Rivers State University Journal of Science and Mathematics Education
(RSUJOSME) 2(1), 171-185

DEVELOPMENT AND STANDARDIZATION OF STATISTICAL ANXIETY SCALE FOR UNDERGRADUATE STUDENTS IN RIVERS STATE UNIVERSITY

Ugada, Catherine

Department of Educational Foundations, Faculty of Education
Rivers State University, Nkpolu – Oroworukwo, Port Harcourt, Nigeria.

Email: catherine.ugada@ust.edu.ng

ABSTRACT

The study was on development and standardization of a Statistical Anxiety Scale (SAS) for undergraduates in Rivers State University. The researcher used the instrumentation research design in carrying out the study. The population of the study consists of 4020 undergraduate students of faculty of Education, Rivers State University. A sample of 200 undergraduate students was drawn from five departments in the faculty of education. This sample size was arrived at using multi-stage sampling procedure. The development of SAS was designed to effectively assess the anxiety level of the student in statistics. Generation of SAS items was done using the multivariate approach. The instrument was designed using a 4-points Likert scale. The scales construction and development of SAS was done following the Classical Testing Theory. Trial testing of the SAS was done on 20 students. After this the 39-30 items questionnaire were subjected to factor analysis. Data were analyzed using factor analysis, Cronbach Alpha, Splits-half, t-test as well as ANOVA. Result showed that construct validity was 0.71. Also, construct validity by hypotheses testing evidence there was no significant difference in the SAS scores of male and female respondents while there was a significant difference in the SAS scores of respondents from the various Educational levels. Cronbach reliability index was 0.76 while Split-half was 0.55. The raw scores was also Standardized using z and t-scores. Recommendation among others was that statisticians should use the instrument in evaluating students especially in tertiary institutions across Rivers State.

Keywords: Anxiety, Statistical Anxiety, Test Development, Standardized Test.

Cite as:
Ugada, C. (2024). Development and Standardization of Statistical Anxiety Scale for Undergraduate Students in Rivers State University. *Rivers State University Journal of Science and Mathematics Education, 2*(1), 171-185

INTRODUCTION

A student who has gone through school should be bold and have high level of mental cognitive skills, but sometimes student fails to reproduce what has been learnt when a task is given due to anxiety. When students are faced with anxiety, they are frightened, distressed, or uneasy during academic activities, this might make it hard for them to get school work done. Nwankwo (2015) stated that anxiety is "emotion, characterized by feelings of anticipated danger, tension and distress, and by arousal of the sympathetic nervous system". Implicitly, anxiety is an emotion just like sadness and happiness. The individual becomes tensed up, troubled, worried and distressed as a result of the sympathetic nervous system. Anxiety disorders are centrally characterized by the tendencies of anxiousness under challenging situations, avoidance responses towards worrisome stimuli, deficit in social relationships and a sense of distress associated responses (Awujo, Ugwu & Amadi, 2019). Psychopathologists define anxiety as a disrupting, fear-mediated avoidance, out of proportion to the danger posed by a particular object or situation. Anxiety is an emotional state arising in situations of worry, nervousness and manifested in expectation of unfavourable events as well as the physiological and psychological reaction to an expected danger, whether real or imagined.

Statistic anxiety is describe by Macher, Papousek, Ruggeri and Paechter (2015) as the apprehension that occurs when an individual is exposed to statistics content or problems and instructional situations, or evaluative contexts that deal with statistics. As statistics-anxious individuals always experience anxiety when doing statistics, statistics anxiety describes an enduring, habitual nature. It is also noted that a large proportion of students identify statistics courses as the most anxiety-inducing courses in their curriculum (Obilor & Ugada, 2023). Especially in subjects such as psychology, education, or sociology, statistics anxiety is widely spread among students. These subjects are often chosen by

students with less interest and more critical self-assessments in mathematics and science related courses. Thus, students often underestimate the extent of statistics in these subjects. Students who experience higher levels of statistics anxiety such as fear and worry towards calculation or mathematical related subjects are assumed to be more likely to procrastinate learning activities such as; postpone writing seminar papers, take time in studying for examinations, or to keep up with the weekly readings (Onwuegbuzie, 2004). Also, statistics anxiety is assumed to be related to less time spent on learning and to less efficient learning and study strategies (Macher et al., 2013). In the examination itself, statistics anxiety is related to worry and rumination and consumes processing capacity that would be needed for task performance (Macher et al., 2013).

