DEVELOPMENT AND FACTORIAL VALIDATION OF TEACHING PRACTICE SCALE FOR ASSESSMENT OF PRE-SERVICE INTEGRATED SCIENCE TEACHERS IN NORTH CENTRAL NIGERIA UNIVERSITIES

[1]Jackson Ochigbudu Ode, [2]John Abah and [3]Msuur Tofi

[1 &3] Department of Science and Mathematics Education
Benue State University, Makurdi
[2] Department of Basic Sciences, School of General Studies,
Alvan Ikoku Federal University of Education, Owerri
odejackson90@gmail.com, jonabah@gmail.com and tmsuur@gmail.com

Abstract

This study developed and validated a teaching practice scale for assessment of pre-service integrated science teachers in North Central, Nigeria universities using instrumentation design. Four universities within North Central Nigeria where integrated science is offered at undergraduate level were purposively sampled for the study. The population of the study was 126 pre-service integrated science teachers and the entire population participated in the study. Four research questions and two hypotheses were raised to guide the study and the hypotheses were tested at 0.05 level of significance. An initial item pool of 130 items on nine major teaching skills identified from literature were presented to three experts in Science Education and two experts in Measurement and Evaluation for face validation. Eighty-three items in six sub-scales survived the face validity. The eighty-three items with four-point scale were subjected to factor analysis to select items that were construct valid and suitable to be included in the final instrument used in the study. The result of the factorial validation showed that 57 items had factor loadings of 0.50 and above on a single factor and were therefore selected. The internal consistency and inter-rater reliability coefficients of the instrument yielded 0.94 and 0.62 through Cronbach alpha and Pearson product moment correlation methods respectively. It was recommended among others that the instrument should be adopted by teacher training institutions for assessing pre-service teachers during teaching practice exercise as the instrument valid and reliable.

Key Words: Instrument development, validation, teaching practice, pre-service teachers and Integrated Science

Introduction

Education is considered in any society as a vital tool for development and advancement in all spheres of human endeavours. As an enterprise, it involves different agents and the teacher occupies a key position. Quality teacher preparation is therefore given a primary attention by nations since it is the bedrock for development and progress in all areas of human life. Quality teachers are essential resources in any educational institution. The integrated science teacher is a professional agent that is responsible for laying a sound foundation for students' successful learning in the field of science. For instance, apart from teaching and supporting students to learn, integrated science teachers introduce the students to culture of science through active engagement in meaningful learning experiences; the learners may therefore suffer or gain from integrated science teachers' teaching quality (Boudersa, 2016). In other to build and develop teaching skills in teachers, teacher training institutions do engage their students in teaching practice for qualitative training of pre-service teachers including that of integrated science.

As a key component of teacher training and becoming a teacher, teaching practice serves as a bridge between theory and practice. Teaching practice is an exercise that provides opportunity for improving the quality of pre-service teachers as they are engaged in the art of teaching before becoming professional teachers (Audet, 2014). It is the crux of pre-service teachers' preparation as it is the nexus between studentship and becoming a professional teacher. As a corollary, teaching practice exercise is a period of varied excitement, expectations and anxiety among pre-service teachers as they are practically exposed to complex mixture of realities involved in becoming a teacher. Through teaching practice, pre-service teachers are expected to perform the responsibilities of a teacher in order to practice in the classroom what they have learnt theoretically in their course programmes. Some of these responsibilities include but not limited to teaching, marking students work books and registers, conducting practical classes, and extra-moral activities, attending staff meetings, setting and administering tests, examinations including giving of feedback and results to students among others. Pre-service teachers' teaching practice is therefore fundamental in teacher preparation and it should be properly organized and thoroughly assessed using valid and reliable instruments for aspiring teachers to meet the demands of teaching.

Assessment is a vital component of the school system. To Martin and Itter (2014), assessment plays a great role to teachers and students by providing them with feedback for them to reflect on their achievement and difficulties particularly when valid instruments are used. Regrettably, apart from the paucity of valid instruments for assessment of pre-service teachers' teaching practice in integrated science (Mehta & Doctor 2013), the few available ones also have limitations. The limitations as discovered from literature (Mackenzie, Podsakoff & Podsakoff, 2011; Prados, 2011) include inadequate definition of construct domain, incorrect specification of measurement model and underutilization of techniques for establishing construct validity; others are weak psychometric properties, improper item editing, while others require extensive time to complete the responses. These limitations are still common with most measurement scales today due to paucity of valid and reliable instruments particularly as it relates to assessment of pre-service teachers' teaching practice. The implication of that according to Morgado, Meireles, Neves, Amaral and Ferreira (2017) is that, it weakens the psychometric quality of the obtained results thereby limiting the applicability and generalizability of the developed scales. The authors further submitted that poorly defined construct equally threatens the understanding of the phenomena as such future studies cannot be built on them. Consequently, apart from addressing the issue of paucity, it becomes necessary too to develop a valid and reliable Pre-service Integrated Science Teachers' Teaching Practice Assessment Scale (PISTTPAS) with enhanced psychometric quality of items in the scale.

