

Assessing Nigerian Secondary School Students' Misconceptions in Five Basic Statistical Concepts

Festus Osadebamwen Idehen ^{1*} 

¹Department of Curriculum and Instructional Technology, Faculty of Education, University of Benin, Benin City, NIGERIA

*Corresponding Author: festus.idehen@uniben.edu

Citation: Idehen, F. O. (2020). Assessing Nigerian Secondary School Students' Misconceptions in Five Basic Statistical Concepts. *Contemporary Mathematics and Science Education*, 1(1), ep20003. <https://doi.org/10.30935/conmaths/8448>

ABSTRACT

The study was designed to assess and document students' right conceptions and misconceptions of five basic concepts in statistics. It employed the survey design to randomly select 4332 participants from 114 Senior Secondary Schools across Edo State, Nigeria. The instrument was a two-tier multiple choice diagnostic test, which was validated and with a K-R Formular 20 reliability coefficient of 0.85. Three research questions were raised to guide the study. Data were analyzed using frequency counts and percentages. Results showed low students' right conceptions of basic statistical concepts and with majority of students having some nine misconceptions. It was recommended that teachers and textbook writers should use conceptual change instructional strategies to identify and change students' misconceptions in statistics in order to improve students' understanding and achievement in mathematics.

Keywords: right conceptions, misconceptions, conceptual change, instructional strategies

Received: 1 Jun. 2020 ♦ Accepted: 1 Jul. 2020

INTRODUCTION

Statistics involves the collection, organization, analysis and interpretation of information called data. According to Doane and Seward (2007), Statistics is the science of collecting, organising, analysing, interpreting, and presenting data. In practice, Statistics involves the use of numbers within a context and involves data collection, summarising these data in some way and making interpretations and decisions. Statistical reasoning involves being able to assess how well data are collected, describe the data, draw conclusions from the data, and allow for the uncertainty that results from the use of a sample.

In Nigeria, Statistics is one of the content areas in the current senior secondary school curriculum. It is considered under the theme Everyday Statistics in the 9-year Basic Mathematics Curriculum (Federal Ministry of Education, 2007a) and under statistics and probability in the Senior Secondary School Mathematics Curriculum (Federal Ministry of Education, 1979, 2007b). The analysis of the nature of the problems which surround people's daily lives, and the knowledge about how to interpret and evaluate critically the statistically information have dominated statistical literacy through various school curriculum. The inclusion of basic statistical concepts and procedures in the school curriculum helps in the development of statistical thinking and competence in students. Mean, median, mode, pie-chart, bar-chart and histogram are some of the basic concepts students learn in statistics, which are useful in their everyday living.

Studies have shown that statistics is one of the mathematics areas teachers and students have difficulties. Oliveira Junior, Zamora, Azevedo de Oliveira and Costa de Souza (2018) observed that analysing the learners' academic status at different levels of education shows that these learners have poor performance in Probability and Statistics, and they have negative attitude towards the subject. In the study where they investigated the relationship between attitude and academic achievement of 134 students in probability and statistics at a Federal University in the state of Sao Paulo, Brazil: on utility, anxiety, trust, pleasure and motivation scales, results indicated students lack confidence in solving Statistical and Probabilistic problems. Reaburn (2011) on students' understanding of Statistical Inference, found that success was measured by grades received without having the ability to explain the principles behind Statistical Inference. That is students were applying procedural knowledge (surface knowledge) without concurrent conceptual knowledge. According to Saidi and Siew (2019), Statistics has always been perceived as one of the challenging subject areas. Among the Statistical concepts that students often have difficulty learning is the measure of central tendency concept. Saidi and Siew conducted a survey research to assess the level of Malaysian students' understanding of the measures of central tendency and attitude towards Statistics. A total of 148 Tenth Graders were sampled from nine secondary rural schools in Sabah, Malaysia. The results revealed students' understanding of measures of central tendency was at a moderate level, while their attitude towards Statistics was found to be positive.

Ishaku and Idris (2017) did a study to analyse perceived difficult topics by primary mathematics teachers ($n=50$) and pupils ($n=200$) in Dutsin- Ma, Katsina state, Nigeria. Results revealed that 10 topics out of the 28 examined were commonly perceived difficult by teachers and pupils. However, in Statistics, teachers (mean=1.92) and pupils (mean =2.01) rated pictogram and bargraphs easy to teach and easy to learn respectively. Ebisine (2010) designed a study to ascertain the level of difficulties senior secondary students encountered in understanding non-technical words (concepts) in multiple-choice tests. The study indicated that students encountered difficulty in understanding the concepts of frequency, mean, sector and quadrant. The study therefore recommended a radical shift from the pedagogical approach adopted by Mathematics teachers to a constructivist one in the teaching of senior secondary mathematics.

