Journal of Education and Practice
ISSN 2222-1735 (Paper) ISSN 2222-288X (Online)
Vol.9, No.32, 2018
www.iiste.org

Parameterization of Teacher-made Physics Achievement Test Using Deterministic-Input-Noisy-and-Gate (DINA) Model

Rose C. Anamezie (Ph.D) Fidelis O. Nnadi (Ph.D)
Department of Science & Computer Education,
Enugu State University of Science and Technology, Enugu-Nigeria

Abstract

Traditional methods of test parameterization have been found defective in terms of assuming one score and not providing information on skills mastery profile of the examinees, in addition to non-estimation of the fourth parameter- slipping parameter of test items. Cognitive diagnostic modeling (CDM), specifically deterministic-input-noisy-and-gate (DINA) model is a robust procedure for measuring the skills mastery profile of the examinees, when there is no attribute hierarchy. DINA model also estimates four item parameters: discriminating, difficult, guessing and slipping which hitherto are not implemented as a whole in traditional techniques including the item response theory (IRT). Physics was chosen because it is commingled with civilization. Also, secondary school Physics teachers predominantly use teacher-made tests in their internal examinations in Nigeria. Such instrument cannot be used to precipitate the true ability levels of the students for purposes of promotion. The data collected plus the harmonized Q-matrix were analyzed using CDM package in r software version 3.4.3 via R-Studio version 1.0.153. The results showed that: (i) eighteen items in the test fit the DINA model while thirty-two items had misfit. (ii) nine attributes were mastered by the examinees while three attributes were not mastered. It was recommended that schools' item banks should be created by the governments for uploading psychometrically robust test items for teachers' use.

Keywords: Physics, teacher-made test, DINA model and parameterization.

Introduction

The cognitive diagnostic model (CDM) is an off-shoot of the item response theory (IRT). However, there are areas of commonalities and divergences between the two theories as far as parameterization of dichotomously scored test items are concerned. Parameterization is a term that is used to represent the statistical quality of test items. In this study the statistical qualities studied included item difficulty index (a-parameter), discriminating index (b-parameter), guessing index (c-parameter), slipping index (d-parameter) and skills mastery profile. Dichotomous items are scored right or wrong for the presence or absence of an attribute. Usually, multiple choice objective items are scored dichotomously. This form of test for an item contains one correct answer, the key and three or four incorrect answers, the distractors. The use of teacher-made objective tests in classroom assessment is a common practice in Nigeria. This kind of test is unstandardized and prone to errors including defective measurement of student's ability.

IRT is anchored on three assumptions: uni-dimensionality, local independence and monotonicity. The uni-dimensional assumption upholds that a test contains a single latent trait, often referred to as ability of the test takers. The IRT is robust in latent traits modeling. This is because of its ability to model interrelated and polytomously-scored manifest and latent variables. Furthermore, with recent advances in IRT, a test can cover few latent traits under multivariate item response theory (MIRT) framework. A purely uni-dimensional test is advantageous in placing the test takers in an ability consortium for purposes of selection.

Second assumption of IRT is local independence. This assumption for example says that the probability of an examinee in an objective test to score item A correctly does not influence the examinee's correct or incorrect choice of item B and vice versa. In other words, the probability of an examinee scoring an item correctly does not predetermine the examinee's ability to score correctly /incorrectly in any item.

The third assumption of IRT is monotonicity. This implies that the test takers ability increases with probability of success in an item, keeping constant the item difficulty level. The monotonicity assumption of CDM shows that the probability that the examinees do not slip has to be greater than the guessing probability (i.e $1\text{-}S>g$) (Choi, Young-Sun & Yoon, 2015).

The CDM framework, just like MIRT is a confirmatory discrete latent class model developed for identifying the presence or absence of multiple cognitive skills (Rupp & Templin,2008). While MIRT is robust for few ability measurement, CDM is robust for measuring multiplicities of abilities (Zhimei, 2011). Zhimei also reported that CDMs are used to analyze test response data in such a way that multivariate classifications of examinees can be achieved based on their latent skills mastery profile. A major point of divergence of CDM from IRT is the estimation of skill mastery profile of the examinees. In the parlance of CDM, the skills are scored either zero (0) or one (1) for the absence or presence of already existing skills in the examinees respectively. Furthermore, unlike IRT which is a latent trait or continuous class model, the CDM is a class of discrete manifest or latent variable model (George & Robitzsch, 2015). Being a discrete manifest/latent variable model, it implies that the CDM is only robust for dichotomously-scored items that do not show meaningful inter-correlation. Also, the discrete nature of the model implies that it traces back an examinees answer based on the skills possessed. Other points of divergences of CDM and IRT are as follows. The IRT is data driven while CDM is not data driven (Zhimei, 2011). CDM uses Markov Chain Monte Carlo (MCMC) to estimate its posterior results at a given iteration. This implies that the error in parameter estimation under the CDM framework is lower than that of IRT when the method of parameter estimation is not Bayesian.