Scale development is a critical and fundamental activity that helps researchers in knowledge acquisition about people, objects, events and processes (Ode & Ezeh, 2019). Assessment scales are therefore useful tools for attributing scores in some numerical dimensions to phenomena that cannot be directly measured; they consist of items revealing level of theoretical variables that are latent. Authors (DeVellis, 2012; Pasquali, 2010) maintained that scale development involves a complex and systematic procedures that requires rigorous theoretical and methodological processes. According to these authors, scale development takes three basic steps: (i) item generation where the theoretical support for the initial item pool is provided (Hutz, Bandeira & Trentini, 2015), (ii) theoretical analysis, where the content validity of the new scale is assessed by ensuring that the initial item pool reflects the desired construct (Arias, Lioreda & Lioreda, 2014) and (iii) Psychometric analysis, where the construct validity and reliability of the developed instrument is assessed.

In instrument development, three criteria have been identified for use in appraising the quality of a scale or test in education research (Ode & Ezeh, 2019). These include objectivity (extent to which the scale or test is free from biases of the researcher and the respondent, reliability (consistency of the instrument) and validity (extent to which the instrument measure what is intended to measure). Construct validity deals with the degree to which scores of a measure permits inferences on the underlying trait; in order words, it is about what the instrument is measuring and it is usually assessed using factor analysis. Factor analysis is a technique used in extracting pure items from impure and complex items in a factor matrix Content validity assesses the adequacy of the proposed items from the item pool; it also concerns the adequacy of sampling the domain of the construct.

Studies on development and validation of instrument for measuring pre-service teachers' teaching practice are relatively scarce; few are available in other science subjects like biology, chemistry, physics and mathematics. Besides, most of the studies were conducted on in-service teachers. These calls for development and validation of pre-service integrated science teachers' teaching practice assessment scale.

Moreover, Audet (2014) submitted that teachers' performance directly affects students learning. The author further maintained that teachers' performance in the classroom is not unconnected to the quality of assessment of their teaching practice during training. Consequently, pre-service teachers' performance could be linked to the quality of assessment of the pre-service teachers during training. However, based on evidence available to the researcher, there is dearth of valid and reliable instruments for assessing pre-service integrated science teachers' teaching competency. It became therefore necessary to develop a valid and reliable teaching practice scale for assessing pre-service integrated science teachers. Hence, against the backdrop of paucity of valid instruments for assessment of pre-service integrated science teachers' teaching practice, this study focused on development and validation of pre-service integrated science teachers' teaching practice assessment scale.

Statement of Problem

Assessment of teaching and learning situation is very essential in education. This is because, it is a means by which information on level of knowledge, learning progress and performance difficulties in a task can be uncovered. However, such information can only be dependable and truthful when valid instrument is used. Unfortunately, few of the instruments available for assessment of pre-service integrated science teacher teaching practice have poor psychometric properties which leads to pre-service teachers' poor assessment during training. This negatively affects the teachers' competence in teaching upon graduation which in turn affects students' performance. This ugly situation may remain if the solution is not sought by developing a valid and reliable scale for assessment of pre-service teachers teaching practice. Moreover, most of the existing valid and reliable instrument for assessing teaching practice from literature did assess one or two facets of teaching skills. This suggests that robust instruments that will comprehensively assess pre-service teachers' teaching skills are lacking. Against this backdrop, this study developed and validated a pre-service teachers' teaching practice scale for assessment of pre-service integrated science teachers.

Purpose of the Study

The general purpose of this study was to develop and validate a pre-service teachers' teaching practice assessment scale. Specifically, the study:

1. Developed a Pre-service Integrated Science Teachers' Teaching practice Assessment Scale (PISTTPAS);
2. Determined the construct validity of the PISTTPAS;
3. Established the internal consistency reliability of PISTTPAS;
4. Ascertained the inter-rater reliability of PISTTPAS

Research Questions

The following research questions were posed for the study

1 .What are the sub-scales that makes up the PISTTPAS?

2. What is the construct validity of the instrument (PISTTPAS)?
3. What is the internal consistency reliability of PISTTPAS?
4. What is the inter-rater reliability of PISTTPAS?

