
2023

The Development, Validation, and Standardization of a New Tool: The Dyscalculia Test

Stella Eteng-Uket

University of Port Harcourt, Nigeria, stella.eteng-uket@uniport.edu.ng

Follow this and additional works at: <https://digitalcommons.usf.edu/numeracy>



Part of the [Educational Assessment, Evaluation, and Research Commons](#)

Recommended Citation

Eteng-Uket, Stella. "The Development, Validation, and Standardization of a New Tool: The Dyscalculia Test." *Numeracy* 16, Iss. 2 (2023): Article 1. DOI: <https://doi.org/10.5038/1936-4660.16.2.1417>

Authors retain copyright of their material under a [Creative Commons Non-Commercial Attribution 4.0 License](#).

The Development, Validation, and Standardization of a New Tool: The Dyscalculia Test

Abstract

This paper describes a study that focused on developing, validating and standardizing a dyscalculia test, henceforth called the Dyscalculia Test. Out of the 4,758,800 students in Nigeria's upper primary and junior secondary schools, I randomly drew a sample of 2340 students, using a multistage sampling procedure that applied various sampling techniques. For data collection, I used the Test of General Reasoning Ability and Paper 1 of the Mathematics Achievement Test section of the 2021 National Common Entrance Examination, as well as the Dyscalculia Test introduced in this paper, which was developed and standardized in stages. My analysis shows that the Dyscalculia Test items effectively zero in on three components: number sense, arithmetic operation, and working memory. Based on my findings, I recommend that school administrators and counselors adopt the Dyscalculia Test to assess students who may be having difficulty in mathematics and arithmetic for proper diagnosis.

Keywords

dyscalculia test, item response theory, development, validation

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial 4.0 License](https://creativecommons.org/licenses/by-nc/4.0/)

Cover Page Footnote

Dr. Stella Eteng-Uket is a distinguished psychometrician and academician, currently serving as a lecturer in the Department of Educational Psychology, Guidance & Counseling at the University of Port Harcourt, Nigeria. With a Ph.D. in Educational Measurement and Evaluation, she has developed a wealth of expertise in psychometrics, which she has leveraged to improve educational assessment and evaluation.

Introduction

Numbers are omnipresent in our lives. Humans are born intuitively able to make meaning out of them. This intuition allows individuals to acquire and develop numerical and arithmetic skills, and this inherent ability, present as early as infancy, makes it possible for specific numerical skills such as recognizing, counting, adding, comparing, understanding quantities, and the like to develop naturally without formal schooling and into complexity with schooling. The ability of infants to discriminate between small numbers and engage in numerical computation encapsulates this inherent trait (Dehaene 2001; Dehaene et al. 2003; Feigenson et al. 2004; von Aster and Shalev 2007; Nfon 2016). However, in some people, this trait appears to be in deficit, making it challenging to learn numerical and arithmetic skills. Difficulty in learning, specifically numerical, arithmetic, and mathematical skills, is called dyscalculia. According to Ogbogo and Orluwene (2021), other terms used by various authors to refer to dyscalculia include mathematical disability, number fact disorder, developmental dyscalculia, arithmetic learning disability, number fact disorder, psychological difficulties in mathematics, developmental dyscalculia, mathematical disorder, impairment in mathematics, and specific disorder of arithmetical skills.

Individuals may sometimes have specific learning difficulty in areas like reading, writing, or arithmetic/mathematics. Difficulty in learning in any of these specific areas is known as specific learning difficulty, a broad term referring to a diverse group of neurobehavioral disorders characterized by significant unexpected and persistent difficulties in the acquisition and application of efficient reading (dyslexia), writing (dysgraphia), and mathematical (dyscalculia) abilities. Specific learning difficulty occurs despite conventional instruction, intact senses, average intelligence, adequate motivation, and adequate socio-cultural opportunity. Dyscalculia is not a result of general mental impairment or inadequate training. (World Health Organization 2010).

Dyscalculia is characterized by impairment in number sense and memorization of arithmetic facts; inaccurate calculation and mathematical reasoning; difficulty understanding mathematical terms, operations, or concepts; difficulty recognizing or reading numerical symbols or arithmetic signs; difficulty learning, observing, and remembering operational signs; long solution times; inability to follow and remember the sequence of steps used in various mathematical operations; inability to concentrate on mentally intensive tasks; and the like (Geary and Hoard 2001; Shalev and Gross-Tsur 2001; Jordan et al. 2003; Dowker 2004; Geary 2004a, 2004b; Landerl et al. 2004; Beacham and Trott 2005; Doyle 2010; Looi and Kadosh 2010; Trott 2010a, 2010b; American Psychiatric Association 2013; Zerafa 2014; Pandey and Agarwal 2015; Zygouris et al. 2017; Chinn 2020; Ogbogo and Opara 2021; Ogbogo and Orluwene 2021).

Lower academic achievement, higher rates of high school dropout, lower rates of post-secondary education, impaired personality development, high levels of psychological distress and poorer overall mental health, higher rates of unemployment and underemployment, difficulty keeping a job and being promoted within employment, lower incomes and posing a high cost on society and the afflicted person are some of the negative functional consequences of dyscalculia (Landerl et al. 2004; Ghesquiere 2004; von Aster and Shalev 2007; Shalev and von Aster 2008; Gross et al. 2009; Heine et al. 2013). Dyscalculia affects 4–15% of the population, and the prevalence is identical for both genders (Hein et al. 2000; Mazzocco and Myers 2003; Desoete 2004; Koumoula et al. 2004; Dirks et al. 2008; Landerl and Moll 2010; Butterworth et al. 2011; Kaufmann and von Aster 2012; Devine et al. 2013; Moll et al. 2014; Nikolaos et al. 2017; Ogbogo and Opara 2021).

Despite the severe consequences of dyscalculia, it is still often not recognized and has not enjoyed the needed research attention, particularly in relation to its counterpart, dyslexia (Ogbogo and Opara 2021; Ogbogo and Orluwene 2021). Most unfortunate is the fact that instruments to identify dyscalculia are almost non-existent. Therefore, most researchers rely on general standardized mathematics achievement tests or general tests of mathematical abilities, sometimes in combination with intelligence (IQ) measures such as the Wechsler Intelligence Scale for Children and the Wechsler Individual Achievement Test for diagnoses, which are not designed for the sole diagnosis of dyscalculia but are usually deployed due to the unavailability of standardized, precise instruments. This situation is inadequate because intelligence tests tend to correlate with socioeconomic status; hence, those in poverty tend to receive lower scores, and therefore may not accurately reflect aptitude or ability. In addition, some authors have posited that the more significant part of intelligence tests includes mathematical skills, of which the average scores could be misleading. Further, some authors assert that standardized tests are limited in what they can measure (Butterworth 2003; Lyon et al. 2003; Barbaresi et al. 2005; Gersten et al. 2005; Cangoz et al. 2018). The lack of precise tests may also lead to late detection of dyscalculia, which can exacerbate negative outcomes. Furthermore, using intelligence tests as a method for detection is ineffective in terms of intervention because it does not address the underlying causes of the problem or how to manage them (Lyon et al. 2003; Gersten et al. 2005).

There have been some notable efforts to measure dyscalculia, including the dyscalculia screener, software developed by Butterworth in 2003. The dyscalculia screener (Butterworth 2003) is a computer-based assessment for children 6–14 years old that identifies features of dyscalculia by measuring response accuracy and response times to test items. The Mathematics

Education Centre at Loughborough University also developed a screening tool known as dyscalculium in 2005. Their screener includes distinguishing factors that distinguish dyscalculia from other specific learning disorders. Another is the instrument developed by von Aster (2001), a standardized arithmetic test called the Neuropsychological Test Battery for Number Processing and Calculation in Children (NUCALC in English, or ZAREKI in German). NUCALC examines basic calculation and arithmetic skills and identifies dyscalculic profiles. The TEDI-MATH is another instrument for dyscalculia designed by Grégoire et al. (2004) as a test for diagnosing arithmetical disorders. Other tests include the Cognitive Assessment Battery for Dyscalculia (CAB-DC), the TeDDy-PC by Schroeders and Schneider (2008), available in Germany, the BADYS1-4 by Merdian et al. (2015) and BADYS 5-8+ by Merdian et al. (2012) (only available in German).

As outlined above, there have been commendable efforts to assess dyscalculia in various ways. However, certain drawbacks are noticeable, including a lack of solid and comprehensive psychometric properties, insufficient information on the standardization process, small sample sizes, possible technological effects on children's performance for the computerized dyscalculia tests, and an absence of norms. Those mentioned above created a need for developing, validating, and standardizing a new tool called the Dyscalculia Test to measure the underlying factors that can help diagnose persons with dyscalculia.