The deterministic-input-noisy-and-gate (DINA) model is one of the sub-models in CDM. The DINA model is non-compensatory and that the examinee's probability of responding to an item correctly increases if the examinee possesses all the required skills or attributes for a particular item (George & Robitzsch, 2015). A student j is considered to have mastered a skill k, if a student's estimated probability for that particular skill is greater than or equal to .5 (Zhimei, 2011). The non-compensatory quality of DINA implies that the examinee cannot compensate for a lack in one skill foe a surplus in another skill. Moreover, there are situations where the examinee can guess the correct option of an objective test without possessing all the required skills. This is called guessing parameter represented by the letter 'g'. Also, an examinee who possesses all the required skills can mistakenly pick the wrong option, leading to what is called slipping/carelessness parameter represented by the letter 's' in CDM. Item parameterization in DINA model are basically achieved using guessing and slipping parameters. However, both difficult and discriminating parameters can be derived from DINA model following George et al's recommendations. The difficult parameter is a measure of the uneasiness of an item while the item discriminating parameter correspond to the difference in the probability of a correct response for a student who has mastered all measured skills and a student who has mastered at least one of the measured skills (Zhimei, 2011). A good test item should not be too easy either too difficult. This implies that the proportion of correct response for a fairly difficult item should not be zero (too difficult) either one (too easy). A range of .30 to .75 is used as the benchmark for this study. Also a good item should be able to highly discriminate between those who possess the skills needed to correctly score an item and those examinees who do not possess the required skills for correct scoring of an item. For the DINA model, de la Torre (2008) reported that the item discriminating is equal to $(1\text{-}S_i)\text{-}g_i$, where S_i represents slipping parameter and g_i represents guessing parameter for item $_i$. Zhimei reported that any good dichotomously scored item should have low (< .5) slipping and guessing parameters to differentiate between students who have mastered all skills from students who have mastered fewer skills (high discrimination).

Effective implementation of CDM is hinged on the skills or attributes profile determined by content experts (Yuling, 2012). The skills or attributes profile correspond to binary matrix of the skills or attributes, with 1 or 0 showing the presence or absence of the desired skill or attribute respectively. In the DINA model, the number of maximum attribute profile is equal to 2^k, where k is the number of diagnosed attributes for a given assessment (Zang, 2015). However, the knowledge of content experts in addition to specification of skills or attributes in the test is needed to verify the existence of nested skills or attributes within the test. The reason is because the procedure for specification of skills or attributes (technically called Q-matrix) of nested skills or attributes is different from the procedure for specification of Q-matrix for discrete skills or attributes. Su (2013), and Leighton, Gierl and Hunka (2004) in their separate studies reported that estimation error would occur if the conventional DINA was used in parameterization of test with nested structures.

The need for continuous assessment of learners heralded the birth of CDM. Continuous assessment posits that learners should be constantly evaluated and feedback given to them, as a way of improving the quality of instructions in schools. The idea of continuous assessment was introduced in Nigeria in 1977 following the adoption of National Policy on Education (Omebe, 2014). The Federal Government of Nigeria through the Federal Ministry of Education's National Policy on Education (2013) reiterated the need for quality assurance measures through the establishment of a standardized assessment system for monitoring and reporting learning achievement. The above policy statement is in keeping with the need of continuous assessment of the learners through feedback mechanisms. CDM is a robust framework for providing feedback on the skills mastery profiles of the examinees in addition to its use for parameterization or item analysis.