Method

The design used for this study was instrumentation. Ezeh (2015) remarked that instrumentation is a study design that aids researchers develop and validate instruments needed for efficacious execution of specific task in education. Therefore, since the major thrust of this study was to develop and validate a new instrument for assessing pre-service integrated science teachers' teaching practice, instrumentation is considered appropriate in this study. The population of this study consists of 126 final year pre-service integrated science teachers (2018/2019 session) in the universities within North

Central Nigeria, where integrated science programme is being offered. This information was obtained through pre-survey visitation to Science Education Departments of the institutions by the researcher. The entire 126 final year pre-service integrated science teachers were used for the study. Purposive sampling was used to select four of the universities that offers integrated science programme out of the 22 universities within the study area.

The instrument developed by the researcher and used for data collection in this study was Pre-service Integrated Science Teachers Teaching Practice Assessment Scale (PISTTPAS). The instrument consists of two sections (A and B). Section A evokes personal information of pre-service integrated science teachers such as name of school, name of student, sex and location of pre-service teachers. Section B consists of statements that measure pre-service teachers teaching practice pertaining lesson planning, lesson presentation, classroom management, communication, ICT usage and the use of instructional materials. Others include learning reinforcement, lesson evaluation and professionalism; it consists of nine clusters, with statements in each cluster relating to a construct. The PISTTPAS is a four-point rating scale to minimize response burden and it comprised 130 items grouped into 9 sub-scales but was reduced to 83 items in six sub-scales after face validation. The six teaching skills on which pre-service integrated science teachers were assessed include – lesson planning, lesson presentation, classroom management, communication, use of instructional materials and lesson evaluation. The raters (lecturers of pre-service integrated science) rated the pre-service teachers using the instrument (scale) with the following points: Very Good (VG) - 4 points, Good (G) -3 points, Fair (F) - 2 points and Poor (P) -1 point.

The development of the scale for this study was based on suggestion by Slavec and Drnovsek (2012) who maintained that the first step in scale development is to bring to fore the theoretical importance and existence of the construct. Therefore, the first step in developing PISTTPAS was to identify and select the teaching skills that pre-service integrated science students are expected to demonstrate in teaching. According to Ode and Iloakasia (2020), the major teaching skills teachers should be able to showcase in teaching include lesson planning, lesson presentation/implementation, classroom management, use of ICTs and instructional materials, learning reinforcement, lesson evaluation, communication and professionalism. Based on the submission of this authority, interview of experts in integrated science teaching at both university and basic education level including other relevant literature in teaching competencies of teachers, the researcher identified and selected six teaching skills.

This was followed by writing or generating the items to get the required information from the respondents under each domain of teaching. The initial item pool in other of priority for pre-service teachers include 18 items under lesson planning component, 20 under lesson presentation, 18 on classroom management, 14 under communication and 16 under incorporation of ICTs. Others include11 items on instructional material component, 12 under learning reinforcement, 10 on lesson evaluation and 11 on professionalism. This brought the total number of initial items to 130. Further to generation of the items, four point scale was attached to the instrument which serves as a guide for the raters in rating the teaching skills of pre-service integrated science teachers.

Face validity: The instrument with 130 items was given to five experts comprising two from measurement and evaluation and three pre-service integrated science lecturers for face validation. Two of the measurement and evaluation experts were from University of Nigeria, Nsukka and the remaining three experts in integrated science education were from Benue State University, Makurdi, The validators were asked to appraise the instrument as it relates to whether the statements can be observed and rated, the appropriateness of the items on each of the domains of teaching captured in the instrument, ambiguity, clarity, adequacy and relevance of the items in the scale. The summary of the suggestions and comments from the validators were that: (i) items under incorporating ICTs should be expunged since most classrooms in our clime are not ICT ready (ii) the length of the instrument should be reduced for it not to be carelessly filled (iii) though some items were relevant under their domains but were unobservable during classroom teaching and so should be removed. Based on these validators' comments and suggestions, items that required minor corrections were corrected, items under ICT teaching practice and all the unobservable items were removed, all items under learning reinforcement were put under lesson presentation while items under professionalism were also discarded since pre-service teachers are still learning how to teach and are yet to be professionals. Arising from these corrections, 47 items were discarded, 8 items were modified. A total of 83 items that survived face validity were subjected to factor analysis to ascertain their construct validity as recommended by Plake and Parker (1982).