Some researchers have hypothesized the causes and etiology of dyscalculia (Dehaene 1992; Dehaene and Cohen 1995; Butterworth 1999; Gathercole and Pickering 2000; Geary et al. 2000; von Aster 2000; Geary and Hoard 2001; Temple and Sherwood 2002; Geary 2004a; Passolunghi and Siegel 2004; D'Amico and Guarnera 2005; Van der Sluis et al. 2005; Rosselli et al. 2006; Bull et al. 2008; Mabbott and Bisanz 2008). Some have hypothesized types of dyscalculia (Kosc 1974; Geary 1993, 2004a). At the same time, others have outlined the dimensions and classification (Karagiannakis et al. 2014). Based on a review of the above and a discussion of the characteristics of dyscalculia from the preceding sections, I drew and outlined three broad core domains of dyscalculia as follows.

Number Sense Domain refers to an intuitive feel for numbers and a common-sense approach to using them. This domain includes sub-domains like subitising and numerizing (the ability to instantaneously recognize the number of objects in a small group without the need to count them), counting (the ability to determine the quantity or the total number of objects or sets in a group), and meaning, or knowledge and understanding of number (conceptual understanding of numbers, number symbols, vocabulary, and meaning as well as the ability to discriminate and make numerical magnitude).

Arithmetic Operations and Computation Domain refers to the four basic operations and computations with the required fluency (accuracy, automaticity, and flexibility). This domain includes the four subsets of addition, subtraction,

multiplication, and division operations and computation (ability to perform addition, subtraction, multiplication, and division tasks fluently).

Working Memory Domain deals with the ability to hold information in short-term memory and manipulate that information. This domain includes the subset of rotated cubes activities (ability to understand visual information, processing of simple cognitive tasks, and also attention), letter-number sequencing activities (ability to hold verbal information in memory while it is being manipulated, short-term storage of linguistic information and other information), digit span activities (short-term memory and attention), and matrix reasoning (non-verbal problem solving and an individual's ability to understand complex visual information).

As I had mentioned earlier, the Dyscalculia Test assesses ability in number sense, arithmetic, computational operations, and working memory. Items in these areas were developed, and their psychometric properties were established using Item Response Theory. Psychometrics is a field concerned with the theories and techniques of psychological measurement of knowledge, abilities, attitudes, traits, and so on: within it are two core theories for test development which are the Classical Test Theory (CTT) and Item Response Theory (IRT). (Kpolovie 2010, 2011, 2014, 2016). The IRT framework is based on some assumptions and uses different statistical approaches. IRT is concerned not only with developing, evaluating, or determining the reliability and validity of tests, but also with holistically improving the quality of test items (Awopeju and Afolabi 2016). In addition, IRT describes the relationship between an examinee's test performance and the traits assumed to underlie such (Awopeju and Afolabi 2016). IRT has the basic assumption of item unidimensionality and item local independence. Furthermore, IRT holds that the probability of a person endorsing an item is a function of some parameters. The parameters are one-parameter ("b," which is the difficulty indices/slope) two-parameter ("a" and "b" for difficulty and discrimination indices), and three-parameter ("a," "b," and "c" for difficulty, discrimination indices, and guessing). Whatever the framework used, the fundamental concern of test developers is that the test items are well-analyzed, suitable, and possess excellent and high psychometric properties.

Psychometrics develops "tests" to make measurements. Opara (2016) describes a test as an instrument or procedure designed to measure the knowledge, intelligence, ability, trait, skills, aptitude, interest, and attitude an individual or thing possesses. Test development refers to the preparation of item writing, item analysis, selection of reliable and valid items for the final test, and testing of the reliability and validity of the developed test (validation). Test standardization deals with uniformity of procedure for administration, test scoring, and norms development after a test has been constructed and

validated. I summarized the following steps as threads observed from literature about the process of test development, validation, and standardization; planning for the test, constructing the test items following the rules guiding the format of the test chosen, evaluating the test items (including experts review, item analysis such as item difficulty and discrimination, preliminary reliability, and validity on items), preparing and administering the final draft, validation (establishing validity and reliability), and standardization (establishing test norms and providing technical manuals) (Kpolovie 2012; Orulwene 2012; Iweka 2014; Opara 2016; Price 2017; Irwing and Hughes 2018).

Over time researchers have developed and validated tests observing the above steps using the IRT framework. They include the Dyslexia Screening Instrument by Coon et al. (1994), the Dyslexia Screening Tool (DEST-1 & 2) by Fawcett and Nicolson (1996), the Computerized Number Sense Scale by Yang et al. (2003, 2008), the Automated Working Memory Assessment by Alloway (2007), the Physics Achievement Test by Metibemu (2016), the Basic Number Processing Test by Olkun et al. (2016), the Test of Dyslexia and Dysgraphia by Cox (2002), the Chemistry Achievement Test by Oku and Iweka (2018), and the abridged version of the Advance Progressive Metric by Kpolovie and Emekene (2016) among others.

From the preceding, dyscalculia is a specific and prevalent learning difficulty in arithmetic with alarming consequences and invariably needs to be accurately identified and treated. Also, measures that have attempted to measure it are include some limitations. Furthermore, a general test of mathematical abilities combined with measurements of intelligence cannot give an accurate picture of dyscalculia. It follows that there is a dire need for developing, validating, and standardizing a specific scale with solid psychometric properties and a robust standardization process for diagnosis, which will trigger prompt necessary interventions. These factors necessitate the development and standardization of a precise new tool called the Dyscalculia Test following the steps outlined in test development, validation, and standardization using the IRT framework.

Research Questions

The following research questions guided the study.

- What are the dimensions/factor structures of the Dyscalculia Test?
- To what extent does the Dyscalculia Test obey the assumption of item unidimensionality and item local independence?
- What are the item difficulty index and item discrimination index of the Dyscalculia Test?
- What are the test response function, item information function, and overall model fit of the Dyscalculia Test?
- What is the internal consistency coefficient of the Dyscalculia Test?
- What is the frequency, percentage, simple ranking, percentile, z scores, and T scores of the Dyscalculia Test?

Methodology

Research Design

The study was designed using the instrumentation research design, which is primarily used for test development based on test theories (like classical test theories, item response theory, and the like) to ensure satisfactorily high validity and reliability. In addition, instrumentation design helps to give the most appropriate norm, criterion, or domain in measuring and evaluating psychological attributes and human abilities (Kpolovie 2010).

The population of the study. I carried out the study in Nigeria within four geopolitical zones. The population of the study comprised all the students (4,758,800) in the (54,434) upper primary and (13,029) junior secondary school students in Nigeria.

Sample and sampling technique. The sample of the study was 2340. This sample size was determined using Taro Yamane's minimum sample as a guide. I employed a multistage sampling technique. Several sampling techniques were employed at different stages using techniques like simple random, clustered, stratified, and purposive. An initial sample size of 3000 was administered with two instruments, the Test of General Reasoning (TOGRA) and the Mathematics Achievement Test (MAT), to screen for those who will form the final sample size. I purposefully excluded students from the sample if they met the following criteria: significantly below average according to the Test of General Reasoning, which was the intelligence test employed; and very significantly low performance in the Mathematics Achievement Test. Respondents were drawn from the MAT test score skewed toward the lower performing end but not substantially lower performing end. I choose this sampling strategy to maximize the sensitivity of the standardization to those pupils likely to be dyscalculic while retaining representativeness overall. However, this process may have omitted students who may be dyscalculic. Those with severely impaired vision and hearing were also not part of the sample. Thus, only 2340 students who scaled the cut-off criteria from the mathematics achievement test and intelligence test with their questionnaires correctly filled formed the final sample size. This 2340 sample comprised students in upper primary and junior secondary school, males and females, whose ages ranged from seven to thirteen years old and spread over four main ethnic groups (Hausa, Igbo, Yoruba, and Minorities). The sample size, therefore, consisted of 1360 Primary and 980 Junior Secondary School Students.

Instruments for data collection. I used three instruments for data collection. The General Reasoning Ability (TOGRA), Paper 1 of the Mathematics Achievement

Test section of the 2021 National Common Entrance Examination (NECO), and the Dyscalculia Test, which was constructed, validated, and standardized. TOGRA was developed by Reynold (2014), the TOGRA is a speeded measure of reasoning ability (intelligence) and problem-solving skills. The TOGRA was used as the intelligence test to screen for only those with average and above intelligence to form the final sample. The second instrument was the adapted Mathematics Achievement Test (MAT) containing multiple choice-type questions constructed by subject experts and developed by NECO into test form. I used the MAT to screen the students such that those who performed too poorly below the set cut-off would not form the final sample size. Respondents were drawn from the MAT test score skewed toward the lower performing end, but not the significantly lower performing end; those with severely poor academic achievement in mathematics.

