data collection processes and the statistical methods were carefully and judiciously applied. Inappropriate use of these processes and methods will obviously lead to inaccurate and fallacious conclusions.

Statistical fallacy refers to the wrong interpretation of data which may arise in the collection, presentation, analysis and interpretation of data. Statistics are supposed to make analysis easier to understand but when used in a fallacious approach can trick the casual observer into believing something other than what the data shows. In some cases, statistical fallacy may be accidental or unintentional. In others, it is purposeful and for the benefits of the analyst.

The fallacies committed intentionally refer to abuse of statistics and the fallacy committed unintentionally refers to misuse of statistics. A misuse occurs when the data or the results of analysis are unintentionally misinterpreted due to lack of comprehension. The fault cannot be ascribed to statistics; it lies with the user (Indrayan, 2007). Some of the basic tenets of statistics are often misunderstood and in some cases, deliberately misused to lend weight to false assertions (Salter, 2012). Statistical fallacies occur when an argument's conclusion is not supported by the numerical evidence provided as premises (Klass, 2008).

Statistics, which has over the years embedded into most fields of study due to its need in research works, has become a discipline of interest. Practitioners and professionals in the fields of study where statistics is needed have delved into statistical analysis in other to remain independent. Due to the inaccurate analysis of data, the world of research has not been without fallacious assertions. Analysts wrongly interpret statistical data which may be out of their selfish interest to prove a false assertion. In media, Journalists sometimes sell their news through wrong statistical captions and analysis of a published data. In politics, analysts misinterpret statistical data in other to win the mind of the people on certain issues of interest or to make the government consent to an idea or forecast because it can be statistically proven.

This research borders on the misuse of statistics by students which is caused by inappropriate understanding of statistics, thereby leading to statistical fallacy. Nevertheless, statistical fallacy cannot be explained or written about without taking cognizance to misconceptions of statistical concepts. Statistical misuse can also be as a result of the misconceptions of its methods and techniques by students and analysts. A solid understanding of inferential statistics is of major importance for designing and interpreting empirical results in any scientific discipline (Sotos, 2007).

Kirk (2001) explained that inferential ideas seem to be especially sensitive to be misunderstood and students are often prone to fall into deep misconceptions. This is because they require students to understand and connect many abstract concepts such as sampling distribution and significance level. Statistical misconceptions are argued to hinder meaningful

learning, impede research progress and interfere with decision making. For students, such misconceptions may be generated by poor understanding reinforced by statements uttered or written by one's mentor or teacher (Huck, 2009). When these misconceptions are applied in research studies, the results of the analysis become fallacious. The use of the inappropriate statistical methods, techniques and the analysis cost time and lost, it gives harm to science and humanity. Even if the study is carefully planned to be conducted as a result of applications with errors, the misleading results might be obtained, and mislead others who will reference the study (Ercan, 2007).

Fear and wrong orientation has contributed to students' misconception about statistical concepts. The course lecturers in the undergraduate level have in most cases failed to properly mentor the students. John and David (2002) added that for one to apply any statistical method correctly, one must have information about the variables used, information about the sampling process used and a sound understanding of the theory and assumptions underlying the method. If a researcher does not use a method correctly, then conclusions may overestimate or underestimate an important relationship or effect.

When statistical fallacies are observed especially by non-scientists, it leads to the distrust in statistics where people will choose not to accept statistical inference, causing a menace to the field of sciences. Statistical fallacy is undoubtedly a very great menace in the field of sciences. Lack of appropriate training, inadequate orientation, misunderstanding and misconceptions has contributed a great deal in fallacious interpretation and data analysis. Since Aristotle's time, the generally accepted method of evaluating informal arguments has been to analyze them for common forms of fallacious reasoning (Klass, 2008). Statistical inference is central to the justification of claims across scientific fields. When statistics serve as the basis for scientific claims, it is essential that researchers interpret them appropriately; otherwise, one of the central goals of science, the justification of knowledge, is undermined. It is therefore critically important to identify and correct errors where researchers believe that a statistic justifies a particular claim when it, in fact, does not (Hoekstra & Morey, 2014).