Some empirical investigations have been carried out on the use of DINA model for parameterization of test items. Instances are drawn from objective test items. Zhimei (2011) investigated the influence of maximizing the potential of multiple-choice items for cognitive diagnostic assessment. The purpose of the study included the determination of the students' profile estimates using the DINA model for 20000 grade 6 students in Ontario in 2016. The instrument was a twenty eight four option multiple choice Mathematics items with four binary non-nested skills. The benchmark for the estimated probability of skills mastery was p > .5 whereas the slipping and guessing indices for ideal item should be less than .5. The result of the study indicated that 24 in 28 items were mastered by the examinees while four items were not mastered by the examinees. 9 in 24 items were dropped on the basis of having high guessing and slipping parameters. In a study that was carried out by Kyong, Young-Sun and Yoon (2015) using the DINA model to determine 43 Mathematics items mastery of US and Korean students, 81.8% of American and 80.3% of Korean students answered an item correctly. 28 in 42 items of the

Mathematics test were not mastered by the American students while 12 in 43 items were not mastered by the Korean sub-sample. Zhang (2015) analyzed hierarchical data with the DINA-HC approach. Hierarchical convergent Q-matrix with six skills was used in a twenty-eight item English proficiency objective test, using CDMs in Mplus software. The result indicated that 5 in 28 items had guessing and slipping parameters within normal range (≤ .5) for DINA model. 9 in 28 items fit the DINA model in terms of guessing and slipping parameters. For the DINA-HL (DINA hierarchical linear) model, 8 in 28 items fit the model interms of guessing and slipping parameters. Yu-Lan, Kyong, Won-Cheng, Taehoon and Malissa (2013) studied hierarchical cognitive diagnostic analysis using simulation. The aim of the study included the determination of guessing and slipping parameters of Mathematics test using the conventional DINA and DINA with hierarchical configurations. The result indicated that 18 in 29 items fit the model for guessing and slipping parameters.

There is the need to apply the DINA model to calibrate Physics achievement test following the importance of Physics to cultural development of societies. Physics is commingled with man's culture. The practice of Physics principles is seen among those that have not received a formalized instruction in Physics. In a typical African setting, for example pieces of firewood are ignited and used as a source of heat energy for cooking. The chemical energy in the firewood is converted to heat by the cook through the process of oxidation. Through the processes of observation and imitation, outside of a formal classroom setting mankind interacts with the elements in the physical environment to solve his/her problems. Apart from roasting food which only uses air as a medium, an informally educated person uses water as a medium to cook food using a vessel. There is the understanding among the class of informally educated people that the amount of heat energy supplied by the burning fuel to the water in the vessel is absorbed by it to do the cooking. The cook is also conscious of the time of cooking and the quantity of heat determined by the volume and the type of wood, lest the food melts or turns to carbon black. This further authenticates that an informally educated person is aware of the concept of melting. Suppose that the cooked food is yam and it is the wish of cook to turn it to a paste, then heavy force is applied on the pestle to crush the yam granules in a mortar in a process called pounding. There is a time during pounding when the pestle makes an angular motion with its associated speed inside the mortar. The essence of the practice is to make finer the yam particles to encourage cohesion between similar molecules of yam. With bare hands, the finely grained usually hot yam paste is rolled into ball-shapes and stored in an earthen vessel or guard with cover to preserve the temperature of the yam paste for a reasonable length of time. From the fore-going, it can be deduced that before any formal Physics instruction takes place, the learners' mindsets are already pre-occupied with some Physics concepts learnt informally. The governments in realization of the need to further develop Physics-related skills among the populace to enhance economic, social and political growth of individuals and nations. To this end,

Purpose of the Study

The study intended to: (i) determine the items of the Physics achievement test that fit the DINA model (ii) determine the skills mastery profile of the examinees in the test

Research Questions

Two research questions guided the study. They included: (i) What are the items of the Physics achievement test that fit the DINA model? (ii). What are the skills mastery profile of the examinees in the test?

Research Method

Ex-post facto evaluation design was used for the study because the attributes existed in the students prior to the study and the researchers did not manipulate them. The population for the study was five hundred and thirteen senior secondary three Physics students in eleven public secondary schools in Nkanu West Local Government Area of Enugu state (Ministry of Education Enugu, 2015). Accidental sampling procedure was used to sample 87 SS3 Physics students from the area. The instrument used for data collection was a four-option Physics achievement test (PAT) developed by the researchers. It had originally 55 items. PAT items were drawn from the six broad areas of the Nigerian Physics curriculum. PAT was subjected to face and content validation by experts in Physics Education from the Department of Science and Computer Education, Enugu State University of Science and Technology. Content validation was achieved using test blueprint. Only 50 items of PAT survived face validation.