Poor performance of students in mathematics has been noted and has been of great concern to teachers, and students, parents and government, and the public. However, studies investigating students' performance in mathematics have been based on achievement – measuring students' success or failure alone. Most of these studies did not investigate reasons why students pass or fail, which may be due to their conceptions or understanding of mathematics basic concepts. Reaburn in his thesis reviewed some literature on students' misconceptions of central tendency. He found that the calculation of the arithmetic mean, the mode and median are simple, yet students from primary school to tertiary level have difficulties with using and understanding these concepts. For some students the mean is defined as the algorithm to calculate it and there is no understanding of what the answer might represent. They did not see the mean as representative of the data nor as this representativeness allows for comparison of data sets.

Therefore, students can hold multiple and contradictory views in concepts in mathematics and Statistics. In the classroom students may use the formal methods, but resort to their own intuitive beliefs and knowledge outside the classroom. To correct these misconceptions in students Reaburn suggested the use of the constructivist theory of learning. According to constructivist theory of learning all students come into any learning environment with their own preconceptions that may or may not be correct. Constructivist theory states that knowledge is personally constructed but socially mediated (Cakir, 2008). Hence, there is the need to assess students' conceptions or misconceptions in mathematics and Statistics. A conception of mathematics is the knowledge students conceive and provide that counts as explanation for mathematics concept. Tamir (1971, as cited in Chandrasegran, Treagust & Mocerio, 2007) had proposed the use of a two tier multiple choice test items that include responses with known alternative conceptions (or misconceptions), and also required students to justify their choice of option by giving a reason. There is therefore the need to investigate and document Nigerian students' conceptions of some basic mathematical concepts, particularly in the area of statistics.

Purpose of the Study

The study aimed at investigating and documenting students' right conceptions and misconceptions of five basic concepts in statistics- namely, histogram, pie-chart, mean, median and mode.

Research Questions

The following research questions have been raised to guide the study:

1. What percentage of students holds the right conceptions of statistical concepts under study?
2. What are the misconceptions held by students in the five statistical concepts under study?
3. What are the significant percentage differences between right answers and right conceptions?

Methodology

The study employed the survey research design. The study is part of a research conducted to assess the conceptions held by students in five mathematics areas of Number and Numeration, Algebra, Measurement Geometry and Statistics (Author, 2011). A total of 4332 subjects from 114 schools across Edo State in Nigeria were sampled. The instrument was a two-tier multiple choice diagnostic test constructed to identify right conceptions and misconceptions in limited and clearly defined content areas. Each question on the test consisted of two sections: the first section contains the answer and the second has the reason. Below is an example of an item in the instrument.

Question

The marks obtained by five pupils in a test are 0, 4, 3, 3, 5. Find the mean mark.

Answer

- (a) 2 (b) 3 (c) 5

Reason

1. It is the largest score in the test.
2. It is the score that appears most.
3. It divides the test scores equally among the pupils.

The instrument was validated for relevance and accuracy, and tested with Kuder-Richardson Formular 20 reliability coefficients of 0.76, 0.70 and 0.85 for answers, reasons and right conceptions, respectively.

Procedure for Data Analysis

Data collected were analyzed using frequency counts and percentages to answer the research questions. A right conception was scored 1 if the respondent gave the right answer and the corresponding correct reason for the item under consideration. A misconception was scored 0 if the respondent gave wrong answer and a wrong reason, or gave the right answer but with a wrong reason. A misconception connotes wrong conception. All the points were added up to give a total score ranging from 0 to 5.

The percentage per item was calculated. Misconceptions were considered significant if found in at least 10% of the subjects (Tan, Taber, Goh and Cho, 2005). A right conception was considered high if found in at least 50% of the participants or if the participants had a total score of at least 50%. A percentage difference was considered significant if found in at least 10% of the participants sampled.

RESULTS

The results are presented in tables.

Research Question One

What percentage of students holds the right conceptions of the statistical concepts under study?

