**Appendix 1**: *2018 UTME Blueprint*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Domain** | **Topic** | **Remembering** | **Understanding** | **Applying** | **Analyzing** | **Evaluating** | **Total Items** | **Item %** |
| Mechanics | Kinematics and Dynamics | 1 (E) | 0 | 1 (M) | 0 | 0 |  |  |
| Work, Energy, and Power | 0 | 0 | 1 (M) | 0 | 0 |  |  |
| Fluid Mechanics | 0 | 0 | 1 (M) | 0 | 0 |  |  |
| Simple Harmonic Motion | 0 | 0 | 1 (M) | 0 | 0 |  |  |
| Equilibrium of Forces | 0 | 1 (M) | 0 | 0 | 0 |  |  |
| Simple Machines | 0 | 0 | 1 (M) | 0 | 1 (M) |  |  |
| Friction | 0 | 0 | 1 (M) | 0 | 0 |  |  |
| Total |  | 1 | 1 | 6 | 0 | 1 | 9 | 22.5% |
| Heat and Thermodynamics | Thermal Properties | 1 (E) | 1 (M) | 0 | 0 | 0 |  |  |
| Heat Transfer | 0 | 0 | 1 (M) | 0 | 0 |  |  |
| Kinetic Theory | 0 | 1 (M) | 0 | 0 | 0 |  |  |
| Thermodynamics | 0 | 0 | 2 (2M) | 0 | 0 |  |  |
| Total |  | 1 | 2 | 3 | 0 | 0 | 6 | 15% |
| Waves | Waves, Wave Properties | 0 | 1 (M) | 1 (M) | 0 | 0 |  |  |
| Optics | 1 (E) | 1 (M) | 2 (2M) | 0 | 0 |  |  |
| Total |  | 1 | 2 | 3 | 0 | 0 | 6 | 15% |
| Electromagnetism | Electrostatics | 0 | 1 (M) | 1 (M) | 0 | 0 |  |  |
| Current Electricity | 2 (E) | 2 (M) | 3 (M) | 1 (H) | 0 |  |  |
| Magnetism | 0 | 0 | 1 (M) | 0 | 0 |  |  |
| Electromagnetic Induction | 0 | 1 (M) | 0 | 0 | 0 |  |  |
| Total |  | 2 | 4 | 5 | 1 | 0 | 12 | 30% |
| Modern Physics | Atomic Structure | 1 (E) | 0 | 0 | 0 | 0 |  |  |
| Photoelectric Effect | 0 | 1 (M) | 0 | 0 | 0 |  |  |
| Radioactivity | 1 (E) | 0 | 2 (2M) | 0 | 1 (H) |  |  |
| Electronics | 0 | 0 | 1 (M) | 0 | 0 |  |  |
| Total |  | 2 | 1 | 3 | 0 | 1 | 7 | 17.5% |
| **G. Total** |  | 7(17.5%) | 10(25%) | 20(50%) | 1(2.5) | 2(5%) | 40 | 100% |

**Notes:** E = Easy, M = Moderate, H = Hard

**Appendix 2**: LLM *Test Blueprint*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Domain** | **Topic** | **Remembering** | **Understanding** | **Applying** | **Analyzing** | **Evaluating** | **Total Items** | **Item %** |
| Mechanics | Kinematics and Dynamics | 1 (E) | 0 | 2 (2M) | 0 | 0 |  |  |
| Work, Energy, and Power | 0 | 0 | 2 (1M,1H) | 0 | 0 |  |  |
| Fluid Mechanic | 0 | 0 | 3 (1M, 2H) | 0 | 0 |  |  |
| Equilibrium of Forces | 0 | 1 (M) | 0 | 0 | 0 |  |  |
| Simple Harmonic Motion | 0 | 0 | 1 | 0 | 0 |  |  |
| Simple Machines | 0 | 0 | 1(M) | 0 | 0 |  |  |
| Friction | 0 | 0 | 1 (M) | 0 | 0 |  |  |
| Total |  | 1 | 1 | 10 | 0 | 0 | 12 | 30% |
| Heat and Thermodynamics | Thermal Properties of Matter | 0 | 0 | 2 (M) | 0 | 0 |  |  |
| Heat Transfer | 0 | 2 (1M, 1H) | 0 | 0 | 0 |  |  |
| Kinetic Theory | 0 | 0 | 0 | 1 (M) | 0 |  |  |
| Thermodynamics | 0 | 0 | 0 | 1 (M) | 0 |  |  |
| Total |  | 0 | 2 | 2 | 2 | 0 | 6 | 15% |
| Waves | Wave Properties | 0 | 1 (H) | 1 (E) | 0 | 0 |  |  |
| Optics | 0 | 1 (M) | 0 | 0 | 0 |  |  |
| Total |  | 0 | 2 | 1 | 0 | 0 | 3 | 7.5% |
| Electromagnetism | Electrostatics | 0 | 0 | 1 (M) | 0 | 1 (M) |  |  |
| Current Electricity | 0 | 1 (M) | 4 (3M, 1H) | 0 | 0 |  |  |
| Magnetism | 0 | 0 | 1 (H) | 0 | 1 (H) |  |  |
| Electromagnetic Induction | 0 | 1 (H) | 0 | 0 | 0 |  |  |
| Total |  | 0 | 2 | 6 | 0 | 2 | 10 | 25% |
| Modern Physics | Atomic Structure | 0 | 1(M) | 0 | 0 | 1(H) |  |  |
| Radioactivity | 0 | 0 | 1 (H) | 0 | 0 |  |  |
| Solid State Physics | 0 | 1 (M) | 0 | 0 | 0 |  |  |
| Photoelectricity | 0 | 0 | 1 (H) | 0 | 0 |  |  |
| Electronics | 0 | 1 (M) | 0 | 0 | 0 |  |  |
| Conduction in Liquids & Gases | 1 (H) | 1 (M) | 0 | 0 | 1 (H) |  |  |
| Total |  | 1 | 4 | 2 | 0 | 2 | 9 | 22.5% |
| **G. Total** |  | 2(5%) | 11(27.5%) | 21(52.5%) | 2 (5%) | 5(12.5%) | 40 | 100% |