Social cognitive theory developed by Albert Bandura in 1986 is a learning theory based on the idea that people learn by observing others. These learned behaviours can be central to one's personality. People learn by observing others, with the environment, behaviour, and cognition acting as primary factors that influence development in a reciprocal triadic relationship. Each behaviour witnessed can change a person's way of thinking (cognition). Similarly, the environment one is raised in may influence later behaviors. For example, a child raised among statisticians tends to develop interest in any statistics related course. When this interest is applied to the study of statistics, there is high probability that the student will understand statistical concepts better and will definitely avoid misuse of statistics.

The core concepts of this theory are explained by Bandura through a schematization of triadic reciprocal causation (Bandura, 2008). The schema shows how the reproduction of an observed behaviour is influenced by the interaction of the three determinants. Personal determinant is based on whether the individual has high or low self-efficacy toward the behaviour (i.e. Get the learner to believe in his or her personal abilities to correctly complete a behavior). Behavioural determinant looks at the response an individual receives after they perform a behavior (i.e. Provide chances for the learner to experience successful learning as a result of performing the behavior correctly). Environmental determinant is the aspects of the environment or setting that influence the individual's ability to successfully complete a behavior (i.e. Make environmental conditions conducive for improved self-efficacy by providing appropriate support and materials).

Cognition is usually defined as the mental processes, representations, and activities involved in the acquisition and use of knowledge. Statistical cognition is accordingly defined as the processes, representations, and activities involved in acquiring and using statistical knowledge (Beyth-Marom, Fidler & Cumming, 2008). The issues relevant in the study of statistical cognition can be explained in three aspects. One aspect is how people acquire and use statistical knowledge and how they think about statistical concepts, this is the descriptive facet of statistical cognition. The study of how people should think about statistical concepts is the normative, also an important aspect of statistical cognition as this is often what we are exposed to (e.g., in school) and it is also the standard to which our performance is usually compared. Finally, the question of closing the gap between the descriptive (the "is") and the normative (the "should") is the prescriptive which is a critical issue in statistical cognition. As such, statistical cognition is a field of theory research and application concerned with normative, descriptive, and prescriptive aspects. It focuses on (a) developing and refining normative theories of statistics and their application, (b) developing and testing theories explaining human thinking about and judgment in statistical tasks, and (c) developing and testing pedagogical tools and ways of communication for the benefit of practitioners and teachers (Beyth-Marom, Fidler & Cumming 2008). Statistical cognition improves the students understanding about statistical concepts and thereby, limits the increase in statistical fallacies committed due to misuse of statistical techniques and methods.

Whitaker (2015) explained that a student who has never been exposed to statistics will likely demonstrate understanding consistent with first level and have to progress through second level before being able to demonstrate understanding consistent with third level. With the varied student backgrounds, curricula, and standards used throughout the country, it is not reasonable to assume that students of any particular age or grade level operate statistically at the three levels. When these levels are properly harnessed by the course lecturers, the students will understand statistical concepts better.

Gupta (2011) explained the six factors that lead to statistical fallacy as bias, inconsistencies in definitions, false generalization, inappropriate comparison, wrong interpretation of statistics measures, and technical errors.

Kuhberger and Fritz (2015), researched on the significance fallacy in inferential statistics. They investigated into the intuitive understanding of the notion of significance. They described the results of two different experiments published in a major psychological journal to a sample of students of psychology, labeling the findings as significant verse non-significant. Participants were asked to estimate the effect size and

sample sizes of the original studies. It was established through the result that labeling the result of a study as significant was associated with estimations of a big effect, but was largely unrelated to sample sizes. Similarly, non-significant results were estimated as near zero in effect size. In the research, it was concluded that after considerable training in statistics, students largely equate statistical significance with medium to large effect sizes, rather than to statistical tricks like increasing sample size.