Consequently, statistics anxiety often is regarded to be one of the most powerful negative factors of influence on performance in statistics courses (Ugada, 2024). In some instances, the developer of the test may not be directly responsible for its administration. As with the development and administration of educational tests, the format and level of difficulty of the tests themselves are highly variable and there is no general consensus or invariable standard for test formats and difficulty. Often, the format and difficulty of the test is dependent upon the educational philosophy of the instructor, subject matter, class size, policy of the educational institution, and requirements of accreditation or governing bodies. In general, tests developed and administered by individual instructors are non-standardized whereas tests developed by testing organizations are standardized. Standardization is the process of high qualities and deep understanding of subject matter. Standardization test specifies the direction for administrating and scoring the test on different groups of individuals like whom that test is designed.

A test or examination whether formal or informal is an assessment process which intend to measure student's knowledge, skill, aptitude, physical fitness, or classification in many other topics. A test may be administered verbally, on paper, on a computer, or in a pre-determined area that requires a test taker to demonstrate or perform a set of skills. Tests vary in style, rigor and requirements. A test may be administered formally or informally. An example of an informal test is a reading test administered by a parent to a child. A formal test might be a final examination administered by a teacher in a classroom or an I.Q. test administered by a psychologist in a clinic. Formal testing often results in a grade or a test score. An exam is meant to test a persons' knowledge or willingness to give time to manipulate that subject. In every testing, it either comes in a standardized or non-standardized format. Also, it is observed that a test may come in various forms. It may be a cognitive one or affective which often requires the use of a questionnaire.

Questionnaire is a set of items designed to measure one or more underlying constructs or latent variables (Uwah, 2021). These latent traits, to a great extent are invisible to mere observation and abstract thereby, requiring special tools carefully constructed before they could be unfolded. In other words, it is a set of objective and standardized self-report questions whose responses are then summed up to yield a score. Scale items are indicators of the measured construct and hence the score is also an indicator of the construct. Over the past years, such instruments became popular in psychology mainly because they provide multiple related pieces of information on the latent construct been assessed (Raykov, 2012).

Construction of scale, is the act of assembling or/and writing the most appropriate items that constitute test questions for a target population. It is noted that the effective scale construction has a serious impact on the research extrapolations, touching first the quality and the size of the effects obtained and second the statistical significance of those effects or in other words the accuracy and sensitivity of the instruments (Price, 2017). According to Irwing and Hughes (2018), successful tests are developed generally due to some combination of the three following conditions; the theoretical advances, empirical advances as well as the practical or market need. In anyway, all these procedures or purposes should be in line with a standard. It should be noted that the development of any instruments involves some basic process which the developer must adhere to. This process of scale development is completed in five steps. They include defining the measured trait, Generating a pool of potential Likert items, having the items rated by a panel of experts, selecting the items to retain for the final scale and finally, administering the scale which will involve reversing items that measure something in the opposite direction of the rest of the scale. Similarly, Furr (2011) also described it as a process completed in five steps: (a) Defining the

Construct as well as the context of measurement, (b) Choosing a response format, (c) Assembling the initial item pool, (d) Select and revise items and (e) Evaluate the psychometric properties. Steps (d) and (e) are an iterative process of refinement of the initial pool until the properties of the scale are adequate. Opara (2016) stated that constructs in psychology are not directly observable, thus developers have to first define a general philosophical foundation to connect the construct to a set of observable traits or behaviors before the construct can be operationalized. Deciding on the construct is usually based on a review of related literature, along with consultation with subject-matter experts. Then a concise, clear and precise definition of the construct is generated. Using this definition, the item content is specified with precision and clarity of scale.

Scale response format denotes the way items are worded and responses are obtained and evaluated (Furr, 2011). Ugada and Ubulom (2024) are also of the view that common scale formats include the Guttman Scaling, the Thurstone Scaling as well as the Likert Scaling. However, Price (2017) stated that the classical measurement model is more suitable for scales with items being approximately equivalent sensors of the measured construct, like Likert etc. Along with specifying the response format, a parallel step in developing a questionnaire is assembling and/or devising items for the initial pool. The content specification of an instrument requires that the developer: 1) operationalizes the construct by specifying an exhaustive list of potential indicators (items) of the target construct, 2) select from this list the representative sample of indicators then vetting of the items by experts (Dimitrov, 2012). This is perhaps one of the most important steps of the process, since no subsequent statistical operation could counterbalance poorly stated or absent items.

The item generation phase is completed when an expert panel reviews the item pool. The items generated are reviewed for quality and relevance by the expert panel or /and by pilot testing. Generally, after reviewing items by expert groups it is also a common practice to pilot test items to acquire data for a first item analysis. Instrument development is just a step in the entire process. Anuku (2011) stated that after the researcher might have developed item statements, he needs to go further to explain if such item has the pre-requisite characteristics that is needed to elicit relevant response from the respondents. This process of ensuring that the instrument possesses these qualities is known as validation of the instrument.