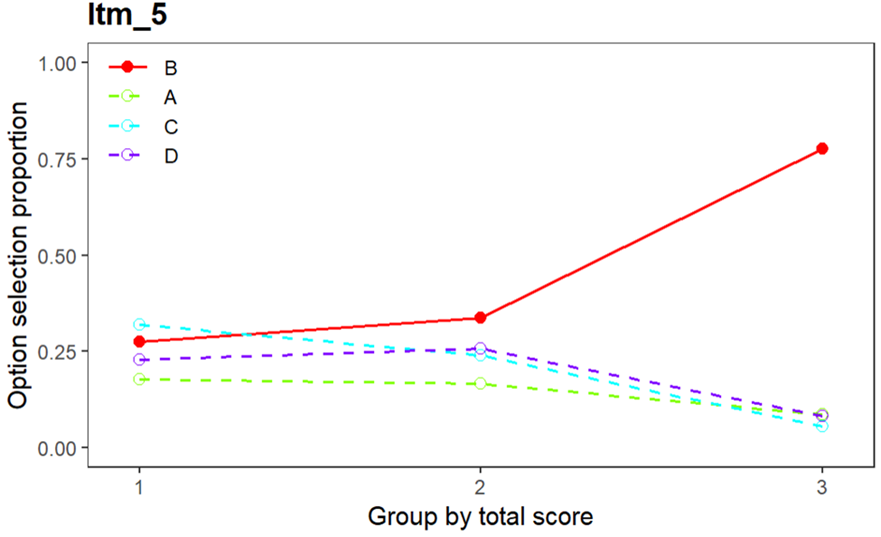
**Notes:** E = Easy, M = Moderate, H = Hard

**Appendix 3**

**Details of psychometrics analysis**

**1.** **Distractor analysis**

All 35 items have distractor frequency above the 5% response threshold as common benchmark for deeming a distractor as functioning well. Moreover, all distractors exhibited positive point-biserial correlations, suggesting that high-ability students were more likely to avoid incorrect options. Figure a shows a sample of distractor plot using the Shiny Item Analysis package in R (Martinková, & Drabinová, 2018).



**Figure a**: Sample of a distractor plot

**2. 1PL and 2PL model comparison**

**Table 1:** Model Comparison

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Deviance | AIC | BIC |
| 1PL | 22531.6 | 22603.6 | 22757.22 |
| 2PL | 22281.5 | 22421.5 | 22720.20 |

The 2PL model has lower AIC and BIC values compared to the 1PL (Rasch) model, indicating a better model fit, that means that both item difficulty and discrimination play important roles in estimating student abilities.

**3. Dimensionality study**

Model comparison using information criteria strongly favors unidimensional model over multidimensional, indicating a simpler factor structure fits the data best. Residuals from the unidimensional model are very small and centered near zero (for Q3 statistics median = -0.018, mean = -0.017) which suggests no large systematic misfit. This further confirms unidimensional as an adequate and parsimonious representation of the test data. Taking this all into account we can consider the test as essentially unidimensional.

**4. Item characteristics for 2PL model**

Item parameters were estimated using the 2PL model, which allows items to vary in both difficulty and discrimination. The discrimination parameter reflects how well an item differentiates between students with different ability levels, while the difficulty parameter represents the point on the latent trait continuum at which a student has a 50% probability of answering correctly.