Construct validation: Construct validation was carried out through factor analysis. According to Ezeh (2011), construct validity refers to extent to which an instrument measures a theoretical construct such as anxiety, verbal fluency, stress, neuroticism and speed of working among others. To do this, the PISTTPAS was subjected to the process recommended by Plake and Parker (1982). This resulted in a factor matrix (6 components). The principal – axis method with Varimax rotation option was employed using Statistical Package for Social Sciences (SPSS -version 20). Varimax rotation maximizes the variance of factor loadings (columns) on all the variables (rows) in a factor matrix and it differentiates the original variable from extracted factors. From the factor analysis, the minimum number of factors that best explained the structure of the instrument including factor loadings of each item in the scale emerged on each factor. Plake and Parker recommended that factor loading of 0.50 should be considered as minimum loading for an item. As a corollary, items with factor loadings of 0.50 and above on a single factor are factorially pure and were retained while those with factor loadings less than 0.50 are factorially impure items; items with factor loadings of up to 0.50 on more than one factor are factorially complex and were discarded. On the basis of these criteria, 20 items and 6 items were factorially impure and complex respectively and were discarded; 57 items deemed factorially pure were retained and renumbered serially for use in the study.

The instrument was trial tested on 15 pre-service integrated science teachers in Benue State University, Makurdi. During trial testing, pre-service integrated science teachers' lecturers used the instrument to observe and rate the pre-service teachers while they taught. The scores obtained by the pre-service integrated science teachers were used to establish the internal consistency and inter-rater reliability coefficients of the instrument. The internal consistency reliability of the instrument was ascertained using Cronbach's alpha and the reliability coefficient of 0.935 was obtained. The choice of Cronbach's alpha technique is due to the fact that the items are polytomously scored. Inter-rater reliability coefficient was determined using Pearson product moment correlation method. The inter-rater reliability is applicable when different raters or scorers are independently engaged in rating individual. This further determines the degree of agreement of the scores by the raters. Since the pre-service integrated science teachers were assessed by two different raters, the scores of the different raters were correlated to establish inter-rater reliability of the instrument and the result yielded 0.616.

The lecturers of pre-service integrated science teachers in the sampled universities were visited and briefed by the researcher on observing and rating the pre-service integrated science teachers using the rating scale during teaching practice. During data collection, the researcher equally visited and frequently called the raters to ensure that the pre-service teachers were actually rated and to ascertain the extent of the progress made in the exercise. A total of 126 pre-service integrated science teachers were rated by 37 supervisors across the sampled universities using the rating scale. The number of pre-service integrated science teachers observed and rated by each lecturers ranged from 1 to 5 in primary and secondary schools. During the teaching practice exercise, the respective supervisors of the pre-service integrated science teachers observed the lesson plans of the pre-service teachers to ascertain their lesson preparation skills. Also, while teaching the students lesson presentation skills were assessed and rated. Other teaching skills of pre-service integrated science teachers that were assessed and rated were classroom management, communication, use of instructional materials and lesson evaluation. The exercise was carried out by the supervisors of the pre-service teachers in all the respective schools where they were posted. The copies of the instrument used in rating the pre-service teachers were collected by the researcher from the supervisors after which the scores obtained were collated for analysis. The data collected were analyzed based on each research question raised for the study. Research question two was answered using factor analysis, research question three was answered using Cronbach's Alpha while Pearson product moment correlation coefficient was used to answer research question four.

Results

Research Question One

What are the sub-scales that makes up PISTTPAS?

Table 1: Summary of Subscales of PISTTPAS

S/NO	Teaching practice Domain	Number Of Items
A.	Lesson Planning	13
B.	Lesson Presentation	12
C.	Classroom Management	8
D.	Communication	12
E.	Instructional Materials	6
F.	Lesson evaluation	6
	Total	57

Table 1 shows the sub-scales that make up PISTTPAS after factor analysis. From the table, a total of 57 items in the instrument were distributed in six subscales that makes up PISTTPAS. These subscales include lesson planning, lesson presentation, classroom management, communication, instructional materials and lesson evaluation. Sub-scale A (lesson planning) has 13 items while subscale B (lesson presentation) has 12 items. Sub-scale C which addresses classroom management has eight items while items under sub-scale D that pertains communication skills are 12. Subscale E which focuses on instructional materials has six items and sub-scale F which is on lesson evaluation comprised of six items. These respective number of items in the six subscales brings the total number of items in PISTTPAS to 57.

Research Question Two

What is the construct validity of the instrument PISTTPAS?

Table 2: Summary of Factor Analysis of PISTTPAS Items and the Factor Loadings.