Table 1. Percentage of Students Selecting each Response Combination for each Item (n=4332)

Item	Content	Content option (Answer)	Reason Option				
			No Reason	Reason 1	Reason 2	Reason 3	No Reason
1	Histogram	No answer	7.8	0.7	0.4	0.3	9.2
		A	1.8	16.0	9.5	20.2	47.6
		B	1.0	17.1*	8.6	2.7	29.5
		C	0.8	4.9	4.7	3.4	13.8
2	Pie-chart	No answer	5.5	0.1	0.7	0.3	6.6
		A	0.9	4.4	3.4	2.2	10.8
		B	1.6	20.4	32.6	25.3*	79.9
		C	0.2	0.9	0.7	0.9	2.7
3	Mean	No answer	6.3	0.4	0.7	0.6	8.0
		A	0.6	1.7	1.9	2.5	6.8
		B	1.2	3.3	24.1	36.8*	65.4
		C	0.6	12.0	2.5	4.6	19.8
4	Median	No answer	6.3	0.4	0.6	0.1	7.4
		A	1.5	3.5	49.3	3.1	57.4
		B	1.0	1.4	16.1	6.6*	25.2
		C	0.4	6.6	2.0	1.0	10.0
5	Mode	No answer	6.4	0.4	0.4	1.0	8.2
		A	0.6	2.9	6.3	1.3	11.1
		B	0.5	2.1	2.6	2.3	7.4
		C	0.8	51.8*	2.4	18.3	73.2

* Percentage of students with right conceptions

. Percentage of students with significant misconceptions

Table 2. A Summary of Right Conceptions in Statistics

Item	Concept	Option	Students with Right Conception (%)	Right Conception	
				Right Answer	Right Reason
1	Histogram	B1	17.1	Total number of straws represented by the histogram is 20	The sum of the number of straws per block for each length
2	Pie chart	B3	25.3	The largest sector shows the class labelled Q	The largest proportion of the number of Pupils in the various classes
3	Mean	B3	36.8	The mean mark of scores for the five pupils in the test is 3	It divides the test scores equally among the pupils
4	Median	B3	46.6	The median of the set of 9 raw scores is 3	It is the 5 th number in order of magnitude
5	Mode	C1	51.8	The mode of the set of numbers is 2	The number appears most often

Table 1 shows that students held right conceptions in the five concepts as follows: histogram (17.1%), piechart (25.3%), mean (36.8%), median (6.6%), and mode (51.8%).

Table 2 shows that at least 50% of the students had right conception of the concept of mode only. Therefore, students' conceptions or understanding of most of the statistical concepts were low.

Research Question Two

What are the misconceptions held by students in the five statistical concepts understudy?

To answer research question two, misconceptions (i.e. items with wrong answers and wrong reasons or right answers and wrong reasons) were extracted from **Table 1** and presented in **Table 3**.

Table 3 shows that students held some 13 misconceptions in the statistical concepts understudy. Students had 9 significant misconceptions in the concepts of histogram (20.2%), pie-chart (20.4%, 32.6%), mean (12.0%, 12.0%), median (49.3%, 24.1%, 16.1%) and mode (18.3%).

Research Question Three

What are the significant percentage differences between right answers and right conceptions?

From **Table 4**, less number of students had right conception (right answer and right reason) than right answer in each of the statistical concept. Therefore, percentage differences existed between right answer and right conception. In pie-chart, for example, the difference was as high as 50.35 percent. It was found that there was significant difference (at least 10 percent) in the percentage of students with right answers and right conceptions. Hence, all students who had right answers did not have right conception for each concept. The conclusion is that solving to arrive at a right answer by a student does not always imply that he has a clear understanding of the underlying concept.

DISCUSSION

The results are discussed under right conceptions and misconceptions.

Right Conceptions

At least, 50 percent of the students had right conception of the concept of mode. However, less than 50 percent of the students had right conceptions of the concepts of histogram, pie-chart, mean and median. These findings are corroborated by those of Ebisine (2010) and Ishaku and Idris (2017) that students and teachers found concepts in statistics difficult as a result of lack of proper understanding of basic

Table 3. A Summary of Misconceptions in Statistics

Concept	Question	Answer	Reason	% of students
Histogram	This histogram below shows drinking straws of various lengths to the nearest cm. How many straws are there altogether?	20	- The sum of the number of each length per number of straws.	8.6
		6	- The number of straws for the highest block.	20.2
		60	- The sum of the number of each length per number of straws.	4.7
Pie-chart	The pie-chart below shows the number of pupils in classes. P, Q, R, S and T. The largest sector show the class labeled	P	- The highest number of pupils in the various classes.	4.4
		Q	- The highest number of pupils in the various classes.	20.4
		Q	- The largest space the pupils occupies in the various classes.	32.6
Mean	The marks obtained by five pupils in a test are 0,4,3,3 and 5. Find the mean score.	5	It is the largest score in the test.	12.0
		3	The score that appears most.	12.0
Median	What is the median of the following set of numbers? 3,2,1,0,4,5,6,7,2	4	- It is the number in the middle position.	49.3
		3	- It is the number in the middle position.	24.1
		3	- It is the number that appears most.	16.1
Mode	What is the mode of the following set of numbers? 1,2,3,4,2,5,2.	4	The number is in the middle.	6.3
		2	- The number appears three times.	18.3