Haller and Krauss (2002), investigated on misinterpretations of significance. The study involved six (6) German Universities and the respondents from Psychology department where sorted into three (3) groups namely; Methodology Instructors, Scientific Psychologists and Psychology Students with sample size of 30, 39 and 44 respectively. In the research it was noted that 90% of Scientific Psychologist, 80% of Methodology Instructors and 100% of Psychology Students perceive at least one of the false meanings of a p-value as true. They concluded that teaching statistics to Psychology students should not only consist of teaching calculations, procedures and formulas, but should focus more on statistical thinking and understanding of the methods. This recommendation agrees with Zaidan and Ismail (2012) on statistical reasoning as one of the concepts that should be taught in class.

Therefore, this research study seeks to find out the level of students misconception and to know if there is a statistical significance difference in statistical conceptions between the students who major in statistics and those who offer more than introductory statistics courses.

II. METHOD

Participants

The population of this study comprises of 400 level students of Ebonyi State University, Abakaliki, who major in statistics and those who offer more than introductory statistics courses in their undergraduate programme. The participants were one hundred and fifty eight (158) students comprising of ninety two (92) females and sixty six (66) males from the four selected departments. 36 students from Statistics department, 38 students from Banking & Finance department, 43 students from Economics, and 41 students from Psychology department.

Instrument

A primary data was collected for this research work through conducting a statistics test using Assessment of Students Conception on Statistical Concepts Questionnaire (ASCSCQ) which was based on statistical reasoning. This was done to find out the level at which the students' understand the basic concepts of statistics and their applications which will suggest the rate at which statistical fallacy can be committed by the students. The instrument consists of four sections; Section A – socio demographic data, which consists of four (4) questions, Section B – students behaviour towards statistics which consists of three (3) questions, Section C – statistical measures and dispersion which consists of five (5) questions, Section D – use of statistical techniques and significance which consists of ten (10) questions. Section B was aimed at understanding the statistical process the students think is most prone to error, their responds towards published statistical data, and statistical concepts. Section C was designed to obtain the students' understanding in the interpretation of statistical measures. Section D was targeted at obtaining the students' ability to use and interpret statistical techniques and significance. Section C and D were basically used to score the students because it consists of a correct answer and wrong answers. The result of the test is a number ranging from 0 to 100 percent which was scored based on the selection of the correct option or the wrong option. A low score indicates the students' low understanding about statistical concepts which suggest a higher tendency to misuse the concepts. Face validity of the instrument was carried out by experts in applied statistics and research methodology. The internal consistency reliability test of the instrument was conducted using a subset of 35 Psychology students which gave a General Cronbach's Alpha of 0.79.

Procedure

Undergraduates in 400 level were randomly selected from four (4) departments that properly defined the population of the study, namely; Statistics, Banking/Finance, Psychology and Economics. A well designed statistics test was randomly sampled to forty five (45) students in each department after a brief explanation on the reason and method of the research. The students were required to answer and return the question to the researchers within 15 minutes. After conducting the test in Statistics, Banking/Finance, Psychology and Economics department, 36, 38, 41 and 43 test sheets were collected respectively. Therefore, the sample size is one hundred and fifty eight (158) respondents.

Design/Statistics

The design adopted for this study is a cross sectional survey design because it is a study which is done at a particular point in time. The data was analyzed descriptively using mean and standard deviation to determine the level of variation in the students' conception. Inferential analysis was employed using One-Way Analysis of Variance (ANOVA) to determine if there exists a significant difference between the student scores with respect to their disciplines.

III. RESULTS

Descriptive analysis was used to present the data collected on students' understanding and behaviour towards statistics and statistical data. The questions are as follows:

1. Which one of these statistical processes is error most observed?

a. Sampling b. Data Collection c. Data Analysis d. Interpretation

2. How will you treat a published statistical data?

a. Reliable b. Unacceptable c. Manipulated d. Biased

3. Statistical concepts are very difficult to understand.

a. Strongly Agree b. Agree c. Disagree d. Strongly Disagree.

Response from the Students based on behaviour towards Statistics and Statistical data.

Response to Question 1

Statistics students believed that error is most observed in data analysis and interpretation with a tie of 28% each, 47% of Banking and Finance students responded to data analysis as the statistical process where error is mostly observed. Data collection was seen as the most error prone statistical process by 64% of Economics students while 40% of Psychology students believed that interpretation of statistical data was the major area of error in statistics.