Factor	Pure Items	Number	Impure Items	Number	Complex Items	Number	Total
A	1,3,5,6,9,10,12,13, 14,15,16,17,18	13	2,4,7,8,11,	5	-	0	18
B	19,20,22,23,25,26, 27,29,34,37,38,41	12	21,24,28,31,35,40	6	30,32,33,36, 39,42	6	24
C	44,45,46,47,48,50, 52,53	8	43,51,49	3	_	_	11
D	54,55,56,57,59,60, 61,62,63,64,65,66	12	58	1	-	0	13
E	67,68,69,71,73,74	6	70,72,75	3	-	0	9
F	77,79,80,81,82,83	6	76,78	2	-	0	8
Total		57		20		6	83

It is imperative to note here that:

i. Factor loading range used in selecting factorially pure items is 0.50 and above (Plake & Parker, 1982).
ii. Items with less than 0.50 factor loading on any of the factors were considered factorially impure and not selected.
iii. Any item with factor loading of 0.50 on more than one factor is considered factorially complex and therefore not selected.

Table two shows the summary of factor analysis of PISTTPAS with factor matrix as shown in Appendix D. Six factors were extracted from the analysis with each factor having items loaded on it. The six extracted factors from the factor matrix include lesson planning, lesson presentation, classroom management, communication, instructional materials and lesson evaluation. From the summary on table 2, 18 items were clustered in factor 1(lesson planning) out of which 13 have factor loadings of 0.50 and above ranging from 0.542 to 0.883 on a single factor and thus were selected. Items 2,4,7,8 and 11 have factor loadings less than 0.50 on all the factors and thus were deemed factorially impure and not selected. Hence, 5 items were not selected under factor one.

Factor 2 which is on lesson presentation has a total of 24 items clustered under it. Out of this number, 12 items (19, 20, 22, 23, 25, 26, 27, 29, 34, 37, 38 and 41) have factor loading ranging from 0.550 to 0.830 on single factor and were selected; conversely, items 21, 24, 28, 31, 35 and 40 were deemed factorially impure as they have factor loading of less than 0.50 on any of the factors. Additionally, items 30, 32, 33, 36 39 and 42 have factor loadings above 0.50 on more than one factor and thus were factorially complex and not selected. This brought the total number of unselected items to 12 under lesson presentation subscale.

Factor 3 is classroom management and it has 11 items clustered under it. Of these 11 items, eight (44, 45, 46, 47, 48, 50, 52 and 53) have factor loading ranging from 0.516 to 0.748 on single factor and were selected. However, three items (43, 49 and 51) were factorially impure and thus were not selected. Factor 4 (communication) has 13 items clustered under it. Under factor 4, 12 items (54, 55, 56, 57, 59, 60, 61, 62, 63, 64, 65 and 66 were factorially pure and were selected since their factor loadings range from 0.557 to 0.904 on single factor. Item 58 has less than 0.05 factor loading on any of the factors and thus was not selected.

Factor 5 is instructional materials. Nine items are in this cluster of which three items (70, 72 and 75) have less than 0.05 factor loadings on any of the factors and thus were factorially impure and were not selected. However, the other six items (67, 68, 69, 71, 73 and 74 have factor loadings of 0.50 and above (0.513 to 0.893) and therefore were selected. Factor 6 which is on lesson evaluation has eight items clustered under it of which six items (77, 79, 80, 81, 82 and 83) have factor loadings of 0.50 and above with the range 0.521 to 0.807 on single factor and thus were selected as pure items. Conversely, two items (76 and 78) were impure items as their factor loading were less than 0.50 across the factors and thus were not selected.

In summary, 57 items of PISTTPAS were selected for having factor loadings of 0.50 and above on single factor; 19 items were factorially impure while 6 items were factorially complex. This brings the total number of discarded items to 25. This implies that the 57 items selected are valid, therefore, have construct validity.

Research Question Three

What is the internal consistency reliability coefficient of PISTTPAS?

Table 3: Summary of Internal Consistency Reliability Coefficient of PISTTPAS Items

S/N	Factor/Cluster	Items	Internal Consistency Reliability
1	Lesson planning	1,3,5,6,9,10,12,14,15,16,17,18	0.797
2	Lesson presentation	19,20,22,23,25,26,27,29,34,37,38,41	0.742
3	Classroom management	44,45,46,47,48,50,52,53	0.839
4	Communication	54,55,56,57,59,60,61,62,63,64,65,66	0.873
5	Instructional materials	67,68,69,71,73,74	0.798
6	Lesson evaluation	77,79,80,81,82,83	0.829
	Overall		0.935

Table 3 above shows the internal consistency reliability coefficients of PISTTPAS items using Cronbach's alpha. It is seen that lesson planning has internal consistency reliability coefficient of 0.797; lesson presentation items have consistency reliability of 0.742; classroom management items have internal consistency reliability of 0.839 while communication items have internal consistency reliability of 0.873. The internal consistency reliability coefficient of instructional materials items is 0.798 and that of lesson evaluation factor is 0.829. The overall internal consistency reliability of the PISTPPAS items is 0.935.