The content experts independently produced binary Q-matrix for the study. Areas of divergences in the Q-matrix specifications by the content experts were sorted out in the course of producing the final Q-matrix that was used with data to determine the item parameters. Table1 shows the Physics attributes and the number of items that measured each attribute. Table 2 shows the harmonized Q-matrix for the test. From Table 2, each of the fifty items measured only one of the twelve attributes. The attributes were not nested. PAT was administered to the respondents by their regular Physics teachers with a return rate of 87/103. The data collected were analyzed with the harmonized Q-matrix using CDM package in r software version 3.4.3. The reliability coefficient of the items of PAT that survived the DINA calibration process was established using Kuder-Richardson's formula-21. An overall coefficient of .67 was obtained. The differential item functioning DIF (based on gender), specifically Mantel Hanszel's method (difMH) in r 3.51 was conducted on the items of the test after the DINA model calibration. The result showed that all the items that fit the model did not show significant DIF based on gender. For the decision rule concerning the acceptability of the item parameters, items whose discriminating parameter is approximately greater than or equal to .60 is flagged as a good item. Difficulty parameter value range between .35 to .75 is flagged as a good item. Guessing and slipping parameter values that are less than .5 are flagged as good items in each case (Zhang, 2015). The estimated probability of skill mastery was accepted when its value was above .5 (adapted from Zhimei, [2011] by the researchers). The reason for the adaptation of rule for the cut-off value of skill mastery was because any skill mastery should be above average (50%) following rule of the thumb.

Table 1: Physics Attributes

Attribute	**Items**	**Number**
1.Conceptual definitions/explanations	1,4,6,8,13,18,20,21,26,28,30, 32,33,34,35,36,37,38,39,40,41,45,46,47,48,49 and 50	27
2.SI unit of physical quantities	2	1
3.Problem solving involving density and upthrust	3,5	2
4.Problem solving involving uniformly accelerated motion	7,10,11,15,16,17	6
5. Problem solving involving coplanar forces	12	1
6. Problem solving involving circular motion/angular momentum	14	1
7. Problem solving involving thermometry and heat exchange	19,24,27	3
8. Problem solving involving expansion	22,23	2
9. Problem solving involving wave motion	29	1
10. Problem solving involving focal length	31	1
11. Problem solving involving electric circuits	41,43,44	3
12. Graphical interpretation	9,25	2
Total		50

Table2: The harmonized Q-matrix for the Physics achievement Test

	Attribute											
Item	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
1	1	0	0	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0	0	0
4	1	0	0	0	0	0	0	0	0	0	0	0
5	0	0	1	0	0	0	0	0	0	0	0	0
6	1	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	1	0	0	0	0	0	0	0	0
8	1	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	0	0	1
10	0	0	0	1	0	0	0	0	0	0	0	0
11	0	0	0	1	0	0	0	0	0	0	0	0
12	0	0	0	0	1	0	0	0	0	0	0	0
13	1	0	0	0	0	0	0	0	0	0	0	0
14	0	0	0	0	0	1	0	0	0	0	0	0
15	0	0	0	1	0	0	0	0	0	0	0	0
16	0	0	0	1	0	0	0	0	0	0	0	0
17	0	0	0	1	0	0	0	0	0	0	0	0
18	1	0	0	0	0	0	0	0	0	0	0	0
19	0	0	0	0	0	0	1	0	0	0	0	0
20	1	0	0	0	0	0	0	0	0	0	0	0
21	1	0	0	0	0	0	0	0	0	0	0	0
22	0	0	0	0	0	0	0	1	0	0	0	0
23	0	0	0	0	0	0	0	1	0	0	0	0
24	0	0	0	0	0	0	1	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0	0	1
25	1	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	1	0	0	0	0	0
28	1	0	0	0	0	0	0	0	0	0	0	0
29	0	0	0	0	0	0	0	0	1	0	0	0
30	1	0	0	0	0	0	0	0	0	0	0	0
31	0	0	0	0	0	0	0	0	0	1	0	0
32	1	0	0	0	0	0	0	0	0	0	0	0
33	1	0	0	0	0	0	0	0	0	0	0	0
34	1	0	0	0	0	0	0	0	0	0	0	0
35	1	0	0	0	0	0	0	0	0	0	0	0
36	1	0	0	0	0	0	0	0	0	0	0	0
37	1	0	0	0	0	0	0	0	0	0	0	0
38	1	0	0	0	0	0	0	0	0	0	0	0
39	1	0	0	0	0	0	0	0	0	0	0	0
40	1	0	0	0	0	0	0	0	0	0	0	0
41	1	0	0	0	0	0	0	0	0	0	0	0
42	0	0	0	0	0	0	0	0	0	0	1	0
43	0	0	0	0	0	0	0	0	0	0	1	0
44	0	0	0	0	0	0	0	0	0	0	1	0
45	1	0	0	0	0	0	0	0	0	0	0	0
46	1	0	0	0	0	0	0	0	0	0	0	0
47	1	0	0	0	0	0	0	0	0	0	0	0
48	1	0	0	0	0	0	0	0	0	0	0	0
49	1	0	0	0	0	0	0	0	0	0	0	0
50	1	0	0	0	0	0	0	0	0	0	0	0