Response to Question 2

Statistics Students with 42% showed trust in a published statistical data. 34% of Banking and Finance students responded that published statistical data is reliable. 33% of the Economics students could treat published statistical data as reliable while 33% other students would treat it as manipulated. On the other hand, 54% of Psychology Students also believes that published statistical data is reliable.

Response to Question 3

High percentage of students from the three departments agreed that statistical concepts are very difficult to understand with Statistics, Banking/Finance, Economics and Psychology responding at 53%, 81%, 53% and 56% respectively.

TABLE 1: The test score of the students from the four (4) departments.

	N	Mean	Std. Deviation	Std. Error	Maximum
STATISTICS	36	59.06	17.465	2.911	93
PSYCHOLOGY	41	36.88	13.130	2.051	60
ECONOMICS	43	40.98	12.749	1.944	60
BANKING/ FINANCE	38	34.53	12.212	1.981	60
Total	158				

From the descriptive table above, we have seen that the students' scores are reasonably low in a 100 percent scoring pattern. Statistics students have the highest mean score ($\bar{X} = 59.06$) in the test with a standard deviation of ($SD = 17.47$) and maximum value (max = 96). Banking and Finance Students had the lowest mean score $\bar{X} = 34.53$ and standard deviation of ($SD = 12.21$).

TABLE 2: The ANOVA Summary Table

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	13678.713	3	4559.571	23.471*	.000
Within Groups	29916.730	154	194.264		
Total	43595.443	157			

*Significance at p<0.01

According to the table above, there was a statistically significant difference between the disciplines as determined by one way ANOVA F(3, 154) = 23.471, p < 0.01. This shows that the mean scores of the students with respect to their departments differ statistically.

TABLE 3: The Multiple Comparisons Table for the Post-Hoc Test (TukeyHSD).

(I) DEPARTMENTS	(J) DEPARTMENTS	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
STATISTICS	PSYCHOLOGY	22.178*	3.183	.000	13.91	30.45
	ECONOMICS	18.079*	3.149	.000	9.90	26.26
	BANKING AND FINANCE	24.529*	3.242	.000	16.11	32.95
PSYCHOLOGY	STATISTICS	-22.178*	3.183	.000	-30.45	-13.91
	ECONOMICS	-4.099	3.042	.534	-12.00	3.80
	BANKING AND FINANCE	2.352	3.139	.877	-5.80	10.50

ECONOMICS	STATISTICS	-18.079*	3.149	.000	-26.26	-9.90
	PSYCHOLOGY	4.099	3.042	.534	-3.80	12.00
	BANKING AND FINANCE	6.450	3.103	.165	-1.61	14.51
BANKING AND FINANCE	STATISTICS	-24.529*	3.242	.000	-32.95	-16.11
	PSYCHOLOGY	-2.352	3.139	.877	-10.50	5.80
	ECONOMICS	-6.450	3.103	.165	-14.51	1.61

*. The mean difference is significant at the 0.001 level.

The Tukey HSD test reveals that the mean difference in the scores of Statistics students was reasonably high when compared to Psychology (22.178), Economics (18.079) and Banking/Finance Students (24.529) while the mean difference between Psychology, Economics and Banking/Finance Students were low. It also showed that there was no statistical significant difference between Psychology, Economics and Banking/Finance Students while Statistics students had a significant difference with Psychology, Economics and Banking/ Finance Students ($p < 0.001$).

IV. Discussions

The findings reveal that majority of students believe that data analysis and interpretation is the major area in statistical processes where error is mostly obtained. This is in line with Midi and Imon (2007), who opined that, one does not just learn formulae and plug in numbers in the formulae, but one has to learn about the conditions or assumptions under which the statistical testing procedures can be applied. This means that students has to be taught on the necessary conditions under which a statistic should be used. Misconception of statistical techniques and methods results to error in data analysis. These in turn intensify statistical fallacy in the statistical results and interpretations.