As stated earlier, any good test should be reliable and valid. A combination of these is the entire validation process. To Obilor (2021), validity of an instrument deals with how well an instrument measures what it was designed to measure. It is the extent to which an instrument is effective in doing what it was designed to do. It is important to note that validity is considered the most important in the selection and use of any instrument. This will of necessity include test validity as well as test reliability and general item analysis. Furthermore, a good item analysis will include determination of the item difficulty, discrimination as well as the item distracter.

There are a variety of techniques for performing an item analysis, which is often used, for example, to determine which items will be kept for the final version of a test. Item analysis is used to help "build" reliability and validity of the test from the start. Item analysis can be both qualitative and quantitative. The former focuses on issues related to the content of the test, eg. content validity. The latter primarily includes measurement of item difficulty and item discrimination. Furthermore, Rivera (2017) stated that writing the test item does not produce an item ready to be tested until it is validated. There have been multiple techniques established for reviewing item-objective congruence based on large-scale assessments to small classroom assessments, such as the use of empirical techniques similar to norm-referenced testing, expert judgment used to calculate the index of item-objective congruence, a rating of item-objective match on a 5-point scale conducted by experts, and the use of a matching task. It should also be noted that writing the test item does not produce an item ready to be tested until it is validated. Hambleton and Rogers (2019) provided three features to focus on when reviewing a criterion reference test (CRT) item's content: (1) item validities, (2) technical quality, and (3) representativeness (p. 18). While these were for CRT tests, the same rules can be applied to multiple-choice question. These three guidelines were based on expert judgment to, "assess the degree to which the sample of items in the test is representative of some defined domain (Hambleton & Rogers, 2019). Haladyna (2014) offered three main characteristics that pertain to item validation: (1) a review of the test item from item development

procedures, (2) pilot testing of the instrument, (3) an analysis of the statistical study of item responses, which involves standardization.

A standardized test is a type of test that all the item properties are being determined and which could be administered to a wide range of testees at the same time and which also has a uniform scoring process. In a standardized test, this is administered and scored in a uniform and consistent way in order to ensure more generalization of its reliability. Standardized tests are often used in education, professional certification, psychology, the military, and many other fields. Orluwene (2012), sees standardized test as a type of test constructed by test experts and subject specialists, it is administered under a uniform set of condition based on a large normative sample.

On the other hand, a non-standardized test is usually flexible in scope and format, variable in difficulty and significance. Since these tests are usually developed by individual instructors, the format and difficulty of these tests may not be widely adopted or used by other instructors or institutions. A non-standardized test may be used to determine the proficiency level of students, to motivate students to study, and to provide feedback to students. In some instances, a teacher may develop non-standardized tests that resemble standardized tests in scope, format, and difficulty for the purpose of preparing their students for an upcoming standardized test. Finally, the frequency and setting by which a non-standardized tests are administered are highly variable and are usually constrained by the duration of the class period. A class instructor may for example, administer a test on a weekly basis or just twice a semester. Depending on the policy of the instructor or institution, the duration of each test itself may last for only five minutes to an entire class period.

Keeping in mind the importance of statistics as a core area in psychology and the occasional anxiety students experience in this area, it has been imperative that one develop a well-structured instrument which this ability could be assessed or diagnosed so that individuals know their abilities. Also developing an adequate instrument that measures the anxiety level of students in this area is most important in enhancing proper learning of the course. This backdrop however is the necessary force that propels the researcher to develop and standardize the statistical anxiety scale for undergraduates in Rivers state.

Purpose of the Study

The purpose of the study is to develop and standardize “Statistical Anxiety Scale” (SAS) for undergraduates in Rivers State University. The specific objectives of the study are to;

1. Develop the items of a Statistical Anxiety Scale (SAS) for undergraduates in Rivers State University.
2. Determine final items of the Statistical Anxiety Scale (SAS) via Principal Component Analysis.
3. Determine the relationship between Statistical Anxiety Scale (SAS) develop using:
    a. Convergent Validation Process.
4. Determine the reliability of the Statistical Anxiety Scale (SAS) using:
    a. Cronbach Alpha method of reliability.
    b. Split-half method of reliability.
5. Standardize the Statistical Anxiety Scale (SAS) using;
    a. Z-score
    b. T-score

Research Questions

The following research questions guided the study;

1. What are the items of the Statistical Anxiety Scale (SAS) for undergraduates in Rivers State University
2. What are the final items of the Statistical Anxiety Scale (SAS) using Principal Component Analysis (PCA)
3. What are the relationship between Statistical Anxiety Scale (SAS) using:
    a. Convergent Validation Process.
4. What are the reliability of the Statistical Anxiety Scale (SAS) using:
    a. Cronbach Alpha method of reliability.

b. Split-half method of reliability.