The third instrument was the Dyscalculia Test. I carried out the development, validation, and standardization processes in the following phases.

Phase one: constructing and preparing the preliminary draft of the test items for the Dyscalculia Test. At this phase, based on a review of research reports and documents on the characteristics, dimensions, and classifications, I outlined three broad domains. The domains are Number Sense, Arithmetic Operation, Computation, and Working Memory Domain. The researcher drew 140 multiple-choice test items from these domains to develop the Dyscalculia Test.

Phase two: editing and reviewing the preliminary draft of the test items for the Dyscalculia Test. After formulating the initial 140 draft items of the Dyscalculia Test, the items were re-edited and revised based on the opinions of test and measurement experts, counseling psychologists, mathematics subject specialists, and language experts. I gave the items to these experts to analyze the content and language critically, correct ambiguities, and check that all the defined objectives were tested. Based on experts' input, I retained only those items that received at least 70% approval from the experts for item analysis.

Phase three: pilot study/item analysis and selection of items for a final draft for the Dyscalculia Test. The 140 Dyscalculia Test items were administered to a trial sample of 150 students from upper primary and junior secondary school students. After that, I carried out a preliminary factor analysis to see initial evidence of the hypothesized dimensions of the test. Then preliminary item analysis of difficulty and the discriminatory index was also done. Of the 140 items, 40 items were loaded on component 1 (arithmetic operations), 42 items were loaded on component 2 (number sense), and 33 items were loaded on component 3 (working memory). Also, from the result, I rejected two items based on their difficulty indices being too high (the test was too easy), and coincidentally these two items did not

load on any factor. Six items were rejected because they could not discriminate well between high and low achievers (had negative signs), two of which coincidentally did not load on any factor. Thus, in total, 25 items were rejected. Therefore, the final draft after the pilot study contained 115 items (see Supplement D in the supplemental documents). After trial testing, based on feedback from fieldwork and expert review, a further revision was made to the items, and 5 items were rejected based on their structuring, leaving 110 items. Although, after item-by-item analysis using the IRT Framework, 85 items formed the Dyscalculia Test.

Instruction, scoring, and classification. The Dyscalculia Test is a 45-minute timed test containing 85 items in an item booklet. It was designed using the multiple-choice question format with four options lettered A, B, C, and D, which has a key (correct response) and three distracters (incorrect response), and test takers are expected to provide the answer in a separate answer sheet (a sample of test items is attached in the Appendix),

A total of 85 items gives a raw score of 85 from the three domains. The total score index from the three core domains of the Dyscalculia Test is summed up to give a score representing an individual's dyscalculia score. The core domain from which an individual has the lowest score indicates the dimension on which an individual is most dyscalculic. The highest possible score from the scale is 85, and the lowest score is 0. Scores from 64 and above show that an individual is not dyscalculic, while scores from 63 and below show dyscalculia.

Based on the DSM-5 recommendation on how specific learning difficulties like dyscalculia are to be categorized, scores from this Dyscalculia Test are classified as follows.

No Dyscalculia (total scores from 64 and above): A score of 64 and above indicates that an individual performed significantly above average in the three subsets of the test. Also, a raw score of 64 and above, when transformed, places individuals between 0.83 to 2.97 deviations above the mean of 50. This result shows that an individual does not have a specific learning disorder in mathematics.

Mild Dyscalculia (total scores between 43–63): A score between 43–63 indicates that an individual performed around average in the three subsets of the test. When transformed, a score in this range places individuals between -0.31 deviations below the mean and 0.77 above the mean of 50. Scores in this range indicate some difficulties in learning skills in the three core domains but, mild enough in severity that the individual may be able to compensate or function well when provided with appropriate accommodations or support services, especially during the school years.

Moderate Dyscalculia (total scores between 23–42): A score between 23–43 indicates that an individual performed below average in the three subsets of the test. Also, when transformed, scores in this range place individuals between -1.19

to -0.17 deviations below the mean of 50. Scores in this range indicate significant difficulties in learning skills in the three core domains, so the individual is likely to become proficient with some intervals of intensive and specialized teaching during the school years. Some accommodations or supportive services may be needed for at least part of the day at school, in the workplace, or at home to complete activities accurately and efficiently.

Severe Dyscalculia (total scores between 0–22): Scores between 0–22 indicate that an individual performed significantly low in the three subsets of the test. Also, when transformed, a score in this range places individuals between -2.54 to -1.24 deviations below the mean of 50. Scores in this range indicate severe difficulties in learning skills in the three core domains, so the individual is likely to learn those skills with intensive, individualized, specialized teaching for most of the school years. However, with an array of appropriate accommodations and services at home, school, and in the workplace, the individual may be able to complete all activities, but could be more efficient.

For scoring norms, see Supplement C in the supplemental documents.

Validity of the Instruments

The Test of General Reasoning Ability (TOGRA) has an internationally acclaimed validity. The test has 0.75 to 0.95 as construct validity via correlation with other tests (RAIT, WISC-IV, WAIS-IV, RIAS, Wonderlic, Beta III, WRAT, and TIWRE). Furthermore, the face and content validity of the Mathematics Achievement Test (MAT), based on Paper 1 of the Achievement section of the Mathematics section of the 2021 National Common Entrance Examination, has been established. This result is due in part because the questions are owned by the examining body NECO and have been validated by experts.

The face and content validity of the Dyscalculia Test was determined using expert judgment. The construct validity of the Dyscalculia Test was estimated using multivariate factor analysis. The principal component analysis was used for processing the data. The varimax Kaiser Normalization extraction method and the rotated factor loading matrix were used to estimate the construct validity. For instance, for the Number Sense Domain, rotated factor loadings ranged between 0.31 to 0.755. For the Arithmetic Operations and Computation Domain, it ranged between 0.30 to 0.54, while for the Working Memory Domain it ranged between 0.30 to 0.50. The eigenvalues of one above were used to select items that measure the similar construct.

Reliability of the Instruments

The reliability of TOGRA, as reported by Reynold (2014), ranges from 0.74 to 0.99 for test-retest reliability, 0.87 to 0.94 for Cronbach's alpha reliability, and 0.85 to 0.94 for alternate form reliability. Although the above instrument has known

reliabilities, the researcher reestablished the reliability. When the parallel form technique was used, a coefficient of $r = 0.617$ was obtained. Split-half reliability analysis showed a reliability estimate of the first half of the test to be 0.827 and the second part of the test to be 0.922. To estimate the reliability of the whole test, Spearman-Brown yielded a coefficient of 0.677 while a Cronbach's alpha coefficient of 0.904 was obtained. The Mathematics Achievement Test (MAT) reliability was established using the split-half reliability analysis. This gave the first half of the test to 0.600 and that of the second part of the test to 0.66. To estimate the reliability of the whole test, Spearman-Brown yielded a coefficient of 0.700.

The split-half reliability analysis for Dyscalculia Test shows the reliability estimate of the first half of the test to be 0.894 and the second part of the test to be 0.780. To estimate the reliability of the whole test, Spearman-Brown yielded a coefficient of 0.824. The Kuder Richardson 20 coefficient was found to be 0.910 for the whole test and 0.86, 0.92, and 0.60 for the three domains, respectively.

Method of Data Analysis

Statistical analyses were done using Item Response Theory logistic models (2 Parameters Model), factor analyses, Kuder-Richardson formulas (Kuder-Richardson 20), z -score, T -score, number counts, frequency, percentage, and percentile ranks. All these were done using statistical software packages such as Item Response Theory application using Excel add-in (EIRT), Xcalibre 4.2 software, Statistical Package for Social Science (SPSS), and Microsoft EXCEL.

Results

Research Question 1: What are the dimensions/factor structures of the Dyscalculia Test? Table 1a (see Supplement A in the supplemental documents) shows the number of factors and, by extension, the domains of the Dyscalculia Test. My analysis from the table shows that the Dyscalculia Test items loaded on three (3) components. It shows that 37 items were loaded on component 1, 32 items were loaded on component 2, and 39 items were loaded on component 3. The values ranged from 0.30 to 0.62 for the test.

Table 1b further shows that all 37 items with values ranging from .37 to .62 loaded on component 1 measured a similar domain of Arithmetic Operations. Similarly, all 32 items with values ranging from .35 to .56 that loaded on component 2, measured a similar domain of Working Memory. In comparison, all the 39 items with values ranging from .30 to .55 loaded on component three measured a similar domain of Number Sense. This shows that the Dyscalculia test has Number sense, Arithmetic Operations and Computations, and Working Memory as its domains.