Research Question Four

What is the inter-rater reliability coefficient PISTTPAS

Table 4: Summary of Inter-Rater Reliability Coefficient of PISTTPAS

Factors/clusters	Number of items	Inter-rater reliability coefficient
6	57	0.616

Table 4 reveals the summary of inter-rater reliability coefficient of PISTTPAS items ascertained using Pearson product moment correlation. This was done by correlating the scores of the two raters that rated the pre-service integrated science teachers. It is seen that the inter-rater reliability coefficient of the items is 0.616.

Discussion of Findings

As shown in table 1, the factor loadings of PISTTPAS items range from 0.50 and above; this indicates that, the 57 items have construct validity with respect to pre-service integrated science teacher's teaching practice assessment. This result is in agreement with Plake and Parker's (1982) recommendations that factor loading of 0.05 and above is the minimum for accepting any item to be valid.

After factorial validation of PISTTPAS items, they have been reduced from 83 to 57 and this agrees with DeVellis (2012) and Nworgu (2015) who submitted that one major characteristic of factor analysis is to condense many items into few underlining constructs in order to ensure that a measure is not polluted by elements of other constructs. This implies that the reduction of the PISTTPAS items from 83 to 57 items clustered in 6 factors has ensured that only items that were related to the 6 factors were retained. This suggests that the retained PISTTPAS items are pure and representative of the teaching practice skills. As a result, the teaching practice skills of pre-service integrated science students can now be assessed and scored using PISTTPAS as opposed to the use of un-validated and or poorly validated instruments currently used for assessing pre-service teachers teaching practice.

Additionally, it is clear from the findings of this study that with the use of PISTTPAS, teaching skills acquired by pre-service integrated science teachers' manifest in their classroom teaching and so such skills can now be identified and scored. In other words, teaching skills of pre-service teachers can be measured to an appreciable level of accuracy using PISTTPAS which is a quality poorly validated teaching practice assessment instruments currently in use lacks.

Internal Consistency Reliability Coefficient of PISTTPAS

The findings of this study have revealed that PISTTPAS items have high internal consistency reliability coefficient and are therefore reliable. In other words, items in PISTTPAS have high level of homogeneity. As indicated in table 3, the overall internal consistency reliability of all the clusters/factors of PISTTPAS items is 0.935 this value is high and therefore implies that PISTTPAS is reliable with respect to assessment of teaching practice assessment of pre-service integrated science teachers. To obtain a high reliability coefficient in an instrument is a confirmation of high inter-item consistency and stability, which makes the instrument dependable. This implies that PISTTPAS is reliable and dependable. Therefore, PISTTPAS can reliably and dependably measure teaching skills of pre-service integrated science teachers to a high level of accuracy. This suggest that, PISTTPAS can reliably measure and score teaching skills of pre-service teachers during teaching practice unlike the ones currently in used by universities and other teacher training institutions that lacks reliability index which is a major psychometric property of a good measurement instrument. This result agrees with Morgado, Meireles, Neves, Amaral and Ferreira (2017) who submitted that the use of instrument with high reliability and validity index in measurement of a construct helps to obtain results that are applicable and generalizable; the author further submitted that further studies can be built on them. This suggest that since PISTTPAS has high reliability index, its use in assessing teaching practice of pre-service integrated science is capable of yielding truthful and generalizable results that future studies can build on. Hence, the use of PISTTPAS in assessing pre-service teacher's teaching practice becomes imperative.

Inter-Rater Reliability Coefficient of PISTTPAS

The result of this study has shown that PISTTPAS has high inter-rater reliability coefficient since there was reasonable agreement in the scores by the different raters. The establishment of inter-rater reliability for PISTTPAS becomes pertinent because the rating of teaching skills of pre-service teachers is more subjective than objective (Nworgu, 2015). Two independent raters were therefore involved in observing and rating the pre-service integrated science teacher's teaching skills using PISTTPAS at the same time. From the result of correlation using Pearson as indicated in table 4, the inter-rater reliability coefficient of all the factors is 0.616. This value indicates extent of agreement in the scores of the raters. Thus, pre-service integrated science teachers' lecturers can use PISTTPAS to assess the teaching skills of their students when there is need to authenticate teaching practice scores of pre-service teachers by different raters without much difference in their scores. In other words, the use of the developed instrument will help teaching practice supervisors to assess the teaching skills demonstrated by their students including identifying their strength and weaknesses in different teaching skills like lesson planning, lesson presentation, classroom management, communication, the use of instrumental materials and lesson evaluation without much variation in the scores of the raters. The results of such assessment can be used as a valid evidence of effectiveness of teacher preparation programmes as well.