A high percentage of students agree that statistical concepts are very difficult to understand. This indicates that most of the statistical concepts which involve techniques and methods prove to be difficult for the students to understand. Zaidan and Ismail (2012) in their research on students misconception in descriptive statistics also found out that students find it very difficult to interpret descriptive statistics. This shows that the students' lack basic knowledge of statistical concepts and thereby believe that statistical concepts are very difficult to understand. Bennett and Anway (2003) agreed that students have the greatest difficulty with probability and sampling variability. Magina and Cazorla (2008) also believed according to their research that Average presents to be a difficult concept to be understood by the students. Students with such misconceptions are prone to committing statistical fallacy because they will possibly interpret result of analysis wrongly.

The mean score of the test indicated that most of the students performed below average. This shows a high misconception about statistical concepts. This result aligns with the result of Hallar and Krauss (2002), whose investigation showed that 90% of Scientific Psychologist, 80% of Methodology Instructors and 100% of Psychology Students perceived at least one of the false meanings of p-value as true. This result also agreed with Bennett and Anway (2003) that students initially held misconceptions and were deficient in their ability to reason correctly with statistics and probability.

It also revealed that Statistics students have a higher mean score in their understanding about statistical concepts. This indicates that the students who major in statistics have a less chance of committing statistical fallacy during data collection, analysis and interpretation. The test of significance showed a statistical significant difference between students in the four departments namely; Statistics, Banking/Finance, Economics and Psychology which is because of the high score margin of the Statistics students. It also shows that there is no significant difference between Economics, Banking/Finance and Psychology Students in their conception about statistical concepts. Studying statistical concepts and techniques reduces misconception. Previous researchers added that for one to apply any statistical method correctly, sound understanding is required because it seems to be especially sensitive to be misunderstood (John & David, 2002; Kirk, 2001; Zaidan & Ismail, 2012) and statistical techniques cannot be used blindly (Midi et al, 2007).

V. Conclusion

Statistics plays a vital role in every fields of human activity (Annapurna, 2017). The challenging aspect is that most students who in no time will become statistical analysts conceive a lot of misconceptions about statistical concepts. In today's society, statistical techniques are being used with increasing rate in Education, Medicine, Social Sciences and Applied Sciences such as Engineering. They are crucial in interpretation of data and making decisions. Based on our experience and observation through seminars, conferences and consultations, we noticed some statistical practitioners often misuse some of the statistical techniques in their researches (Midi et al, 2007) which results to statistical fallacy. This no doubt indicates that most of the fallacies committed in statistics are due to lack of statistical reasoning and misconception of statistical concepts. Since students who major in statistics prove to have greater understanding about statistical concepts, it will be of great help if statistical analysis should rest only on the shoulders of Statisticians so as to reduce the rate of statistical fallacy. On the other hand, since the field of psychology majors its research on empirical analysis, it is very important that the tutors of statistics courses in the undergraduate level improves the rate at which the course is been handled. Considering the trust students have in a published statistical data, it is required that statistical assertions should be replicated before application so as to detect any possible error in the analysis.

Further research studies should be carried out in specific statistical concepts to evaluate the areas in which the students are more prone to misunderstand. Studies on teachers' misconceptions of statistical techniques and methods should also be considered in further studies.

We therefore, recommend that statistical techniques should not to be used as tools for deceit, so that the science of statistics will not lose its reputation. Analysts should be very careful while working with statistical tools. Teaching statistical courses should not only be based on formulas and calculations but also emphasis should be laid on statistical reasoning. Statistics course lecturers in the undergraduate level should devote time to embed courage in the student in learning statistics.

REFERENCES

[1] Annapurna I. (2017). Importance of Statistics and Mathematical Models in the Field of Social Sciences Research, *Imperial Journal of Interdisciplinary Research (IJIR)*, Vol. 3, Iss. 4, 619 - 627

[2] Bandura, A. (2008). Social cognitive theory of mass communication. In J. Bryant & M. B. Oliver (Eds.), *Media Effects: Advances in Theory and Research* (pp. 94-124). New York, NY: Routledge.

[3] Bennett, E. C. & Anway, D. W. (2003). Statistical and Probabilistic Reasoning and Misconceptions among Selected College Students, *University of Wisconsin-Superior McNair Scholars Journal*, volume 4, 2003.