Results

The results are presented according to the research questions that guided the study.
Research question 1 (RQ1) sought information on the items of the Physics achievement test that fit the DINA model. Table 3 was used to answer RQ1.

Table 3: What are the items of the Physics achievement test that fit the DINA model?

Item	Discriminating	Difficulty	Guessing	Slipping	Remarks
1	-0.076576577	0.20495495	0.8333333	2.432432e-01 *	Restructure
2	0.999999714*	0.49999986 *	2.859958e-07 *	1.366428e-16 *	Fair
3	-0.003743316	0.22540107	7.764706e-01	2.272727e-01*	Restructure
4	0.256756757	0.37162162 *	5.000000e-01	2.432432e-01*	Restructure
5	1.000000000 *	0.50000000*	1.836115e-15 *	3.027881e-13*	Fair
6	0.214714715	0.83708709	5.555556e-02 *	7.297297e-01	Restructure
7	0.583725029 *	0.30273208*	4.054054e-01 *	1.086957e-02 *	Fair
8	0.402402402	0.35435435 *	4.444444e-01 *	1.531532e-01*	Restructure
9	0.112244898	0.05612245	8.877551e-01	0.000000e+00*	Restructure
10	0.415981199	0.25146886	5.405405e-01	4.347826e-02*	Restructure
11	1.000000000 *	0.50000000*	1.902088e-52 *	1.322088e-11*	Fair
12	0.999905564*	0.50004721*	1.028840e-08 *	9.442533e-05*	Fair
13	0.109609610	0.33408408*	6.111111e-01	2.792793e-01*	Restructure
14	1.000000000 *	0.50000000*	5.537146e-12 *	2.577253e-13*	Fair
15	0.046122209	0.49045241 *	4.864865e-01 *	4.673913e-01*	Restructure
16	0.117508813	0.31962397*	6.216216e-01	2.608696e-01*	Restructure
17	0.176263220	0.47943596*	4.324324e-01 *	3.913043e-01 *	Restructure
18	0.747747748*	0.62612613*	0.000000e+00 *	2.522523e-01 *	Fair
19	0.153174603	0.61230159*	3.111111e-01 *	5.357143e-01	Restructure
20	0.397897898	0.52327327*	2.777778e-01 *	3.243243e-01 *	Restructure
21	0.774774775 *	0.61261261*	0.000000e+00 *	2.252252e-01 *	Fair
22	0.999782234 *	0.50010888 *	6.512387e-12 *	2.177657e-04 *	Fair
23	0.046705348	0.39538791*	5.812594e-01	3.720352e-01 *	Restructure
24	1.000000000*	0.50000000*	7.398735e-18 *	0.000000e+00 *	Fair
25	1.000000000 *	0.50000000*	2.248564e-25 *	2.995740e-13*	Fair
26	0.070570571	0.24249249	7.222222e-01	2.072072e-01 *	Restructure
27	0.231746032	0.30634921*	5.777778e-01	1.904762e-01 *	Restructure
28	0.067567568	0.13288288	8.333333e-01	9.909910e-02 *	Restructure
29	1.000000000 *	0.50000000 *	4.816565e-18 *	0.000000e+00 *	Fair
30	0.468468468	0.76576577	0.000000e+00 *	5.315315e-01	Restructure
31	1.000000000 *	0.50000000*	2.157223e-16 *	0.000000e+00 *	Fair
32	0.764264264 *	0.56231231*	5.555556e-02 *	1.801802e-01 *	Fair
33	0.328828829	0.66891892*	1.666667e-01 *	5.045045e-01	Restructure
34	0.403903904	0.74249249*	5.555556e-02 *	5.405405e-01	Restructure
35	0.602102102 *	0.64339339*	5.555556e-02 *	3.423423e-01 *	Fair
36	0.348348348	0.38138138 *	4.444444e-01 *	2.072072e-01 *	Restructure
37	0.409909910	0.29504505*	5.000000e-01	9.009009e-02 *	Restructure
38	0.003003003	0.44294294*	5.555556e-01	4.414414e-01 *	Restructure
39	0.357357357	0.37687688*	4.244444e-01 *	1.981982e-01 *	Restructure
40	-0.177177177	0.83708709	8.888889e-01	2.882883e-01 *	Restructure
41	0.214714715	0.55126081*	5.555556e-02	7.297297e-01	Restructure
42	0.559322838*	0.32112346*	1.690778e-01*	2.715994e-01 *	Fair
43	0.863416686 *	0.56829166 *	4.265840e-29*	1.365833e-01 *	Fair
44	0.505225618 *	0.25261281*	4.947744e-01	0.000000e+00 *	Restructure
45	0.283783784	0.35810811*	5.000000e-01	2.162162e-01 *	Restructure
46	0.837837838 *	0.58108108*	0.000000e+00*	1.621622e-01 *	Fair
47	0.261261261	0.86936937	0.000000e+00 *	7.387387e-01	Restructure
48	0.054054054	0.30630631*	6.666667e-01	2.792793e-01 *	Restructure
49	0.818318318 *	0.53528529*	5.555556e-02	1.261261e-01 *	Restructure
50	0.150150150	0.48048048*	4.444444e-01*	4.054054e-01*	Restructure