On the part of the students, the use of PISTTPAS in micro teaching and teaching practice assessment will enable them ascertain the extent to which they have acquired and can showcase teaching skills while still undergoing teaching training. Consequently, the use of the instrument will help the students to understand and improve on the teaching skills there are expected to acquire in each domain of teaching due to the comprehensive and observational nature of the instrument. Hence, because of the robust nature of this instrument, with respect to teaching practice assessment in teacher training institutions like colleges of education, national teachers institute and universities, the reliability coefficients of the instrument developed in this study are considered adequate enough for use in supervision of pre-service integrated science students.

Conclusion

Fifty-seven items of PISTTPAS were found valid for assessing the teaching practice of pre-service integrated science teachers. Six teaching skills (factors) used as constructs were found to explain teaching practice skills. These skills include lesson planning, lesson presentation, classroom management, communication, instructional materials and lesson evaluation. The inter-item consistency reliability analysis of PISTTPAS using Cronbach's alpha technique indicates that the instrument has high internal consistency reliability coefficient of 0.935; and therefore, it is reliable and dependable. The inter-rater reliability coefficient of PISTTPAS using Pearson moment correlation indicates that the instrument has high inter-rater reliability index of 0.616 which implies that there agreement among the raters of the pre-service integrated science teachers.

Recommendations

The following recommendations have been made based on the findings of the study:

1. Teacher training institutions and teachers of pre-service teachers should adopt this instrument for assessing pre-service teacher's teaching practice since it is stable across gender and location.
2. This instrument since it is valid and reliable can be modified and used for assessing the teaching skills of in-service teachers instead of using un-validated/poorly validated and unreliable measurement instruments for assessing the teachers during promotion and other professional exercises.
3. Due to the high internal consistency reliability and validity of the instrument

developed in this study, it should be used for formative instruction of teachers in teacher training institutions.

References

Arias, M.R.M., Lioreda, M.J.H; & Lioreda, M.V.H. (2014). *Psicometria.* South African: Alianza Editorial.

Audet, K.K. (2014). *Assessing pre-service teachers' beliefs: Development of the effective teacher efficacy survey*. Doctoral dissertation, Montana State University, Bozeman, Montana.

Boudersa, N. (2016). *The importance of teachers' training programme and professional development in Nigeria Educational Context: Towards informed and effective teaching practices*. Experiences Retrieve January 6, 2017 from http://exppedagoiens-brandz.

Ezeh, D.N. (2015). Reliability and validity of test. In B.G. Nworgu (Ed). *Educational measurement and evaluation: theory and practice*. Nsukka: Hallman Publisher.

Devellis, R.F. (2012). *Scale development: theory and applications*. Thousand Oaks: Sage.

Hutz, C.S., Bandeira, D.R. & Trntini, C.M. (2015). *Psicometria*. Porto Alegre: Artmed.

Mackenzie, S.B., Podsakoff, P.M. & Podsakoff, N.P. (2011). Construct validation procedures in behavioral research: Integrating new and existing techniques. *Mis Quaterly, 35*(2), 293-334.

Martin, D.L. & Itter, D. (2014). Valuing assessment in teacher education – Multiple choice competency testing. *Australian Journal of Teacher Education, 39*(7), 1-14.

Mehta, J. & Doctor, J. (2013). Raising the bar for teaching. *Phi Delta Kappan,* April, 8-13.

Morgado, F.F.R., Meireles, J.F.F., Neves, C.M., Amaral, A.C.S. & Ferreira, M.E.C. (2017). Scale development: Ten main limitations and recommendations to improve future research practice. *Psicologia: Reflexaoe Critica, 30,* 1-11.

Nworgu, B.G. (2015). *Educational measurement and evaluation: Theory and practice*. Nsukka; Hallman publisher.

Ode, J. O. & Ezeh, D. N. (2019). Scale development in science education research: Validity and reliability related issues. *Education Review Letters, 5*(7), 1-6.

Ode, J. O. & Iloakasia, A. O. (2020). Teacher education programme: The place of teaching practice. *Education Review Letters*, *5*(11), 1-11.