5. What are the Standardize Statistical Anxiety Scale (SAS) using:
   c. Z-score
   d. T-score

METHODOLOGY

The researcher used the instrumentation research design in carrying out the study. Here, the researcher developed and validated the "statistical Anxiety Scale (SAS) for evaluating of undergraduates' level of anxiety in statistics. The population of the study consists of all undergraduate students in Faculty of Education, Rivers State University. As at the time of the study, there are 4024 students in the faculty as at 2021/2022 academic session. (Source: Faculty office, 2022). This population was chosen because they had all the qualities to provide adequate data that will help the researcher in completing the study, and undergraduate students were considered in the faculty of education because of the extent of anxiety they exhibit in cause of leaning research and statistics.

A sample of 200 undergraduate students were drawn using 5% from the population. The multi-stage sampling technique was adopted to assist in getting representative from five departments as shown in table one below. First, the researcher used purposive sampling technique to select only students from the Faculty of Education. At stage two, the researcher applied simple random sampling technique by ballot to draw five departments from the faculty. Names of all the departments were written and folded in a piece of paper and with blindfolded the researcher drew 5 pieces which revealed the department used. These departments included; Library and Information Science (LIS), Adult Education and Community Development, Educational Management, Business Education and Science Education. Also, the researcher used stratified non-proportionate sampling technique to select 10 undergraduates' students from all the 4 levels. This gave a total of 40 students per department and a grand total of 200 for the five departments.

The development of the instrument SAS was designed to effectively assess the anxiety level of the student in statistics. Therefore, it followed the procedure below. Generation of SAS items was done using the multivariate approach. Individual items were generated by the researcher while some were adapted from the works of others. The instrument was designed using a 4-piont Likert scale divided into two sections (A and B). Section A of the instrument contains personal details of the respondents which include gender and educational level of the students. It also contains instructions on how to go about the instrument, and reported to it. DeMaio and Landreth, (2004) as well as Willis, Schechter, and Whitaker (2000) all stated that before carrying out pilot testing of items, it is common practice to subject items to a process of review. Probably the most common and arguably useful form of review is the use of expert groups. In the light of this, the researcher subjected the instrument to test experts in measurement and evaluation. Thorough vetting was carried out before arriving at the final component of the test before pilot testing. The scales construction and development of SAS was done following the Classical Testing Theory. Test and non-test experts would be able to score the constructed and standardized SAS by simply calculating the total score in each sub-scale. Note that the instrument is rated in a 4-point Liker. For all positively keyed items, the calculated weight is on the scale of 4, 3, 2, 1 and 1,2,3,4 for negatively scored items. Hence, minimum and maximum score of SAS in each of the sub-scale was on the total number of items finally accepted.

Trial testing of the SAS was done as stated by Orluwene (2012) who identified this process as administration of generated test items to group of persons similar to those intended to take the final test. Kpolovie (2012) also stated that trial testing gives a clear picture of what the main process will be and also adjudge if to modify the items in any meaningful way before carrying out the main study. Therefore, the researcher administered the instrument to 25 students and scores generated were subjected to various statistical analysis with factor analysis as the pre-requisite analysis which all other analysis were based. The respondents scores on the 39 items questionnaire were subjected to factor analysis. Before the items, the Kaiser-Meyer-Olkin value indicated the following.

Table 1: Representing the Sample Size from the Five Departments

	YR 1	YR 2	YR 3	YR 4	TOTAL
LIS	10	10	10	10	40
ACD	10	10	10	10	40
BUS EDU	10	10	10	10	40
EDM	10	10	10	10	40
SED	10	10	10	10	40

Validation

Determination of validity of SAS was done using factor Analysis.

Table 2: KMO and Bartlets Test confirming Proceeding to PCA

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.810
Bartlett's Test of Sphericity	Approx. Chi-Square	6209.997
	Df	4005
	Sig.	.000

As observed, KMO is significant by being greater than .60 (.81> .60) as stipulated by Kaiser (1974) meaning that it is proper to proceed with the factor analysis. This indicates that it is suitable to proceed for the factor analysis.

Table 3: Shows the result of the factor analysis at the validation stage.