Table 1b
Summary of The Dimensions/Factor Structures of The Dyscalculia Test Using Factor Analysis

Factors	Dimension/Domain measured by factor	Items loaded	Items Loaded Value Ranges
1	Arithmetic Operations and Computation Domain	41,42,43,44,45,46,47,48,49,50,51,52,53,54,55,56,57,58, 59,60,61,62,63,64,65,66,67,68,69,70,71,72,73, 74,75,76,77 (total of 37)	0.37–0.62
2	Working Memory Domain	78,79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,94,95, 96,97,98,99,100,101,102,103,104,105,106,107,108,109 (total of 32)	0.35–0.56
3	Number Sense	2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,2 3,24,25,26,27,28,29,30,31,32,33,34,35,36,37,38,39,40 (total of 39)	0.30–0.55

Research Question 2: To what extent does the Dyscalculia Test obey the assumption of item unidimensionality and item local independence?

Furthermore, dichotomous test items are unidimensional when the first-factor loading for all items is significantly greater than one. This was the case in this result. Also, unidimensionality is assumed when the first eigenvalue is substantially more significant than the next. For example, the distance between the first eigenvalue of 13.31 and the next 7.31 shows that the distance between the two components is significant. As the difference between the two components was large enough and substantially greater than 1, it suggests that the Dyscalculia Test items are unidimensional.

Table 2
The Extent Dyscalculia Test Obeys the Assumption of Item Unidimensionality and Item Local Independence Using Eigenvalues and Screen Plot of Factor Analysis

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	13.317	12.106	12.106	13.317	12.106	12.106
2	7.316	6.651	18.758	7.316	6.651	18.758
3	4.819	4.381	23.139	4.819	4.381	23.139
4	3.069	2.790	25.929			
5	2.669	2.427	28.356			
+	+	+	+			
+	+	+	+			
+	+	+	+			
105	.341	.310	98.580			
106	.327	.297	98.878			
107	.322	.293	99.171			
108	.317	.288	99.458			
109	.306	.278	99.736			
110	.290	.264	100.000			

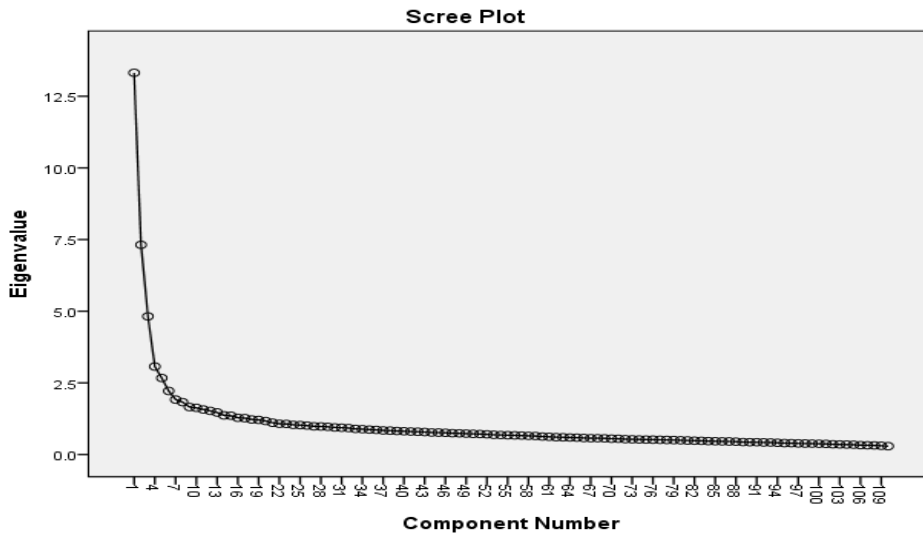


Figure 1: Scree plot confirming unidimensionality.

A careful examination of the scree plot (Fig. 1) shows only one construct before the breaking point or elbow joint, therefore succinctly showing that the Dyscalculia Test measures the underlying construct of specific difficulty with arithmetic. The underlying construct is effectively examined by the test, and it ensures its unidimensionality. Moreover, since this model meets the assumption of unidimensionality, it invariably means that local independence holds.

Research Question 3: What are the item difficulty index and item discrimination index of the Dyscalculia Test?

The IRT parameters table (see Supplement B in the supplemental documents) presents the IRT item parameters. The parameter a indexes the discrimination of the item, as larger values for a will result in a greater slope of the IRF and indicate that the item differentiates examinees well. Negative values are considered poor discriminators, and positive values are good discriminators. My analysis presents the assessment of the Dyscalculia Test using the set criteria for $a \geq 0.2$. Using this criterion, items whose a fell below or equal to 0.199 or had negative values were considered “poor,” while outside this range was considered “good.” A careful examination of the a -parameter column shows that values obtained ranged from 0.90 to 0.02. Also, for the good items, the values of the Dyscalculia test ranged from 0.200 to 0.911, and 25 items were viewed as poor.

The b parameter is the item difficulty parameter. Higher b parameters (> 1.0) indicate that the item is more difficult; a value below -1.0 indicates that the item is very easy. According to the Xcalibre manual, the difficulty index “ranges in theory from negative to positive infinity, but in practice from -3.0 (very easy) to $+3.0$ (very

difficult). Using set criteria for b parameter of -2.99 (very easy) to +2.99 (very difficult) as “good” items and outside these ranges as “poor” items, the table presents the assessment of the 85 items that were good and 25 items that were poor in the Dyscalculia Test using the set criteria for item difficulty. Careful examination of the b parameter column shows that the values of b for the Dyscalculia Test range from 4.00 to -2.81. Also, for the good items, the values of b for the Dyscalculia test range from 2.900 to -2.81. A total of 25 items were considered poor.

Research Question 4: What are the test response function, test information function, and the Overall Model Fit of the Dyscalculia Test?

Table 4 presents the Overall Model Fit with a Chi-Square value of 22655.415, a degree of freedom (df) of 1430, a probability of 0.000, and a two logistic likelihood of 290242. The total calibrated Dyscalculia Test items fit the 2 Parameter Logistic Model as the p -value of 0.000 is less than the chosen alpha and therefore is not significant.

Table 3
Overall Model Fit

Test	Items	Chi-square	Df	p -value	-2LL
Full Test	110	22655.415	1430	0.000	290242

The Test Information Function (TIF) summarizes the item information functions where the test provides information. The Test Response Function (TRF) predicts the proportion or number of items an examinee would answer correctly as a function of theta. In Figure 2, the left Y-axis is in proportion-correct units, while the right Y-axis is in number-correct units. In this instance, TRF predicts 88.5%. That is approximately 89% or is equivalent to the score of each examinee on the Dyscalculia Test.

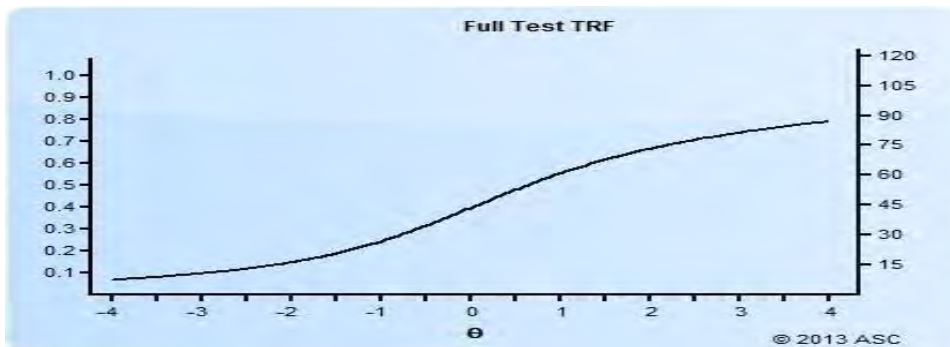


Figure 2. A graph of the Test Response Function (TRF) for all calibrated items.

TIF is a graphical representation of how much information the test provides at each level of theta. For example, the test information function in Figure 3 shows

that the maximum amount of information provided by the Dyscalculia Test was 19.516 at a theta, i.e., an ability level of 0.000, or the point at which the curve peaks. Therefore, maximum information was 19.516 at $\theta = 0.000$. At the cutpoint of $\theta = 0.650$ (EPC = 0.500), the TIF equaled 17.086. In this case, the TIF provides satisfactory information over the ability trait range since it takes the shape of a normal distribution curve.