[4] Cumming, G., & Finch, S. (2005). Inference by eye: Confidence intervals and how to read pictures of data. *American Psychologist, Vol. 60,* No. 2, 170 – 180.

[5] Cumming, G., Williams, J., & Fidler, F. (2004). Replication and researchers' understanding of confidence interval and standard error bars. *Understanding Statistics*, 3, 299 – 311.

[6] Ercan, I., Yazıcı B., Yang, Y., Özkaya, G., Cangur, S.,Ediz, B.,Kan, I. (2007). Misusage of statistics in medical research. *European Journal of General Medicine; Vol. 4* No. 3, 128 – 134.

[7] Gupta, S. C. (2011). *Fundamentals of statistics*, 6th Ed. Himalaya publication house PVT. Ltd., Delhi, India.1.13 – 1.14 & 24.1 – 24.11.

[8] Haller, H. & Krauss, S. (2002). Misinterpretations of Significance: A Problem Students Share with their Teachers. *Method of Psychological Research Online*. Vol. 7, No. 1. 1 – 20.

[9] Hoekstra, R. & Morey, R. D. (2014). Robust misinterpretation of confidence intervals. *Psychonomic Society, Inc., 3, 1 – 8.*

[10] Huck, S. W. (2009). *Statistical Misconceptions.* Psychology press, Taylor & Francis group, LLC, New York. xv – xvi.

[11] Indrayan A. (2007). Statistical fallacies in orthopedic research. *Indian Journal of Orthopaedics*, 41(1), 37-46. doi: 10.4103/0019-5413.30524.

[12] John, G. & David, G. (2002). The Misuse of statistics: Concepts, tools, and a research agenda. *Accountability in Research: Policies and Quality Assurance, Vol 9*, No 2. 65 –74.

[13] Kirk, R. E. (2001). Promoting good statistical practices: some suggestions. *Educational and Psychological Measurement, Vol. 61* No. 2, 213 – 218.

[14] Klass, G. (2008). Just plain data analysis: Common statistical fallacy in analysis of social indicator data. *Department of politics and government, illinois state university*. Avaliable at http://polmeth.wustl.edu

[15] Kuhberger, A., Fritz, A., Lermer, E., &Scherndl, T. (2015). The Significance Fallacy in Inferential Statistics. *BioMed Central Research Notes.* Vol. 84, No. 8, 1 – 8.

[16] Margina, S. & Cazorla, I. (2008). Conceptions and Misconceptions of Average: A comparative study between teachers and students. *In proceeding of the eleventh international congress on mathematical education*, Mexico.

[17] McGraw−Taylor Harris (2007). Why you need to use statistics in your research, 1–10, 355503E1/production/mcgraw−hill/booksxml/brownsaunders/ch1

[18] Midi, H., Imon, A. H. M. R. & Jaafar, A. (2007). The misconceptions of some statistical techniques in research. *Jurnal Teknologi, Vol. 47, 21 – 36.*

[19] Salter, J. (2012). Misinterpretation of statistics – an introduction. *Pharmaceutical Programming (Phuse) Journals.* Available at: http://www.phusewiki.org/docs/2012/paper/sp/sp02.pdf

[20] Sotos, A. E. C., Vanhoof, S., Van den Noortgate, W., & Onghena, P. (2007). Students' misconception of statistical inference: A review of the empirical evidence from research on statistics education. *Education Research Review. 2,* 98 – 113.

[21] Whitaker, Douglas; Foti, Steven; & Jacobbe, Tim (2015). The Levels of Conceptual Understanding in Statistics (LOCUS), *Results of the Pilot Study*, Vol. 8, No. 2, Article 3. DOI: http://dx.doi.org/10.5038/1936-4660.8.2.3. Available at: http://scholarcommons.usf.edu/numeracy/vol8/iss2/art3

[22] Zaidan, A., & Ismail, Z. (2012). Misconceptions in descriptive statistics among postgraduates in social sciences. *Sci-Verse Science-Direct, Elsevier Ltd. 46*, 3535 – 3540.