Rotated Component Matrix

	Component	
	1	2
VAR00008	**.765**	
VAR00036	**.729**	
VAR00034	**.704**	
VAR00006	**.672**	
VAR00032	**.653**	
VAR00014	**.639**	.570
VAR00009	**.630**	
VAR00007	**.619**	
VAR00013	**.607**	
VAR00035	**.603**	
VAR00020	**-.581**	.573
VAR00018	**-.559**	
VAR00033	**-.552**	
VAR00022	**.546**	.414
VAR00019	**.542**	
VAR00010	**.516**	.422
VAR00005	**.487**	
VAR00037	**.455**	
VAR00031	**.445**	
VAR00017	**.423**	
VAR00029	**.418**	
VAR00002		
VAR00028		
VAR00027		
VAR00025		
VAR00004		
VAR00024		**.816**
VAR00016	.427	**.674**
VAR00023		**.643**
VAR00039		**.599**
VAR00015		**.571**
VAR00021		**.517**
VAR00003		**.479**
VAR00001		**.468**
VAR00012	.409	**.414**
VAR00030		
VAR00038		
VAR00026		
VAR00011		

The table above has shown that items 8, 36, 34, 6, 32, 14, 9, 7, 13, 38, 35, 20. 18, 33, 22, 19, 10, 5, 37, 31, 17, 29 loaded well in factor 1 while items 24, 16, 23, 39, 15, 21 3, 1 and 12 loaded well into factor 2. On the other hand, items 2, 28, 27, 25, 4, 30, 38, 26 and 11 did not load in the scale. This means that in all 9 items were rejected while 30 were accepted based on the criterion of 0.40 (Kaiser, 1974). Hence the questionnaire below is the final version of the pilot study showing the items numbers deleted.

Table 4: Final version of the pilot study showing the items numbers deleted

S/N	ITEMS (I feel anxious When)	SA	A	D	SD
1	Studying for an examination in a statistics course.				
2	Interpreting the meaning of a table in a journal article.				
3	Going to ask my statistics teacher for individual help with material I am having difficulty understanding.				
4	Realizing the day before an exam that I cannot do some problems that I thought were going to be easy.				
5	Asking a private teacher to explain a topic that I have not understood at all.				
6	Reading a journal article that includes some statistical analyses.				
7	Asking the teacher how to use a probability table.				
8	Trying to understand a mathematical demonstration.				
9	Doing the final examination in a statistics course.				
10	Reading an advertisement for an automobile which includes figures on gas mileage, compliance with population regulations, etc.				
11	Walking into the classroom to take a statistics test				
12	Asking the teacher about how to do an exercise in statitics.				
13	Getting to the day before an exam without having had time to revise the syllabus.				
14	Waking up in the morning on the day of a statistics test.				
15	Realizing, just before you go into the exam, that I have not prepared a particular exercise.				
16	Copying a mathematical demonstration from the blackboard while the teacher is explaining it.				
17	Asking one of my teachers for help in understanding a printout in statistics is difficult.				
18	Seeing a classmate carefully studying the results table of a problem he has solved.				
19	Going to a statistics exam without having had enough time to revise.				
20	Asking a teacher for help when trying to interpret a results table.				
21	Trying to understand the statistical analyses described in the abstract of a journal article.				
22	Going to the teacher’s office to ask questions.				
23	Asking a private teacher to tell me how to do an exercise.				
24	I wonder why I have to do all these things in statistics when in actual life I'll never use them.				
25	Statistics is worthless to me since it's empirical and my area of specialization is abstract.				
26	Statistics takes more time than it's worth				
27	Statisticians are more number oriented than they are people oriented.				
28	I can't tell you why, but I just don't like statistics.				
29	Statistics teachers talk so fast you cannot logically follow them.				
30	Statistical figures are not fit for human consumption				

On the other hand, the following are the discarded items

Table 5: Discarded items after PCA

S/N	ITEMS	SA	A	D	SD
1	Interpreting the meaning of a table in a journal article.				
2	Statistics takes more time than it's worth				
3	Statistics is worthless to me since it's empirical and my area of specialization is abstract				
4	I wonder why I have to do all these things in statistics when in actual life I'll never use them				
5	Realizing the day before an exam that I cannot do some problems that I thought were going to be easy				
6	Since I've never enjoyed maths I don't see how I can enjoy statistics				
7	Statistics takes more time than it's worth				
8	Walking into the classroom to take a statistics test				
1	I can't tell you why, but I just don't like statistics				

RESULTS

Research Question 1: What are the items of Statistical Anxiety Scale (SAS)?

Table 6: shows the items of Statistical Anxiety Scale (SAS)

S/N	ITEMS (I feel anxious When)	SA	A	D	SD
1	Studying for an examination in a statistics course.				
2	Interpreting the meaning of a table in a journal article.				
3	Going to ask my statistics teacher for individual help with material I am having difficulty understanding.				
4	Realizing the day before an exam that I cannot do some problems that I thought were going to be easy.				
5	Asking a private teacher to explain a topic that I have not understood at all.				
6	Reading a journal article that includes some statistical analyses.				
7	Asking the teacher how to use a probability table.				
8	Trying to understand a mathematical demonstration.				
9	Doing the final examination in a statistics course.				
10	Reading an advertisement for an automobile which includes figures on gas mileage, compliance with population regulations, etc.				
11	Walking into the classroom to take a statistics test				
12	Asking the teacher about how to do an exercise in statistics.				
13	Getting to the day before an exam without having had time to revise the syllabus.				
14	Waking up in the morning on the day of a statistics test.				
15	Realizing, just before you go into the exam, that I have not prepared a particular exercise.				
16	Copying a mathematical demonstration from the blackboard while the teacher is explaining it.				
17	Asking one of my teachers for help in understanding a printout in statistics is difficult.				
18	Seeing a classmate carefully studying the results table of a problem he has solved.				
19	Going to a statistics exam without having had enough time to revise.				
20	Asking a teacher for help when trying to interpret a results table.				
21	Trying to understand the statistical analyses described in the abstract of a journal article.				
22	Going to the teacher’s office to ask questions.				