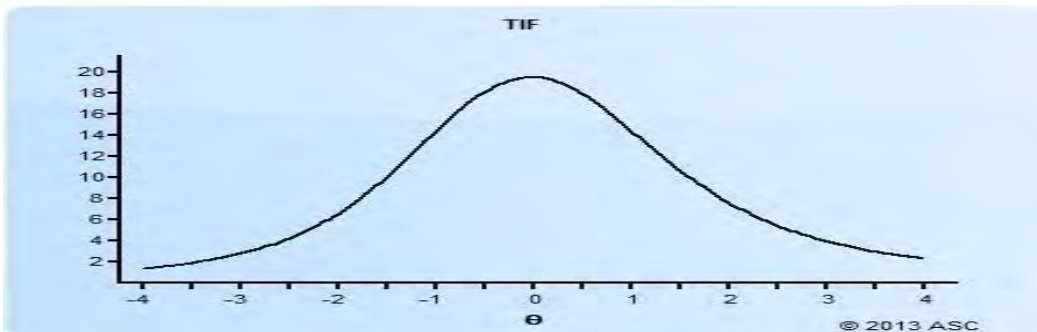


Figure 3. Test Information Function (TIF).

Research Question 5: What is the internal consistency coefficient of the Dyscalculia Test?

Table 5 shows that the Number Sense Domain has a Kuder-Richardson 20 coefficient of 0.888, the Arithmetic Operations and Computation have a KR-20 coefficient of 0.931, and Working Memory has a KR-20 coefficient of 0.910. As a whole, the Dyscalculia Test has a KR-20 coefficient of 0.930, a very high coefficient indicating the Dyscalculia Test is very highly reliable. The internal consistency of the Dyscalculia Test was further established using the split-half method.

Table 4
Internal Consistency Coefficient of Dyscalculia Test Using the Kuder-Richardson 20 Method

Total Number	Domain	Total Score	Mean X	Standard Deviation	Variance	K	KR ₂₀
2340	Dyscalculia	109634	46.852	18.35	336.78	85	0.93
2340	Number Sense	3388666	13.20	8.90	79.21	31	0.88
2340	Arithmetic Operations	3654466	15.33	9.00	81.00	37	0.93
2340	Working Memory	3350066	12.43	7.11	50.55	17	0.89

Table 6 shows the split-half reliability analysis for Dyscalculia Test, with the reliability estimate of the first half of the test to be 0.907 and the second part being 0.922. To estimate the reliability of the whole test, Spearman-Brown yielded a coefficient of 0.886. Therefore a split-half coefficient of 0.886 was obtained.

Table 5**Internal Consistency Coefficient of Dyscalculia Test Using the Split-Half Method**

Cronbach's Alpha	Part 1	Value	0.907
		N of Items	43 ^a
	Part 2	Value	0.922
		N of Items	42 ^b
Total N of Items			85
			0.321
Correlation Between Forms			
Spearman-Brown Coefficient	Equal Length		0.886
	Unequal Length		0.886
Guttman Split-Half Coefficient			0.885

Research Question 6: What are the frequency, percentage, simple ranking, percentile, z scores, and T scores of the Dyscalculia Test?

A table containing figures relating to the Dyscalculia Test for Upper Primary Students in Nigeria is attached in Supplement C in the supplemental documents.

Discussion of Findings

Establishment of the Factor Structure, Unidimensionality, and Local Independence of the Dyscalculia Test

One of my findings from this research shows the number of factors and, by extension, the domain of the Dyscalculia Test. It shows that the Dyscalculia test items loaded on three factors. This shows that the Dyscalculia test has Number Sense, Arithmetic Operations and Computations, and Working Memory as its domain. This study was able to establish the domains of the Dyscalculia Test empirically, a significant finding as persons with dyscalculia can now be identified effectively. Also, the dimensions on which intervention is needed can easily be identified and supplied. This result is somewhat in tandem with the dyscalculia screener developed by Butterworth in 2003. Butterworth developed a dyscalculia screener that assesses four primary areas; one of the areas, which is the test of achievement (arithmetic achievement test, i.e., addition and multiplication), is similar to the arithmetic and computational (addition, subtraction, multiplication, and division) domain of this study. Although similar, vast differences exist: no information regarding the statistical procedure employed in establishing these areas of assessment and no report concerning the item-by-item analysis carried out on the dyscalculia screener. This finding is also in line with that of Yang et al. (2008). They developed a computerized number sense scale (CNST) which was empirically and theoretically supported via confirmatory factor analysis and literature review. Also similar to this is the Grégoire et al. (2004) Test for Diagnostic Assessment of Mathematical Disabilities, the Romanian Screening Test (ST) for dyscalculia by Gliga and Gliga (2012), the Alloway (2007) Automated Working Memory Assessment (AWMA), Beacham and Trott's (2005) DyscalculiUM and von Aster's

(2001) Neuropsychological Test Battery for Number Processing and Calculation in Children. Although these works are similar in that they were designed to identify arithmetic and mathematics difficulties, vast differences exist. They include the issue of the development of these scales not having information regarding the statistical procedure employed in establishing the assessment of these areas, no report concerning the item-by-item analysis carried out on the test, and they were not developed in Nigeria. In all, none of the tests were developed and validated using IRT.

Another finding from this research is that the Dyscalculia Test obeyed the assumption of unidimensionality and local independence. This succinctly shows that the Dyscalculia Test measures the underlying construct of specific difficulty with arithmetic. This result was in line with the set condition for assessing the unidimensionality of items in a test by Hambleton (2004). According to him, dichotomous test items are unidimensional when first-factor loading for all items is significantly greater than the eigen value 1. This value suggests that the dyscalculia items are unidimensional. Since the assumption of unidimensionality holds, it invariably means that local independence holds (Ubi 2006). This finding is in line with that of Orulwene and Asiegbu (2016), who checked the assumptions of unidimensionality and local independence using factor analysis. The study's result revealed that the test items met the assumptions of unidimensionality and local independence when they investigated bias in a test using the IRT model. This result is also in tandem with the findings of Kpolovie and Emekene (2016), Metibemu (2016), and Emekene 2017 regarding establishing unidimensionality for their test instruments.

Item by Item Analysis for the Dyscalculia Test Using Particular Item Response Theory Characteristics

My analysis considered a total of 25 items “poor” based on their a (discriminatory) and b (difficulty) parameters. This result aligns with the study of Kpolovie and Emekene (2016). The result also shows that the total calibrated Dyscalculia Test items fit the 2 Parameter Logistic Model. The TRF could predict the proportion or number of items an examinee would answer correctly as a function of theta. In this instance, TRF predicts 88.5%. That is approximately 89% or equivalent to the score of each examinee on the Dyscalculia Test. That is, TRF predicts 88.5%. This means that approximately 89% of the score of each examinee is predicted by the theta (trait) of dyscalculia. The test information function shows that the maximum amount of information provided by the Dyscalculia Test was 19.516 at a theta, i.e., an ability level of 0.000. The TIF provides sufficient information over the ability trait range since it takes the shape of a normal distribution curve. Using the test information function, the standard error and an assessment of the information each contributes to the overall information supplied by the full test and how each item

fit the model, 25 items were seen as poor. The remaining items were seen as good. The selected items also were based on their order of difficulty and fair contributions to the item-total correlation (i.e., discrimination). These 85 good items were what we were subjected to validity and reliability analysis. This finding is in sync with Oku and Iwaka's (2018) findings and Emekene's (2017), who used TRF and TIF to get the test information function and item response function of their assessment instruments. This is also in line with the findings of Metibemu (2016) who used TIF to get information about the PAT instrument he developed.

Establishment of Validity and Reliability (Validation) of the Dyscalculia Test: The split-half reliability estimate, the Kuder-Richardson 20 internal consistency estimate of the test as a whole, and the domains as well were high as seen by the values obtained. The construct validity of the Dyscalculia Test was also estimated using multivariate factor analysis. This result revealed that the Dyscalculia Test fulfilled the assumption of unidimensional as factor analysis results aligned with the set condition for assessing it. It succinctly showed that the Dyscalculia Test measures the underlying construct of specific difficulty with arithmetic. These findings further significantly affirm that the Dyscalculia test items are empirically fit to assess specific learning difficulties in arithmetic. This finding is consistent with Kpolovie and Emekene (2016). They found a high-reliability index when KR-20 and split-half were employed to determine reliability as well as factor analysis and hypothesis testing to determine validity.

Frequency, percentage, simple ranking, percentile, z scores, and T scores were obtained separately for Dyscalculia Test for upper primary and junior secondary in Nigeria. This presents a range of norms. Therefore, any future test on the Dyscalculia Test by any upper primary or junior secondary school student in Nigeria or anybody aged 7 to 13 can be easily transformed from the raw score to any of the norms. For instance, the single percentile ranks table generated through this research work is suitable for upper primary and junior secondary school students because all the likely differences in the raw scores have already been evened out with their transformation into a normalized score.