**Table 2**: Item characteristics under 2PL model

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Entry** | **Discrimination** | **Location (difficulty)** | **Outfit** | **Infit** |
| 1 | 1.60 | -0.74 | 0.971 | 1.008 |
| 2 | 1.11 | -0.35 | 1.017 | 0.995 |
| 3 | 0.80 | 0.27 | 1.001 | 1.0 |
| 4 | 0.92 | 0.50 | 1.015 | 0.996 |
| 5 | 1.15 | 0.17 | 0.986 | 1.003 |
| 6 | 1.08 | 1.19 | 1.039 | 0.991 |
| 8 | 0.45 | 1.91 | 0.998 | 1.001 |
| 9 | 0.70 | 2.49 | 1.034 | 0.99 |
| 10 | 0.60 | 1.89 | 1.018 | 0.993 |
| 12 | 0.66 | 0.18 | 1.01 | 0.997 |
| 13 | 0.90 | 0.60 | 1.009 | 0.997 |
| 14 | 0.42 | 0.32 | 1.001 | 0.999 |
| 15 | 0.09 | -2.11 | 1.0 | 1.0 |
| 16 | 0.58 | 1.05 | 1.008 | 0.997 |
| 17 | 0.86 | 0.47 | 1.0 | 1.0 |
| 18 | 0.77 | 0.61 | 1.005 | 0.998 |
| 19 | 0.18 | 4.50 | 1.0 | 1.0 |
| 21 | 0.66 | 1.33 | 0.995 | 1.002 |
| 22 | 0.93 | 0.81 | 1.016 | 0.995 |
| 23 | 0.65 | 0.91 | 1.004 | 0.999 |
| 24 | 0.95 | 1.02 | 1.028 | 0.993 |
| 25 | 0.64 | 0.95 | 1.0 | 1.0 |
| 26 | 0.76 | 0.81 | 1.018 | 0.994 |
| 27 | 0.36 | 1.35 | 1.003 | 0.999 |
| 28 | 0.73 | 0.21 | 1.0 | 1.0 |
| 29 | 0.18 | 2.62 | 1.0 | 1.0 |
| 30 | 1.99 | 0.42 | 1.1 | 0.992 |
| 32 | 0.42 | 0.87 | 1.0 | 1.0 |
| 33 | 0.35 | 2.43 | 1.002 | 0.999 |
| 34 | 0.50 | 0.14 | 0.998 | 1.001 |
| 35 | 0.39 | 2.04 | 1.008 | 0.996 |
| 37 | 0.35 | 4.98 | 1.004 | 0.998 |
| 38 | 0.35 | 2.69 | 1.002 | 0.999 |
| 39 | 0.34 | 2.54 | 1.006 | 0.998 |
| 40 | 0.74 | 1.69 | 1.021 | 0.995 |

**Appendix 4**

**Sample of Items in the test**

(See the supplementary file to see the complete file)

1. Consider the following physical quantities:

I.     Mass II.     Time III.     Temperature IV.   Displacement

Which of the above is a vector quantity?

A. I only B. II only C. III only D. IV only

1. A rocket is launched from the Earth's surface. What is the minimum velocity it must attain to escape the Earth's gravitational pull completely? (Take the radius of the Earth as R and the acceleration due to gravity at the surface as g)

A. √(gR) B. √(2gR) C. gR D. 2gR

1. A uniform ladder of length 5 m and weight 200 N leans against a smooth vertical wall. The coefficient of static friction between the ladder and the rough horizontal ground is 0.4. If the ladder makes an angle of 53° with the ground, what is the maximum distance a person of weight 600 N can climb up the ladder before it starts to slip?
2. 1.5 m B. 2.5 m C. 3.5 m D. 4.5 m
3. A block of mass 2 kg is released from rest at the top of a frictionless inclined plane of height 1 m. At the bottom of the incline, it collides elastically with another block of mass 4 kg, initially at rest. What is the velocity of the 4 kg block after the collision?

A. 1 m/s B. 2 m/s C. 2.67 m/s D. 4 m/s

1. A spring is stretched by 10 cm when a mass of 2 kg is hung from it. If the mass is then pulled down further by 5 cm and released, what is the period of oscillations? (Take g = 10 m/s²)

A. 0.2 s B. 0.45 s C. 0.63 s D. 0.9 s

1. A thermos flask is designed to minimize heat transfer. Which of the following statements about a thermos flask is **FALSE**?

A. The vacuum between the walls prevents heat transfer by conduction and convection

B. The silvered inner walls minimize heat transfer by radiation

C. The insulated stopper reduces heat transfer by conduction

D. The vacuum between the walls allows for efficient heat transfer by radiation

1. A body of mass 36kg falls through a viscous liquid which offers a drag force of 260N on the body. The upthrust on the body at terminal velocity is [g = 10 ms-2 ]

A. 50N B. 100N C. 310N D. 620N

1. A projectile is launched at a 30° angle with an initial velocity of 20 m/s, what is its initial vertical velocity component?
2. A train is approaching a stationary observer at a speed of 30 m/s. The train's whistle emits a sound with a frequency of 500 Hz. If the speed of sound in air is 340 m/s, what frequency will the observer hear? The air temperature is 25 °C, and the relative humidity is 50%.

A. 455.9 Hz B. 500 Hz C. 548.8 Hz D. 550 Hz

1. Which of the following is NOT an application of eddy currents?

A. Induction furnace

B. Magnetic braking in trains

C. Metal detectors

D. Transformers

Other relevant data (with the analysis code) used in this study are available in the supplementary file. Supplementary data to this article can be found online at <https://github.com/MOO-DIO/LLM-paper.git>.