Table 6 Continued

23	Asking a private teacher to tell me how to do an exercise.				
24	I wonder why I have to do all these things in statistics when in actual life I'll never use them.				
25	Statistics is worthless to me since it's empirical and my area of specialization is abstract.				
26	Statistics takes more time than it's worth				
27	Statisticians are more number oriented than they are people oriented.				
28	I can't tell you why, but I just don't like statistics.				
29	Statistics teachers talk so fast you cannot logically follow them.				
30	Statistical figures are not fit for human consumption				

Research Question 2: What are the final items of Statistical Anxiety Scale (SAS) using Principal Component Analysis (PCA)?

Table 7: shows Final Items of Statistical Anxiety Scale (SAS) Via Principal Component Analysis

Rotated Component Matrix

	Component	
	1	2
VAR00029	**.574**	
VAR00027	**.565**	
VAR00017	**.546**	
VAR00028	**.537**	
VAR00009	**.514**	
VAR00010	**.508**	
VAR00025	**.502**	
VAR00022	**.502**	
VAR00016	**.491**	
VAR00024	**.480**	
VAR00026	**.480**	
VAR00030	**.473**	
VAR00023	**.467**	
VAR00003	**.464**	
VAR00002	**.441**	
VAR00018	**.439**	
VAR00011	**.437**	
VAR00019	**.421**	
VAR00015		
VAR00020		
VAR00021		
VAR00013		
VAR00014		
VAR00007		**.655**
VAR00005		**.653**
VAR00006		**.646**
VAR00008		**.557**
VAR00004		**.502**
VAR00001		
VAR00012		

The table above has shown that items 29, 27, 17, 28, 9, 10, 25, 22, 16, 24, 26, 30, 23, 3, 2, 18, 11, 19, 7, 5, 6, 8 and 4 loaded well in both factors while items 15, 20, 21, 13, 14, 1 and 12 did not load in the

scale. This means that in all 7 items were rejected 23 were accepted based on the criterion of 0.40 (Kaiser, 1974).

Table 8: Final items of SAS is revealed below;

S/N	ITEMS (I feel anxious When)	SA	A	D	SD
1	Interpreting the meaning of a table in a journal article.				
2	Going to ask my statistics teacher for individual help with material I am having difficulty understanding.				
3	Realizing the day before an exam that I cannot do some problems that I thought were going to be easy.				
4	Asking a private teacher to explain a topic that I have not understood at all.				
5	Reading a journal article that includes some statistical analyses.				
6	Asking the teacher how to use a probability table.				
7	Trying to understand a mathematical demonstration.				
8	Doing the final examination in a statistics course.				
9	Reading an advertisement for an automobile which includes figures on gas mileage, compliance with population regulations, etc.				
10	Walking into the classroom to take a statistics test				
11	Copying a mathematical demonstration from the blackboard while the teacher is explaining it.				
12	Asking one of my teachers for help in understanding a printout in statistics is difficult.				
13	Seeing a classmate carefully studying the results table of a problem he has solved.				
14	Going to a statistics exam without having had enough time to revise.				
15	Going to the teacher's office to ask questions.				
16	Asking a private teacher to tell me how to do an exercise.				
17	I wonder why I have to do all these things in statistics when in actual life I'll never use them.				
18	Statistics is worthless to me since it's empirical and my area of specialization is abstract.				
19	Statistics takes more time than it's worth				
20	Statisticians are more number oriented than they are people oriented.				
21	I can't tell you why, but I just don't like statistics.				
22	Statistics teachers talk so fast you cannot logically follow them.				
23	Statistical figures are not fit for human consumption				

Table 9: Below are the items discarded.

15	Trying to understand the statistical analyses described in the abstract of a journal article				
20	Asking a teacher for help when trying to interpret a results table.				
21	Realizing, just before you go into the exam, that I have not prepared a particular exercise				
13	Studying for an examination in a statistics course				
14	Asking the teacher about how to do an exercise in statistics.				
1	Getting to the day before an exam without having had time to revise the syllabus.				
12	Waking up in the morning on the day of a statistics test.				