KEY: * implies accepted within the column

From Table 3, the items which are flagged with asterisks are the accepted items within each parameter estimate. An item which bears an asterisk for each and every four parameters: difficulty, discriminating, guessing and slipping fits the DINA model and therefore regarded as fair in Table 3. Eighteen items namely: 2, 5, 7, 11, 12, 14, 18, 21, 22, 24, 25, 29, 31, 32, 35, 42, 43 and 46 fit the DINA model whereas thirty-two items had a misfit and labeled " Restructure" in the Table3. The graphical plot of the guessing and non-slipping probabilities of the fifty items is also shown below. From Figure 1, one minus the value of non-slipping probability for each item gives the value of the slipping probability estimated in Table 1. Appendix A shows the r codes which yielded the results in Tables1-2 and figures1-2.

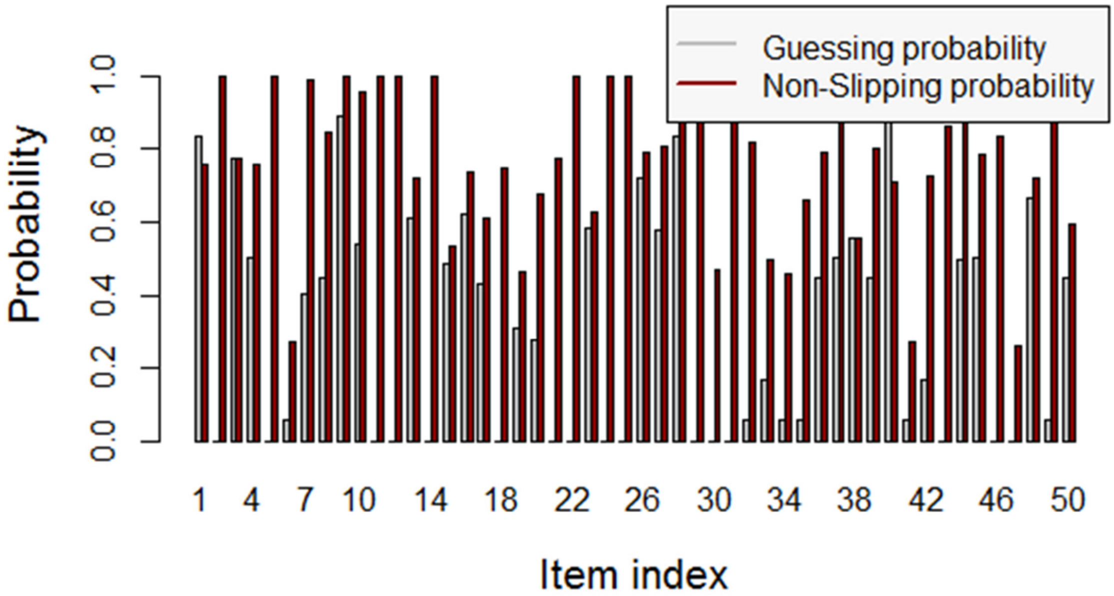

Figure 1: Guessing and Non-slipping probability plots

Research question 2 (RQ2) sought information on the skills mastery profile of the examinees in the test. The data presented in Table 4 were used to answer RQ2.