Table 4. A summary of Percentage Differences between Right Answers and Right Conceptions (n=4332)

Concept	No. of Students (Right answers)	% (Right answers) P_1	No. of Students (Right conceptions)	% (Right conceptions) P_2	% Difference $(P_1 - P_2)$
Histogram	1274	29.41	758	17.50	11.91*
Pie-chart	3452	76.69	1141	26.34	50.35*
Mean	2821	65.42	1590	36.70	28.72*
Median	1094	25.52	296	6.81	21.71*
Mode	3157	72.88	2246	51.85	21.03*

* Significant percentage difference ($\geq 10\%$)

concepts, principles, terms and symbols involved. However, saidi and Siew (2019) found that students' understanding of measures of central tendency was at a moderate level while their attitude towards Statistics was positive.

Misconceptions

The misconception of the total frequency represented by a histogram as the sum of the number of each length per number of straws, must have resulted from students counting the number of blocks and adding the scores instead of adding the frequencies represented of all the blocks or bars. Another misconception was of mode as the total frequency may be due to lack of understanding.

The misconception of the largest sector of a pie-chart were: one, the highest number of pupils in the various classes, and two, the largest space the pupils occupies in the various classes. Students' thinking is that "highest number" or "largest space" are synonymous with the largest "proportion of the number" which the largest sector represents. These wrong conceptions must have resulted from, one, textbooks not

defining concepts before solving problems involving them. Two, teachers therefore following the textbooks as the curriculum, associates "highest" or "largest" frequency with the sector with the largest area, and the class with the lowest frequency represented by the sector with smallest area.

The conceptions of the mean being defined by the "largest score in the test" or the "score that appears most" resulted from students thinking that the mode is the same as the mean. This is an error that must have resulted from lack of understanding of the concepts of mean and mode. Also, the median was misconceived as the number (score) in the middle position must have resulted from calculating the median from a set of raw scores not grouped in order of magnitude. These findings are supported by Ebisine (2010), Reaburn (2011) and Ishaku and Idris (2017) that students lack understanding and therefore are unable to explain the underlying principles in the conceptions of Statistical concepts and inferences. This is a mechanistic error as students follow rules or formulae which are not personally internalized by them (Federal Ministry of Education, 2007c). The right conception

in the median is the middle score in order of magnitude. The conception that the mode appears three times yields the correct answer but the explanation does not generally define the concept mode. This alternative conception therefore resulted from the inability of the students to generalize their specific or particular knowledge to other situations. The study reveals that mean, median and mode has each been misconceived by students for one another.

SUMMARY

The right conceptions, misconceptions and alternative conceptions are as follows:

Right Conceptions

1. The total frequencies represented by a histogram is the sum of the frequency per block for each score.
2. The largest sector of pie-chart is represented by the largest proportion of the number of pupils in the various classes.
3. The mean mark divides the test scores equally among the pupils.
4. The median is the middle score in order of magnitude.
5. The mode is the score that appears most or the score with the highest frequency.

Misconceptions

1. The total number of frequency represented by a histogram is the sum of each score per block or bar.
2. The total number of frequency represented by a histogram is the frequency of the highest block or bar.
3. The largest sector of a pie-chart represents the highest number of items in the various classes.
4. The largest sector of a pie-chart represents the largest space the items occupies in the various classes.
5. The mean of a set of scores is the score with the highest frequency (i.e., the mean is the same as mode).
6. The mean of a set of scores is the largest score.
7. The median is the score in the middle position.
8. The median is the score that appears most (i.e., median is the same as mode).
9. The mode is the score in the middle position (i.e., the mode could be the same as the median).
10. The mode is the frequency of the middle score.

CONCLUSION AND RECOMMENDATION

A number of conclusions could be drawn from the study. Students' conceptions of basic statistical concepts are low, and therefore they are at variance with conventionally defined statistical concepts. There are positive and significant relationships between right answers and right conceptions, but the percentage of students with right answers differ from those with right conceptions. Therefore, right answers do not always lead to right conceptions or clear understanding of some basic mathematical concepts. Hence, majority of students have misconceptions of these basic statistical concepts.