Research Question 3: What are the relationship between Statistical Anxiety Scale (SAS) using:

a. Convergent Validation Process.

The Statistical Anxiety Scale" (SAS) was administered alongside a similar instrument which was named Statistical Anxiety Rating Scale (STARS) by Vigil-Colet, Lorenzo-Seva and Condon (2008).

Table 10: PPMC Correlation analysis between SAS and STARS

Variable	N	r	Df	α	Remark
SAS	200	0.71	198	0.05	Significant Reject Ho
STARS	200				

The table shows that correlation evidence reveals an r-value of 0.71 at an alpha value of 0.05. From the r-value, it is clear that SAS correlate strongly and highly with a similar instrument, hence, the instrument is adjudged to be construct valid.

Research Question 4: What are the reliability of Statistical Anxiety Scale (SAS) using:

a. Cronbach Alpha method of reliability.

Table 11: Cronbach Coefficient Alpha Reliability table

Cronbach r	N of Items	Valid N	Excluded		Remark
0.76	23	200	0	200	Very High Coefficient

From table 11 above, it is seen that Cronbach Alpha correlation coefficient was .76. This value was remarked to be very high and according to Kpolovie (2010) is good enough to guarantee the reliability of the overall SAS.

b. Split-half method of reliability.

Table 12: Split-Half Coefficient Reliability

N of items (Half)	N of Items	Rht.	Rft.	Remarks
11	22	.38	.55	High Reliability

The table revealed that split-half reliability of the half test (rht) was .38. However, when spearman brown prophecy formula was used to substitute the half test with a formula of

$$\text{rft} = \frac{2x\ rht}{1 + rht}$$ (Opara, 2016).

A Guttman reliability test (rft) was gotten as .55. Again Kpolovie (2010) also indicated that this was a high coefficient that guarantees the reliability of SAS for use by undergraduates.

Research Question 5: What are the standard scores of Statistical Anxiety Scale (SAS) using:

a. Z-score

b. T-score

Table 13: shows z and t-score tables of the responses
Note: Raw Score Mean=42.06, St.D=13.74
Where; **Z-score** = z-mean/St.D, **T-score** = 10z+50

Raw Scores	Frequency	Z-scores	T-scores (To whole Numbers)
70	1	2.032	**70**
64	3	1.596	**66**
63	4	1.523	**65**
62	8	1.450	**65**
60	7	1.305	**63**
59	4	1.232	**62**
58	21	1.159	**62**
57	15	1.086	**61**
56	10	1.014	**60**
55	12	0.941	**59**
54	14	0.868	**59**
53	20	0.795	**58**
52	24	0.723	**57**
51	24	0.650	**56**
50	19	0.577	**56**
23	14	-1.387	**36**
	N Total= 200		

SUMMARY OF FINDINGS

1. The present study has developed, validated and standardized Statistics Anxiety Scale (SAS) for undergraduates. Put succinctly, it is satisfactorily and of no doubt that the SAS has twenty three (23) components which formed the final component the instrument;
2. Convergent Validity index indicating construct validity was 0.71.
3. Cronbach Reliability was 0.76 while Split-half was 0.55 all indicating high reliabilities.
4. Standard Scores of SAS for use in tertiary institution are as shown in table 13.

DISCUSSION OF FINDINGS

The current study dealt with developing and standardizing statistical anxiety scale for undergraduate students in tertiary institutions in Rivers State. The scale was developed with an initial item of 39. During the pilot stage, the items of the instrument were reduced to the 30 and finally during analysis, the items were further reduced to 23 items. The final 23 items however measured student anxiety level in a general sense and this could be used and applied to undergraduate students at all levels of tertiary institutions. Validation of the instrument was done validation via convergent process revealed an index of 0.71. This index as noted by Kpolovie (2020) indicated that the instrument is valid. These validity index also correspond with the index as reported by Vigil-Colet, Lorenzo-Seva and Condon (2008) when they developed the statistical anxiety rating scale showing that the current instrument develop here in Nigeria is equally valid.

The analysis also revealed in terms of reliability of the instrument a Cronbach reliability index of 0.76. This index indicated a high reliability coefficient signifying that the instrument has high reliability. Similarly, a split half reliability of 0.55 was realized also indicating a high reliability. It therefore means that the current instrument so developed and validated is highly reliable and could be applicable to every tertiary institution in Rivers State and in general. Finally, the scores of the student were converted to standard scores as shown in table 14. For instance, a student with a raw score of 54 have a standard score of 56 while those with a score of 15 or 14 will have a standard score of 36. The standardization indicates

that the instrument could be used to evaluate students within the same characteristics and could be used to predict their level of statistical anxiety.