Table 4: Attribute Mastery probability

Attribute	Mastery probability	Remarks
1	0.8604651	Mastered
2	0.8062015	Mastered
3	0.3410853	Not mastered
4	0.7131783	Mastered
5	0.3488701	Not mastered
6	0.6279070	Mastered
7	0.6511628	Mastered
8	0.6668119	Mastered
9	0.8914729	Mastered
10	0.8062016	Mastered
11	0.8619083	Mastered
12	0.2403101	Not mastered

From Table 4, eight attributes namely: 1, 2, 4, 6, 7, 8, 9, 10 and 11 were mastered by the examinees whereas three attributes: 3, 5 and 12 were not mastered by the examinees. The proportion of attribute mastery in the test was 9/12 ~ .75. The plot of the skill mastery is shown in figure 1 below. The graphical plot of the skills mastery probabilities provides visual interpretation of the examinees' skills mastery profile.

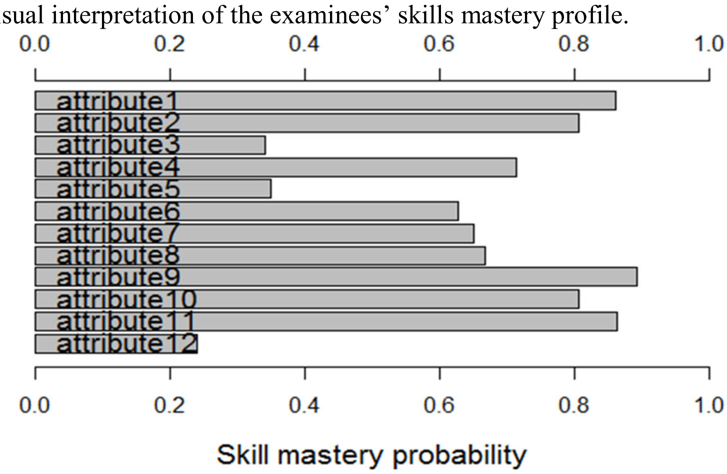

Figure 2: Skill mastery probability

Discussion of Findings

The result of research question one indicated that eighteen in fifty items fit the DINA model. The procedure adopted in extracting the model's parameters is in tandem with earlier procedures adopted by Zhimei (2011) and, Kyong,Yong-Sun and Yoon (2015). The researchers in their independent studies blended non-hierarchical Q-matrix with data to produce the parameter estimates. Only two parameters: guessing and slipping parameters were extracted in earlier studies. In the present study, George and Robitzsch's (2015) suggested equations for difficulty and discrimination parameters were adopted to extract the parameters. However, the procedures adopted by Yu-Lan, Kyong, Won-Cheng, Taehoon and Malissa (2013) and Zang (2015) in extracting the DINA model's parameters were at variance to the procedure adopted in the present study. This is because the Q-matrices in those studies were hierarchical.

The result of research question two indicted that nine attributes were mastered by the examinees whereas three attributes were yet to be mastered. The attributes that were yet to be mastered included problem solving involving uniformly accelerated motion, coplanar forces and graphical interpretation. This result has serious implications for the Physics teachers and students. This is because they need to improve on the skills needed in those attributes to ensure their mastery and improvement in Physics achievement. This is indeed a major benefit of the cognitive diagnostic modeling, CDM in providing immediate fine-grained feedback both at the item and person levels. Although, the other nine attributes were classified as having been mastered by the examinees, they vary in their levels of mastery. For example, problem solving involving uniformly accelerated motion, circular motion and angular momentum, thermometry and heat exchange and expansion were below 80% mastery. There is need for a concerted effort for their improvement in their mastery by the examinees. Other attributes that have mastery levels from 80% and above but not reaching 100% included conceptual definitions and explanations, units of physical quantities, problem solving involving wave motion, focal length and electric circuits need slight improvements in skill mastery level as well.

Conclusion

The result of the study showed that eighteen items of the Physics achievement test fit the DINA model. The items were not biased against the examinees' gender. Three attributes were not mastered by the examinees. It was recommended that item banks should be created by the government where psychometrically robust test items can be uploaded and accessed based on subjects by the teachers for students' testing purposes. This is a way of reducing to its barest minimum teacher-made tests from being used in classrooms for testing.