Teachers and textbooks could be sources of these misconceptions. It is recommended that pre-service teachers and classroom teachers should be encouraged to examine their own conceptions and actions in class. There is the need to take measures that help teachers and text writers identify their misconceptions, and on how to replace them or develop them into accepted mathematical knowledge. These measures can include training and retraining of teachers and organizing workshops and seminars for practicing teachers and text writers. In these seminars and workshops, identified misconceptions are listed and discussed, textbooks used in schools scrutinized and marked for improvements, and strategies to correct these wrong conceptions in students' agreed upon. To improve students' understanding of statistics and improve achievement in mathematics, it is necessary to develop strategies for conceptual change instruction in mathematics.

REFERENCES

- Cakir, M. (2008). Constructivist approaches to learning in science and their applications for science pedagogy: a literature review. *International journal of environmental and science education*, 3(4), 193-206. Retrieved from <https://files.eric.ed.gov/fulltext/EJ894860.pdf>
- Chandrasegran, A. L., Treagust, D. F., & Mocerino, M. (2007). The development of a two-tier multiple-choice diagnostic instrument for evaluating secondary school students' ability to describe and explain chemical reactions using multiple levels of representation. *Chemistry education research and practice*, 18(3), 293-307. <https://doi.org/10.1039/B7RP90006F>
- Doane, D., & Seward, L. (2007). *Applied statistics in business and economics*. Boston, MA: McGraw-Hill Irwin. Retrieved from <https://www.amazon.com/Statistics-Business-Economics-Mcgraw-hill-Operations/dp/0073373699>
- Ebisine, S. S. (2010). Difficulties encountered by senior secondary school students in understanding of non-technical words in mathematics multiple-choice tests. *Abacus, The Journal of Mathematical association of Nigeria*, 35(1), 97-105.
- Federal Ministry of Education (1979). *Senior secondary school mathematics curriculum*. Lagos: Federal Ministry of Education.
- Federal Ministry of Education (2007a). *9-year basic education curriculum: mathematics primary 1-3*. Abuja: Nigerian Educational Research and Development Council (NERDC).
- Federal Ministry of Education (2007b). *Senior secondary education curriculum: mathematics for JS 1-3*. Abuja: Nigerian Educational Research and Development Council (NERDC).
- Federal Ministry of Education (2007c). *Nigeria Certificate in Education (NCE) course book on mathematics, year 1*. Kaduna: National Teachers Institute.
- Idehen, F. O. (2011). *An assessment of secondary school students' conception of selected mathematical concepts in Edo State* (Unpublished Ph.D thesis), Faculty of Education, University of Benin, Benin City. Retrieved from <https://www.iwemi.com/an-assessment-of-secondary-school-students-conceptions-of-selected-mathematical-concepts-in-edo-state>

- Ishaku, A., & Idris, M. (2017). Analysis of perceived difficult topics by primary school mathematics teachers and pupils in Dutsin-Ma, Katsina State. *Abacus, Journal of the Mathematical Association of Nigeria (MAN)*, 42(1), 164-177. Retrieved from <https://www.man-nigeria.org.ng/abstract?sid=ABA-EDU-2017-1-17>
- Oliveira Junior, A. P., Zamora, P. R., Azevedo de Oliveira, L., & Costa de Souza, T. (2018). Students' attitudes towards probability and statistics on higher education. *Acts Didactica Napocensia*, 11(2), 43-56. <https://doi.org/10.24193/adn.11.2.4>
- Reaburn, R. (2011). *Students' understanding of statistical inference: implications for teaching* (Ph.D. Thesis), University of Tasmania, Australia. Retrieved from <https://iase-web.org/documents/dissertations/11.RobynReaburn.Dissertation.pdf>
- Saidi, S. S., & Siew, N. W. (2018). Assessing students' understanding of the measures of central tendency and attitude towards statistics in rural secondary schools. *International Electronic journal of Mathematics Education*, 14(1), 73-86. Retrieved from <https://files.eric.ed.gov/fulltext/EJ1227361.pdf>
- Tan, K. C. D., Taber, K. S., Goh, N. K., & Chia, L. S. (2005). The ionisation energy diagnostic instrument: a two-tier multiple-choice instrument to determine high school students' understanding of ionisation energy. *Chemistry Education Research and Practice*, 6(4), 180-197. <https://doi.org/10.1039/B5RP90009C>