Pasquali, L. (2010). *Instrumenta cao psicologica: Fundamentos epraticas.* Porto Alegre: Artmed.

Plake, B.S. & Parker, C.S. (1982). The development and validation of a revised version of the mathematics anxiety rating scale. *Educational Psychological Measurement*, *42*(2), 551-557.

Prados, J.M. (2007). Development of a new scale of belief about the worry consequences. *Annals of Psychology, 23*(2), 226-230.

Slavec, A. & Drnovsek, M. (2012). A perspective on scale development in entrepreneurship research. *Economic and Business Review, 14*(1), 39-63.

APPENDIX

INSTRUMENT AFTER FACTOR ANALYSIS (CONSTRUCT VALIDATION) PRE-SERVICE INTEGRATED SCIENCE TEACHERS' TEACHING PRACTICE ASSESSMENT SCALE (PISTTPAS)

Section A: Personal Information of Pre-service Integrated Science Teachers.
Please tick (√) as applicable
Name of School: ____

Section B: For raters (Lecturers of Pre-service Integrated Science Teachers)
Instruction: Please tick (√) the appropriate column in the scale to indicate the extent of teaching skill pre-service integrated science teachers demonstrate in teaching as it relates to the following item statements. The four point Likert type scale options are: 4 (Very Good), 3 (Good), 2 (Fair) and (1 Poor).

S/NO	Teaching practice Domain	4	3	2	1
A.	**Lesson Planning**				
1.	Logical arrangement of parts of the lesson plan				
2.	Stating the lesson objectives in measurable terms				
3.	Provision of specific expectations in lesson objectives				
4.	Stating lesson objectives that are achievable within lesson time.				
5.	Provision of relevant assignments				
6.	Clarity of instruction				
7.	Appropriateness of chosen teaching aids				
8.	Provision for real-life application of the lesson				
9.	Relevance of lesson introduction to the topic				
10.	Conciseness of the lesson summary				
11.	Clarity of the procedures in each segment of the lesson				
12.	Provision of pertinent details in the content of the lesson				
13.	Provision of appropriate teacher's activities				
B.	**Lesson Presentation**				
14.	Ability to start with motivating activity to excite the students				
15.	Harmony of ideas in lesson presentation				
16.	Treating each part of the lesson without needless digression.				
17.	Arousing students' attention in learning experiences				
18.	Use of varied teaching methods				
19.	Appropriateness of teaching methods				
20.	Provision of cooperative learning activities				
21.	Use of random selection in calling students to answer questions.				
22.	Use of appropriate methods to take care of individual differences in lesson presentation.				
23.	Rewarding students' efforts with praise				
24.	Addressing areas of students' difficulties in the lesson				
25.	Provision of incentives to students when necessary				
C.	**Classroom Management**				
26.	Appropriate organization of physical space in the classroom				
27.	Elimination of disruptive behaviour among students				
28.	Creating a threat free learning environment in the classroom				

29.	Use of warning signs to control students' disruptive behaviour				
30.	Responding to help signal during class activities				
31.	Using positive feedback on students' response instead of insultive statements.				
32.	Asking students on what to do to improve their learning experiences.				
33.	Use of multi-step approach in handling students with disruptive behaviour.				
D.	**Communication**				
34.	Clarity of expressions				
35.	Use of inclusive words to convey meaning of concepts				
36.	Maintaining eye contact with students during teaching				
37.	Fluency of explanations of concepts during teaching				
38.	Use of positive words on the students				
39.	Having warm facial expression during teaching				
40.	Use of appropriate pace in teaching				
41.	Suitability of the words used based on the developmental level of the students.				
42.	Ability to demonstrate confidence in knowledge of the subject.				
43.	Asking high order questions.				
44.	Giving students time to think before asking them to answer questions.				
45.	Appropriate pronunciation				
E.	**Instructional Materials**				
46.	Appropriateness of the instructional material				
47.	Relevance of the teaching material to the lesson objectives				
48.	Legibility of the instructional material				
49.	Appropriateness of the instructional material to learners' age.				
50.	Provision of guide to students on how to use the instructional material.				
51.	Appropriateness of instructional material to learner's intelligence.				
F.	**Lesson evaluation**				
52.	Determining if lesson was successful by asking students questions relevant to lesson objectives.				
53.	Use of appropriate assessment methods to gather information on students' learning.				
54.	Ability to use multiple assessment techniques to evaluate learning.				
55.	Providing timely feedback to students on learning progress.				
56.	Using the responses to questions for improving learning				
57.	Ability to construct valid assessment questions for learning evaluation.				