Conclusion

The Dyscalculia Test was successfully developed and constructed, and it measures the dimensions and domain of number sense, arithmetic operations, and working memory. The Dyscalculia Test fulfills and obeys the assumption of unidimensional and local independence holds. Under the IRT Framework, it is concluded that the Dyscalculia Test items have good item difficulty and discrimination indices. The fully calibrated test items fitted the 2 Parameter Logistic Model. The Test Response Function of the Dyscalculia Test predicts 88.5%. The test information function shows the maximum amount of information provided by the Dyscalculia Test. It is

concluded that the TIF provides sufficient information over the ability trait range. It is concluded that the content and construct validity were sufficiently established using expert judgment, factor analysis, and high-reliability indices. The Dyscalculia Test was sufficiently standardized and normed using frequency, percentage, simple ranking, percentile, z scores, and T scores in Nigeria.

This has implications for psychologists, counselors, students, teachers, school administrators, and parents as I believe this will make possible early identification of persons having difficulty with arithmetic skills and subject them to timely interventions. Psychologists, counselors, and teachers have an instrument they can use to quickly identify people with dyscalculia and then promptly apply necessary interventions. This allows teachers, administrators, parents, and even students to be aware of and actively participate in the necessary interventions that will help the students overcome this specific learning difficulty. Hopefully this will raise awareness of the existence of dyscalculia and promote the advancement of dyscalculics in society.

References

- Alloway, Tracy P. 2007. *Automated Working Memory Assessment*. London: Harcourt Assessment.
- American Psychiatric Association. 2000. *Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) (4th ed. online)*. Arlington, VA, APA.
- American Psychiatric Association. 2000. *Diagnostic and Statistical Manual of Mental Disorders Test Revision (4th ed., text rev.)*. Washington, DC. <https://doi.org/10.1176/appi.books.9780890423349>
- American Psychiatric Association. 2013. *Diagnostic and Statistical Manual of Mental Disorders (5th ed.)*. Washington, DC. <https://doi.org/10.1176/appi.books.9780890425596>
- Awopeju, O. A., and E. R. I. Afolabi. 2016. “Comparative Analysis of Classical Test Theory and Item Response Theory Based Item Parameter Estimates of Senior School Certificate Mathematics Examination.” *European Scientific Journal*, 12(28): 263–284. <https://doi.org/10.19044/esj.2016.v12n28p263>
- Barbarese, W., S. Katusic, R. Colligan, A. Weaver, and S. Jacobsen. 2005. “Learning Disorder: Incidence in a Population-Based Birth Cohort, Rochester” *Minn. Ambulatory Pediatric* 5: 281–289. <https://doi.org/10.1367/A04-209R.1>
- Beacham, N., and C. Trott. 2005. “Screening for Dyscalculia Within HE.” *MSOR Connections* 5(1): 5–8. <https://doi.org/10.11120/msor.2005.05010004>
- Bull, R., K. A. Espy, and S. A. Wiebe. 2008. “Short-term Memory, Working Memory, and Executive Functioning in Preschoolers: Longitudinal Predictors

- of Mathematical Achievement at Age 7 Years.” *Dev Neuropsychol*, 33(3): 205–228. <https://doi.org/10.1080/87565640801982312>
- Butterworth, B. 1999. *The Mathematical Brain*. Papermac.
- Butterworth, B. 2003. *Dyscalculia Screener*. London: Nelson Publishing Company Limited.
- Butterworth, B., S. Varma, and D. Laurillard. 2011. “Dyscalculia: From Brain to Education.” *Science*, 332(6033): 1049–1053. <https://doi.org/10.1126/science.1201536>
- Cangoz, B., S. Olkun, A. Altun, and F. Salma. 2018. “Comparing Two Cut-off Based Criteria While Investigating the Discriminatory Characteristics of a Tablet-Based Dyscalculia Screening Battery for 5–9 Age Group.” *Noro Psikiyatrs Ars*. 55(2): 177–182.
- Chinn, S. 2020. “Dyscalculia.” <http://www.stevechinn.co.uk/dyscalculia/what-is-dyscalculia>. (Accessed March 27, 2020)
- Coon, K. B., M. M. Waguespack, and M. J. Polk. 1994. *Dyslexia Screening Instrument*. San Antonio: TX. The Psychological Corporation: Harcourt Brace & Company.
- Cox, E. A. 2002. An evaluation of the psychometric properties of the test of dyslexia and dysgraphia. Doctoral dissertation. University of Tennessee, Knoxville. https://trace.tennessee.edu/utk_graddiss/6218.
- Dehaene, S., & Cohen, L. 1995. Towards an anatomical and functional model of number processing. *Mathematical Cognition*, 1, 83–120
- D’Amico, A., and Guarnera, M. 2005. “Exploring Working Memory In Children With Low Arithmetical Achievement”. *Learning and Individual Differences*, 15(3):189–202. <https://doi.org/10.1016/j.lindif.2005.01.002>
- Dehaene, S. 1992. “Varieties of Numerical Abilities.” *Cognition*, 44(1-2): 1–42. [https://doi.org/10.1016/0010-0277\(92\)90049-N](https://doi.org/10.1016/0010-0277(92)90049-N)
- Dehaene, S. 2001. “Précis of the Number Sense.” *Mind Language* (16): 16–36. <https://doi.org/10.1111/1468-0017.00154>
- Dehaene, S., M. Piazza, P. Pinel, and L. Cohen. 2003. “Three Parietal Circuits for Number Processing.” *Cognitive Neuropsychology* (20): 487–506 <https://doi.org/10.1080/02643290244000239>
- Desoete, A., H. Roeyers, and A. De Clercq. 2004. “Children with Mathematics Learning Disabilities in Belgium.” *Journal of Learning Disabilities* (37): 50–61. <https://doi.org/10.1177/00222194040370010601>
- Devine, A., F. Soltész, A. Nobes, U. Goswami, and D. Szucs. 2013. “Gender Differences in Developmental Dyscalculia Depend On Diagnostic Criteria.” *Learning and Instruction* (27): 31–39. <https://doi.org/10.1016/j.learninstruc.2013.02.004>
- Dirks, E., G. Spyer, E. C. D. M. Van Lieshout, and L. De Sonneville. 2008. “Prevalence of Combined Reading and Arithmetic Disabilities.” *Journal of*

- Learning Disabilities (41)*: 460–473.
<https://doi.org/10.1177/0022219408321128>
- Dowker, A. 2004. *What Works for Children with Mathematics Difficulties?* London: DfES. Retrieved on 4/9/2020 from <http://www.catchup.org/LinkClick.aspx?fileticket=59GXj0uNY1A%3d&tabid=105>
- Doyle, A. 2010. “Dyscalculia and Mathematical Difficulties: Implications for Transition to Higher Education in the Republic of Ireland.” Paper presented at the Disability Service Symposium, Trinity College Dublin.
<https://www.researchgate.net/publication/273137402>. (Accessed October 4, 2020).
- Emekene, Clement O. 2017. *Psychometric Analysis of the Advanced Progressive Matrices Scale for Use in Nigeria*. Ph.D Thesis, University of Port Harcourt.
- Fawcett, A. J. & Nicolson, R. I. 1996. *The Dyslexia Screening Test Manual*. London: UK; The Psychological Corporation: Harcourt Brace & Company.
- Feigenson L., S. Dehaene, and E. Spelke. 2004. “Core Systems of Number.” *Trends in Cognitive Sciences (8)*: 307–314.
<https://doi.org/10.1016/j.tics.2004.05.002>
- Gathercole, S. E., and S. J. Pickering. 2000. “Working Memory Deficits in Children with Low Achievements in the National Curriculum at 7 Years of Age.” *British Journal of Educational Psychology*, 70(2): 177–194.
<https://doi.org/10.1348/000709900158047>
- Geary, D. C. 1993. “Mathematical Disabilities: Cognitive, Neuropsychological, and Genetic Components.” *Psychological Bulletin*, 114: 345–362.
<https://doi.org/10.1037/0033-2909.114.2.345>
- Geary, D.C. 2004a. “Mathematics and Learning Disabilities.” *Developmental Neuropsychology*, 32(1): 471–519.
<https://doi.org/10.1080/87565640701360924>
- Geary, D. C. 2004b. “Mathematics and Learning Disabilities.” *Journal of Learning Disabilities*, 37(1): 4–15.
<https://doi.org/10.1177/00222194040370010201>
- Geary, D. C., C. O. Hamson, and M. K. Hoard. 2000. “Numerical and Arithmetical Cognition: A Longitudinal Study of Process and Concept Deficits in Children with Learning Disability.” *Journal of Experimental Child Psychology*, 77(3): 236–263. <https://doi.org/10.1006/jecp.2000.2561>
- Geary, D. C., and M. K. Hoard. 2001. “Numerical and Arithmetical Deficits in Learning-Disabled Pupils: Relation to Dyscalculia and Dyslexia.” *Aphasiology*, 15(7): 635–47. <https://doi.org/10.1080/02687040143000113>
- Gersten R., N. C. Jordan, and J. R. Flojo. 2005. “Early Identification and Interventions for Students with Mathematics Difficulties.” *Journal of Learning Disability*, 38: 293–304. Gliga, F., and T. Gliga. 2012. “Romanian