RECOMMENDATIONS

Based on the findings of this study, it is recommended that;

1. The statistical anxiety scale so develop should be able to applied undergraduate students to measure their level of anxiety towards statistics as this will help the lecturers to develop new ways and methods that will help them learn better.
2. It is also recommended that other statisticians should use the instrument in evaluating students especially in tertiary institutions across Rivers State.
3. It is also recommended that whenever similar instruments will ever be developed, the developer should use the current instrument, or as compared in the current study with STARS by Vigil-Colet, Larenzo-Seva and Tondon in determining convergent and divergent validity of such similar instrument.
4. Finally, scores of students in the statistical anxiety scale should be used to predict scores of student in related areas.

CONCLUSION

Statistics as a core course in education should be encouraged by students in faculty of education in order to reduced their fear, based on this, developing a well-structured instrument that can diagnosed those problem is inevitable and as such standardizing the instrument in determining the psychometric properties which gave high reliability index when compared to similar instruments.

References

Awujo, C.G, Ugwu C.J & Amadi G.N (2019). *Psychology of phobic disorder for psychotherapists and counselors*. University of Port Harcourt Press.

Chadha, N. K. (2009). *Applied Psychometry*. New Delhi, IN: Sage Publications. Retrieved from https://www.scirp.org/pdf/PSYCH_2018102216111334.pdf.

Dimitrov, D. M. (2012). *Statistical methods for validation of assessment scale data in counseling and related fields. Alexandria, VA: American Counseling Association.*

Furr, R. M. (2011). *Scale Construction and Psychometrics for Social and Personality Psy-chology*. New Delhi, IN: Sage Publications. https://doi.org/10.4135/9781446287866.

Haladyna, T. M. (2014). *Developing and Validating Multiple-Choice Test Items*. Mah-wah, NJ: Erlbaum.

Hambleton R, Swaminathan H & Rogers J. (2019). *Ability and item parameter estimation. In: Foster D, editor. Fundamentals of item response theory*. California: Sage Publications. pp. 32-52.

Irwing, P., & Hughes, D. J. (2018). Test Development. In P. Irwing, T. Booth, & D. J. Hughes (Eds.), *The Wiley Handbook of Psychometric Testing: A Multidisciplinary Reference on Survey, Scale and Test Development.* Hoboken, NJ: Wiley.

Kpolovie, P.J (2010) *Advance research method*. Owerri: Spring field publishers.

Krosnick, J. A., & Presser, A. (2010). Question and questionnaire design. In P. V. Marsden, & J. D. Wright (Eds.), *Handbook of Survey Research* (2nd ed., pp. 264-313). Bingley, UK: Emerald.

Macher D., Paechter M., Papousek I., Ruggeri K., Freudenthaler H. H., Arendasy M. (2013). Statistics anxiety, state anxiety during an examination, and academic achievement. *Br. J. Educ. Psychol*. 83, 535–549.

Macher, D., Papousek, I. Ruggeri, K & Paechter, M. (2015). Statistics anxiety and performance: blessings in disguise. *Front Psychol.* 2015; 6: 1116.

Nwankwo, O.C (2015). *Abnormal psychology*. (The clinical approach) Port-Harcourt Pam Unique Publishing Coy, Ltd.

Obilor, E.I. (2021). *Principles of test construction*: Sabcos printers and publishers.

Obilor, E.I., & Ugada, C. (2023). Problems in measurement and evaluation by lecturers in Ignatius Ajuru University of Education: Implication for students assessment. *Faculty of Education Conference Journal*, 2(1).

Opara, I. M. (2016). *Test Construction and Measurement Concepts and Applications*. Owerri. Career Publisher.

Orluwene, G.W. (2012). *Fundamentals of testing and non-testing tools in educational psychology*, printed by: Sunny Alade printed product in Port Harcourt.

Price, L. R. (2017). *Psychometric Methods: Theory into Practice*. New York: The Guilford Press.

Ramalingham, P (2016). *Dictionary of Psychology*. Academic Publishers. New Delhi (India).

Raykov, T. (2012). *Scale construction and development using structural equation modeling.* In Hoyle, R. (Ed.), Handbook of structural equation modeling (pp. 472-492). New York, NY: Guilford.

Ugada, C. & Ubulom, W.J. (2024). Development and standardization of instrument to measure students attitude towards statistics. *Rivers state university faculty of Education Conference Journal*, 3(2), 58-79.

Ugada, C. (2024). Analyzing the Dimensions and Difficulty Level of the 2022 West African Examination Certificate Chemistry Objective Test. Faculty of Natural and Applied Sciences, Ignatius Ajuru University of Education Port Harcourt, Nigeria, *FNAS 5th Annual Conference Proceedings:* February 2024; Page No. 133-139.