AUTHORS

First Author - Kalu Emmanuel Ogbonnaya, BSc Applied Statistics, Ebonyi State University, Abakaliki, ogbonnaya_emmanuel@yahoo.com
Second Author - Benard Chibuike Okechi, BSc Psychology, University of Nigeria, Nsukka, chibuokenikem@gmail.com
Third Author - Benedict Chimezie Nwankwo, PhD Social Psychology, Ebonyi State University, Abakaliki, nwankwomezie@gmail.com

Corresponding Author:
Ogbonnaya Kalu Emmanuel
ogboonnaya_emmanuel@yahoo.com, kaluexcel01@gmail.com
2348066796687

APPENDIX A: ASSESSMENT OF STUDENTS CONCEPTION ON STATISTICAL CONCEPTS QUESTIONNAIRE (ASCSCQ)

Note: Please fill in the blank spaces

SECTION A: Socio Demographic Data

1. Faculty: _____
2. Department: _____
3. Level: _____
4. Gender: _____

Note: Tick the option that seems most appropriate to you.

SECTION B: Behaviour towards statistics

1. Which one of these statistical processes is error most observed?
    a. Sampling
    b. Data Collection
    c. Data Analysis
    d. Interpretation
2. How will you respond towards a published statistical data?
    a. Reliable
    b. Unacceptable
    c. Manipulated
    d. Biased
3. Statistical concepts are very difficult to understand
    a. Strongly agree
    b. Agree
    c. Disagree
    d. Strongly disagree

SECTION C: Test on Statistical Measures

1. Mean (Average) of a distribution shows the
    a. Deviation
    b. *Central tendency
    c. Variation

d. Dependence

2. When the mean is far from the center of the distribution curve, it means that the distribution is

   a. Normal

   b. *Skewed

   c. Correlated

   d. Binomial

3. A high value of standard deviation indicates

   a. *High level of variation

   b. Low level of variation

   c. High level of dependence

   d. High level of independence

4. Variance tells us how

   a. How real the mean is

   b. How reliable the data is

   c. *How far each number in the set is from the mean

   d. No idea

5. What does kurtosis tell us

   a. *The level of the standard deviation

   b. The mean is typically greater than the mode

   c. The data is randomly collected

   d. No idea

SECTION D: Test on the use of Statistical Techniques and Statistical Significance

1. Assuming income has a strong relationship with expenditure, can we conclude that income is the main cause expenditure

   a. Yes

   b. *No

   c. Sometimes

   d. No idea

2. If the average score in percentage of performance for male is 43% and female is 62%. Can we conclude that the females performed better than males.

   a. Yes

b. *No

c. Sometimes

d. No idea

3. We use students t-test when we encounter

a. *Small sample size

b. Biased sample

c. Violation of assumptions

d. Sampling error

4. Which of the following can Chi-square be used for

a. Difference in mean

b. Difference in variance

c. Degree of relationship

d. *Independence

5. A test with 5% level of significance indicates

a. 5% confidence in the test

b. 0.05 confidence in the test

c. *95% confidence in the test

d. 95% error was committed

6. To compare or contrast error value (α) and p-value, we can say that

a. Error value is the same as p-value

b. Error value can be used to in estimate effect size while p-value explains the effect size

c. *Error value suggests the confidence level while p-value is the probability of chance outcome

d. Error value is always 0.05 while p-value is always 0.10

7. Small sample size can

a. Yield better result

b. Be easy to interpret

c. Make analysis more interesting

d. *Bias the conclusion

8. Conclusions drawn from sample data about a population are subject to uncertainty because

a. The data are not reliable

b. Calculations are not accurate

c. * Only part of the population is available

d. Don't know

9. To avoid biasing the results of a survey, a sample selected from a population should be

a. Large

b. * Random

c. Representative

d. Don't know

10. In doing a statistical test we

a. * Draw conclusions about populations from sample data

b. Draw conclusions about populations and then collect sample data to support these conclusions

c. Collect sample data and use the data to make assumptions about a population

d. Don't know

SCORING

Section A and B are not scored but rather presented.

Section C and D is scored based on one point per item for a correct answer and zero for any wrong answer. The total score is presented as a percentage.