- Screening Instrument for Dyscalculia.” *Procedia Social and Behavioral Sciences*, 33: 15–19. <https://doi.org/10.1016/j.sbspro.2012.01.074>
- Grégoire, J., M. Noël, and C. Van Nieuwenhoven. 2004. “TEDI-MATH.” Antwerpen: Harcourt.
- Hein, J., M. Bzufka, and K. Neumärker. 2000. “The Specific Disorder of Arithmetic Skills. Prevalence Studies in a Rural and an Urban Population Sample and Their Clinical-Neuropsychological Validation.” *European Child & Adolescent Psychiatry* 9: S87–S101. <https://doi.org/10.1007/s007870070012>
- Hambleton, R. K. (2004). Theory, methods, and practices in testing for the 21st century. *Psicothema*, 16(4), 696–701.
- Heine, A., et al. 2013. “An Electrophysiological Investigation of Non-Symbolic Magnitude Processing: Numerical Distance Effects in Children with and Without Mathematical Learning Disabilities.” *Cortex*, 49: 2162–2177. <https://doi.org/10.1016/j.cortex.2012.11.009>
- Irwing, P., and D. J. Hughes. 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*, Vol. I (pp. 4–47). Hoboken, NJ: Wiley. <https://doi.org/10.1002/9781118489772.ch1>
- Iweka, F. (2014). *Comprehensive Guide to Test Constructions and Administration*. Rivers State: Chifas.
- Jordan, N. C., L. B. Hanich, and D. Kaplan. 2003. “A Longitudinal Study of Mathematical Competencies in Children with Specific Mathematics Difficulties Versus Children with Comorbid Mathematics and Reading Difficulties.” *Child Development*, 74(3): 834–850. <https://doi.org/10.1111/1467-8624.00571>
- Karagiannakis, G., A. Baccaglini-Frank, and Y. Papadatos. 2014. “Mathematical Learning Difficulties Subtypes Classification.” *Frontiers in Human Neuroscience* 8(57): 1–5. <https://doi.org/10.3389/fnhum.2014.00057>
- Kaufmann, L., and M. von Aster. 2012. “The Diagnosis and Management of Dyscalculia.” *Deutsches Ärzteblatt International*, 109(45): 767–778. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3514770/>. (Accessed July 30, 2020).
- Koontz, K. L., and D. B. Berch. 1996. “Identifying Simple Numerical Stimuli: Processing Inefficiencies Exhibited by Arithmetic Learning-Disabled Pupils.” *Mathematical Cognition*, 2(1): 1–23. <https://doi.org/10.1080/135467996387525>
- Koumoula, A., et al. 2004. “An Epidemiological Study of Number Processing and Mental Calculation in Greek School Children.” *Journal of Learning Disabilities*, 37: 377–388. <https://doi.org/10.1177/00222194040370050201>

- Kosc, L., 1974. “Developmental Dyscalculia.” *Journal of Learning Disabilities*, 7(3): 164–177. <https://doi.org/10.1177/002221947400700309>
- Kpolovie, P. J. 2010. *Advance Research Method*. New Owerri: Springfield Publishers Ltd.
- Kpolovie, P. J. 2011. *Statistical Techniques for Advanced Research*. New Owerri: Springfield Publishers Ltd.
- Kpolovie, P. J. 2012. *Education Reforms without Evaluation Designs*. Nigeria at Risk, New Owerri: Springfield Publishers Ltd.
- Kpolovie, P. J. 2014. *Tests, Measurement and Evaluation in Education*. Second Edition. New Owerri: Springfield Publishers Ltd.
- Kpolovie, P. J. 2016. *Excellent research methods*. Indiana: Partridge Publishing.
- Kpolovie, P. J., and C. O. Emekene. 2016. “Psychometric Advent of Advanced Progressive Matrices—Smart Version (Apm-Sv) for Use in Nigeria.” *European Journal of Statistics and Probability*, 4(3): 20–30
- Landerl, K., A. Bevan, and B. Butterworth. 2004. “Developmental Dyscalculia and Basic Numerical Capacities: A Study of 8–9-Year-Old Students.” *Cognition*, 93(2): 99–125. <https://doi.org/10.1016/j.cognition.2003.11.004>
- Landerl, K., and K. Moll. 2010. “Comorbidity of Learning Disorders: Prevalence and Familial Transmission.” *Journal of Child Psychol Psyc*, (51): 287–294. <https://doi.org/10.1111/j.1469-7610.2009.02164.x>
- Looi, C. Y., and R. C. Kadosh. 2010. *The Cambridge Encyclopaedia of Child Development*, 2nd edition (*In Press*). (Accessed October 4, 2020).
- Lyon, G. R., S. E. Shaywitz, and B. A. Shaywitz. 2003. “Defining Dyslexia, Comorbidity, Teachers’ Knowledge of Reading and Writing: A Definition Of Dyslexia.” *Annals of Dyslexia*, 53: 1–14. <https://doi.org/10.1007/s11881-003-0001-9>
- Mabbott, D. J., and J. Bisanz. 2008. “Computational Skills, Working Memory, and Conceptual Knowledge in Older Children with Mathematics Learning Disabilities.” *Journal of Learning Disabilities*, 41(1): 15–28. <https://doi.org/10.1177/0022219407311003>
- Mazzocco, M., and G. Myers. 2003. “Complexities in Identifying and Defining Mathematics Learning in the Primary School-Age Years.” *Annals of Dyslexia*, 53(1), 218–253. <https://doi.org/10.1007/s11881-003-0011-7>
- Moll, K., Kunze, S., Neuhoff, N., Bruder, J., & Schulte-Korne, G. 2014. “Specific learning disorder: Prevalence and gender differences.” *PLoS ONE*, 9(7), e103537. <https://doi.org/10.1371/journal.pone.0103537>
- Merdian, G., F. Merdian, and K. Schardt. 2012. *Bamberger Dyskalkuliediagnostik 1–4+ (R) (BADYS 1–4+)*. Bamberg: *PaePsy*.
- Merdian, G., Merdian, F., Schardt, K 2015. *Bamberger Dyskalkuliediagnostik 5–8+ (BADYS 5–8+)*. Bamberg: *PaePsy*.

- Metibemu, M. A. 2016. *Comparing Classical Test Theory and Item Response Theory in the Development, Scoring and Equating of Secondary School Physics Test in Ondo State, Nigeria*. Ph.D. International Centre for Educational Evaluation, Thesis University of Ibadan.
- Nfon, N. F. 2016. "A Survey of the Mathematical Problems (Dyscalculia) Confronting Primary School Pupils in Buea Municipality in the South West Region of Cameroon." *International Journal of Education and Research*, (4) 437–450.
- Ogbogo, S., and I. M. Opara. 2021. "Quantitative and Qualitative Analysis of Dyscalculia Test." *International Journal of Education, Learning and Development*, 9(7): 1–15.
- Ogbogo, S., and G. W. Orluwene. 2021. "Psychometric Analysis of Dyscalculia Test." *British Journal of Education*, 9(9): 55–70.
- Oku, K., and F. Iweka. 2018. "Development, Standardization and Application of Chemistry Achievement Test Using the One-Parameter Logistic Model (1-Plm) of Item Response Theory (IRT)." *American Journal of Educational Research*, 6(3): 238–257.
- Olkun, S., A. Altun, S. G. Sahin, and G. Kaya. 2016. "Psychometric Properties of a Screening Tool for Elementary School Student's Math Learning Disorder Risk." *International Journal of Learning, Teaching and Educational Research*, 15(12): 48–66.
- Opara, I. J. 2016. *Test Construction and Measurement: Concepts and Applications*. Owerri: Career Publishers.
- Orluwene, G. W. 2012. *Introduction to Test Theory and Development Process*. Port Harcourt: Chris-Ron Integrated Services.
- Orluwene, G.W., and C. Asiegbu. 2016. "Detecting Item Bias in Rivers State JSSCE Business Studies Using Item Response Theory (IRT) Approach." *European Journal of Statistics and Probability*, 5(2): 17–27.
- Passolunghi, M. C., and L. S. Siegel. 2004. Working Memory and Access to Numerical Information in Children with Disability in Mathematics. *Journal of Experimental Child Psychology*, 88(4): 348–358.
<https://doi.org/10.1016/j.jecp.2004.04.002>
- Pandey, S., and S. Agarwal. 2015. "Assessment of Learning Disability (Dyscalculia) among School Going Children: Today's Concern." *International Journal of Current Research*, 7(10): 22033–22035.
- Price, L. R. 2017. *Psychometric Methods: Theory into Practice*. New York: The Guilford Press.
- Reynold, C.R. 2014. *Test of General Reasoning Ability*. Lutz, FL: PAR
- Rosselli, M., E. Matute, N. Pinto, and A. Ardila. 2006. "Memory Abilities in Children with Subtypes of Dyscalculia." *Dev Neuropsychol*, 30(3): 801–818.
https://doi.org/10.1207/s15326942dn3003_3