References

Choi, K.M.,Young-Sun, L. & Yoon, S. P. (2015).What CDM can tell about what students have learned: An analysis of TIMSS Eight grade Mathematics. *Eurasia Journal of Mathematics, Science and Technology Education*, 11(6), 1563-1577.

De la Torre, J. (2008). An empirically based method of Q-matrix validation for the DINA model: Development and application. *Journal of Educational Measurement*, 45, 343-362.

Federal Government of Nigeria, (2013*). National Policy on Education and a four–year strategic plan for the development of the education sector*. Abuja: Government press

George, A.C.& Robitzsch, A. (2015). Cognitive diagnosis in R: A didactic. *The Qualitative Methods for Psychology*, 11(3), 189-295

Leighton, J.P., Gierl, M.J., & Hunka, S.M. (2004). The attribute hierarchy method for cognitive assessment: A variation of Tatsuoka's rule space approach. *Journal of Educational Measurement*, 41, 205-237

Omebe, C.A. (2014). Continuous Assessment in Nigeria: Issues and Challenges. *International Journal of Humanities, Arts, Medicine and Sciences,*2(9), 89-96

R core Team (2015). *R: A language and environment for statistical computing*. ISBN 3-9000051-07-0. R Foundation for statistical computing, Vienna, Austria. http://www.R-project.org

Rash, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen Denmark: Danmarks pedagogic Institute

Rupp, A. A & Templin, J. (2008). Unique characteristics of diagnostic models: A review of the current state-of–art. *Measurement*, 6,219-262

Su, Y. (2013). Cognitive diagnostic analysis using hierarchically structured skills. *Unpublished Ph.D Thesis*, University of Iowa:

Yu-Lan, S., Kyong,M.C.,Won-Chan, L. & Melissa, M. (2013). Hierarchical cognitive diagnostic analysis for TIMSS 2003 Mathematics. *Center for advanced studies in measurement and evaluation, CASMA Research report*, Number 35

Yuling, F. (2012). Estimation and Q-matrix validation for diagnostic classification models. *Unpublished Ph.D Thesis*, College of Arts and Sciences, University of South Carolina.

Zang, J. (2015). Analyzing hierarchical data with the DINA-HC approach. *Unpublished Ph.D Thesis*, The

Graduate school of Arts and Sciences, Columbia University
Zhimei, G. (2011). Maximizing the potential of multiple choice items for cognitive diagnostic assessment. *Unpublished Ph.D Thesis*, Department of Human Development and Applied Psychology, University of Toronto.

Appendix A: R codes for this study's DINA model implementation

```
#import sample 2 and qmatrix3 with 12 attributes into R
library("CDM") #run
Dinasample2<-read.delim(file.choose(),header=FALSE) # load data and run
Dinaqmatrix3<-read.delim(file.choose(),header=FALSE) # load data and run
colnames(Dinasample2)=c("item1","item2","item3","item4","item5","item6",
                "item7","item8","item9","item10","item11","item12",
                "item13","item14","item15","item16","item17","item18",
                "item19","item20","item21","item22","item23","item24",
                "item25","item26","item27","item28","item29","item30",
                "item31","item32","item33","item34","item35","item36",
                "item37","item38","item39","item40","item41","item42",
                "item43","item44","item45","item46","item47","item48",
                "item49","item50") #run
colnames(Dinaqmatrix3)=c("attribute1","attribute2","attribute3",
                 "attribute4","attribute5","attribute6",
                 "attribute7","attribute8","attribute9",
                 "attribute10","attribute11","attribute12") #run
DM<-din(Dinasample2,Dinaqmatrix3) #run the DINA model
pvalues<-colMeans(Dinasample2,na.rm = TRUE)  #run p values
DM$guess  #run guessing index
DM$slip #run slipping index
DM$IDI #run item discriminating index(confirms George & Robitzsch (2015) Discrim equation.
Discrim<- 1-DM$guess$est-DM$slip$est # run item discrimination index
Easiness<-(DM$guess$est+(1-DM$slip$est))/2 #run item easiness
Difficulty<-1-Easiness # run item difficulty
DM$skill.patt  #run probability for skills
DM$attribute.patt  # run skill class probability
skill.cor(DM)
fit.DM<-IRT.modelfit(DM)
fit.DM<-DM$itemfit.rmsea
DM$mean.rmsea  #skill class prob plot
DM$coef
DM$subj.pattern
DM$posterior
DM$model.type
fit.DM
coef(DM)  #extracts estimated parameters
```

Sis, publish in Eurasia Journal of Mathematics, Science and Technology Education