- Schroeders and Schneider. 2008. *Test zur Diagnose von Dyskalkulie (TeDDy-PC)*. Göttingen: Hogrefe.
- Shalev, R. S., and V. Gross-Tsur. 2001. “Developmental Dyscalculia.” Review article, *Pediatric Neurology*, 24, 337–342. [https://doi.org/10.1016/S0887-8994\(00\)00258-7](https://doi.org/10.1016/S0887-8994(00)00258-7)
- Shalev, R., and M. von Aster. 2008. “Identification, Classification, and Prevalence of Development Dyscalculia.” *Encyclopedia of Language and Literacy Development*: published online. <https://doi.org/10.5167/uzh-12874>.
- Temple, C. M., and S. Sherwood. 2002. “Representation and Retrieval of Arithmetical Facts: Developmental Difficulties.” *Quarterly Journal of Experimental Psychology*, 55A(3): 733–52. <https://doi.org/10.1080/02724980143000550>
- Trott, C. 2010a. “Dyscalculia: A Practitioner’s View.” *Assessment & Development Matters*, 2(2): 19–21. <https://doi.org/10.53841/bpsadm.2010.2.2.19>
- Trott, C. 2010b. “Dyscalculia in Further and Higher Education.” *CETL MSOR Conference Proceedings* (pp. 68–73). University of Birmingham.
- Ubi, I. O. 2006. “Item Local Independence, Dimensionality and Trend Candidates’ Mathematics Performance on University Matriculation Examination in Nigeria.” Unpublished Ph.D Dissertation, University of Calabar, Nigeria.
- Van der Sluis, S., A. van der Leij, and P. F. de Jong. 2005. “Working Memory in Dutch Children with Reading- and Arithmetic-Related LD.” *Journal of Learning Disabilities*, 38(3): 21–9. <https://doi.org/10.1177/00222194050380030301>
- von Aster, M. 2000. “Developmental Cognitive Neuropsychology of Number Processing and Calculation: Varieties of Developmental Dyscalculia.” *European Child and Adolescent Psychiatry*, 9: 41–57. <https://doi.org/10.1007/s007870070008>
- von Aster, M. 2001. *Die neuropsychologische Testbatterie für Zahlenverarbeitung und Rechnen bei Kindern (ZAREKI)*. [The Neuropsychological Test Battery for Number Processing and Calculation in Children (NUCALC)]. Frankfurt, Germany: Swets & Zeitlinger.
- Von Aster, M. G., and R. S. Shalev. 2007. “Number Development and Developmental Dyscalculia.” *Developmental Medicine & Child Neurology*, 49: 868–873. <https://doi.org/10.1111/j.1469-8749.2007.00868.x>
- World Health Organization. 2010. International Statistical Classification of Diseases and Related Health Problems (ICD-10). <http://www.who.int/classifications/icd>. (Accessed October 5, 2020).

- Yang, D. C., and M. N. Li. 2003. "The Development of a Computerized Number Sense Scale for 3rd-graders." A number sense project supported by National Science Council, Taiwan.
- Yang, D., M. F. Li, and W. Li. 2008. "Development of A Computerized Number Sense Scale For 3rd Graders: Reliability and Validity Analysis." *International Electronic Journal of Mathematics Education*, 3(2): 111–124. <https://doi.org/10.29333/iejme/222>
- Zerafa, E. 2014. "Helping Children with Dyscalculia: A Teaching Programme with Three Primary School Children." *Procedia - Social and Behavioral Sciences*, (191): 1178–1182. <https://doi.org/10.1016/j.sbspro.2015.04.516>
- Zygouris, N., et al. 2017. "A Neuropsychological Approach of Developmental Dyscalculia and a Screening Test Via a Web Application." *International Journal of E. P.*, 7(4): 51–65. <https://doi.org/10.3991/ijep.v7i4.7434>

Appendix

DYSCALCULIA TEST (DYSCALT)

STELLA ETENG-UKET

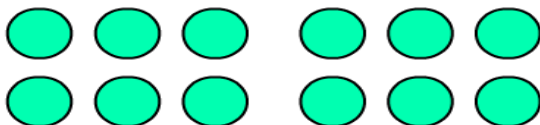
Item booklet

In total, this test will take you approximately **45 minutes** to complete. Each of the three sections is timed separately (approximately **15 minutes each**).

You will need a pencil for this test. Please look at your answer sheet now. The box on the top of the page should be filled in.

For each section, wait until the examiner tells you to begin before turning to the first page of that section. Within each section, you will be asked to pick the best answer to each question. Please **do not mark** your answers directly on the booklet pages. Instead, record your responses on the answer sheet provided.

Look at the sample item below. Identify the number of circles from the options supplied below.



- A. 12 B. 13 C.16 D.15

The best answer for this item is **A**, for this item, shade the letter **A** on your answer sheet as pictured below.

Example	<input checked="" type="checkbox"/> A	<input type="checkbox"/> B	<input type="checkbox"/> C	<input type="checkbox"/> D
----------------	--	----------------------------	----------------------------	----------------------------

If you make an error when marking your answer sheet, completely erase the mistake and shade the option for the correct answer.

For each section, begin with the first item in that section. If a question is too difficult or you need more time to think about your answer, you may skip ahead to the next question. There is no penalty for guessing on this test. You may return to skipped items within each section if there is time remaining for that section. You may use scratch paper, but not a calculator or the internet. After your time for each section has ended, the examiner will tell you to stop. If you finish a section before the time

has ended, please stop at the stop sign and tell the examiner that you have finished with that section.

You have completed the instructions. Please wait for the examiner to tell you to continue.


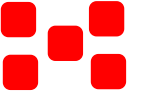


Sample items from the Dyscalculia Test

Number sense Domain (Items here assess number sense which is an intuitive feel for numbers and a common sense approach to using them).

Sample items from the Substising and numerosity Subset (This subset assesses the ability to instantaneously recognize the number of objects in a small group without the need to count them, i.e., estimating a small number of objects and the ability to sequence numbers from small to large accurately).

Example 1

1. Which of the boxes below have the highest numbers?

Box 1	
Box 2	
Box 3	
Box 4	

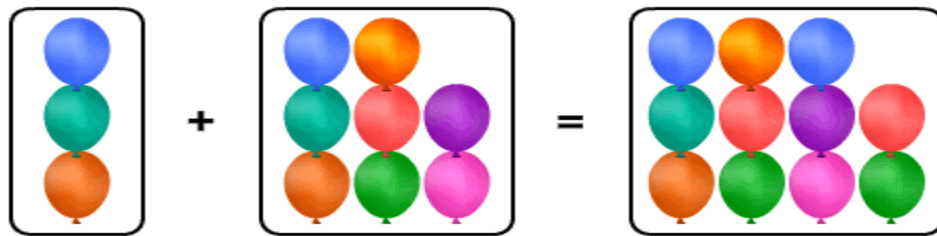
A Box 1 B Box 2 C. Box 3. D. Box 4

Sample items from the Addition operations and computation Subset of the Dycalculia Test

Addition operations and computation Subset- assess an individual's ability to perform addition task fluently, i.e., with accuracy, *automaticity*, and *flexibility*.

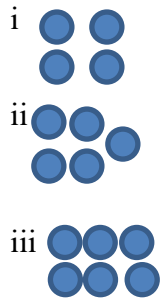
Example 2

Choose the addition sentence which the model shows.



- A. $9+5=14$ B. $9+3=12$ C. $3+8=11$ D. $10+5=15$

Example 3



The dots presented as i, ii, and iii above will give a sum when added in the following order from the options below:

- A. 4, 5, 15 B. 5, 4, 6 C. 4, 6, 5 D. 4, 5, 6