The interplay between psychometric properties of items and Cognitive Levels of Bloom's Taxonomy in assessment design: A case of five MCQ examination papers from the ECZ

Subtheme

Leveraging educational assessment data for decision-making and accountability.

Simwinga Simwinga

Examinations Council of Zambia

PRESENTATION OUTLINE

- Background information
- Justification of the Study
- Research Objectives
- Research Questions
- Methodology (survey instruments, sample size, data collection, analysis)
- Findings
- Discussion & Conclusion
- Recommendations

BACKGROUND INFORMATION

- MCQs are widely used in educational assessments due to:
  1. Efficiency in scoring
  2. Cost-effectiveness
  3. Ability to assess a broad range of cognitive skills
- Effectiveness depend on:
  1. The quality of psychometric properties
  2. Alignment to the Test Blueprint (Bloom's Taxonomy and Test Specifications)
- Psychometric Properties
  - **Item Difficulty** (ID): Proportion of students answering an item correctly (0 to 1).
    Discrimination Index (DI): Distinguishes strong vs. weak students (-1 to 1)
  - **Distractor Efficiency** (DE): Effectiveness of incorrect options (Plausible Options)
    - Functioning distractors (FD): chosen by ≥5% of students
    - Non-functioning distractors (NFD): chosen by <5%

BACKGROUND INFORMATION (Cont..)

- In Zambia,
  - acceptable ID range: **0.20 ≤ ID ≤ 0.80**
  - acceptable DI range: **0.20 < DI ≤ 1.00**
- MCQs are used at PSLE and SCOLE
- PSLE: All papers consist of MCQs
- SCOLE: Five papers (*Geography P1, Physics P1, Chemistry P1, Biology P1, Principles of Accounts P1*) consist MCQs
- Items are developed by trained item writers
- Paper development is guided by
  - Test Specification Tables (TSTs)
  - Psychometric Properties
  - Expert judgement
- TSTs guide distribution of items by content and cognitive level

BACKGROUND INFORMATION (Cont..)

- Distribution of Items based on the TST.

Subject	Remember (%)	Understand (%)	Apply (%)	Analyse (%)	Evaluate (%)	Create (%)
Geography	36	32	20	12	0	0
Physics	12	30	25	25	5	3
Chemistry	13	35	17	25	5	5
Biology	7	45	15	15	10	8
Principles of Accounts	15	70	15	0	0	0

JUSTIFICATION OF THE STUDY

- When properly constructed, MCQs can reliably assess students' knowledge, understanding, and application of learned material *(Haladyna & Rodriguez, 2013)*.
- Despite the structured approach to MCQ development, issues still persist.
- For example, in the 2022, Physics Paper 1 had some items with ID outside the acceptable range.
- Limited research

OBJECTIVES OF THE STUDY

- The study seeks to evaluate the quality of MCQs in terms of their psychometric properties at SCOLE in Zambia
- Assess the extent to which items are aligned with the TSTs and cognitive levels defined in the Bloom's Taxonomy.

RESEARCH QUESTIONS

The study is guided by three research questions:

1) What is the quality of MCQs at SCOLE level in Zambia, as measured by ID, DI, and DE?
2) To what extent do the MCQs align with the TSTs?
3) Does the ID of test items correspond appropriately to the cognitive demands outlined in Bloom's Taxonomy?

METHODOLOGY

Study Design

- Descriptive research design to evaluate item quality and alignment with cognitive skill levels

Sample Selection

- 840 four-option MCQs from five public examination papers administered between 2018-2022

Data Collection

- Item analysis data from post-examination reports generated by the ECZ
- Two teachers independently mapped each item to one of six cognitive levels of the bloom's taxonomy with inter-rater reliability determined using Cohen's Kappa (κ)

METHODOLOGY (Cont..)

Interpretation of Cohen's kappa

Kappa (κ)	Level of Agreement	Data Reliability
0.00 - 0.20	None	00 - 04%
0.21 - 0.39	Minimal	04 - 15%
0.40 - 0.59	Weak	15 - 35%
0.60 - 0.79	Moderate	35 - 63%
0.80 - 0.90	Strong	64 - 81%
Above 0.90	Almost Perfect	82 - 100%

Inter-Rater reliability across the five subjects

Subjects	Percent Agreement	Kappa (κ)
Geography	0.84	0.73
Physics	0.91	0.86
Chemistry	0.83	0.76
Biology	0.79	0.71
Principles of Accounts	0.91	0.80

Data Analysis

- Microsoft Excel and STATA used to analyze ID, DI, DE, and cognitive skill alignment

FINDINGS

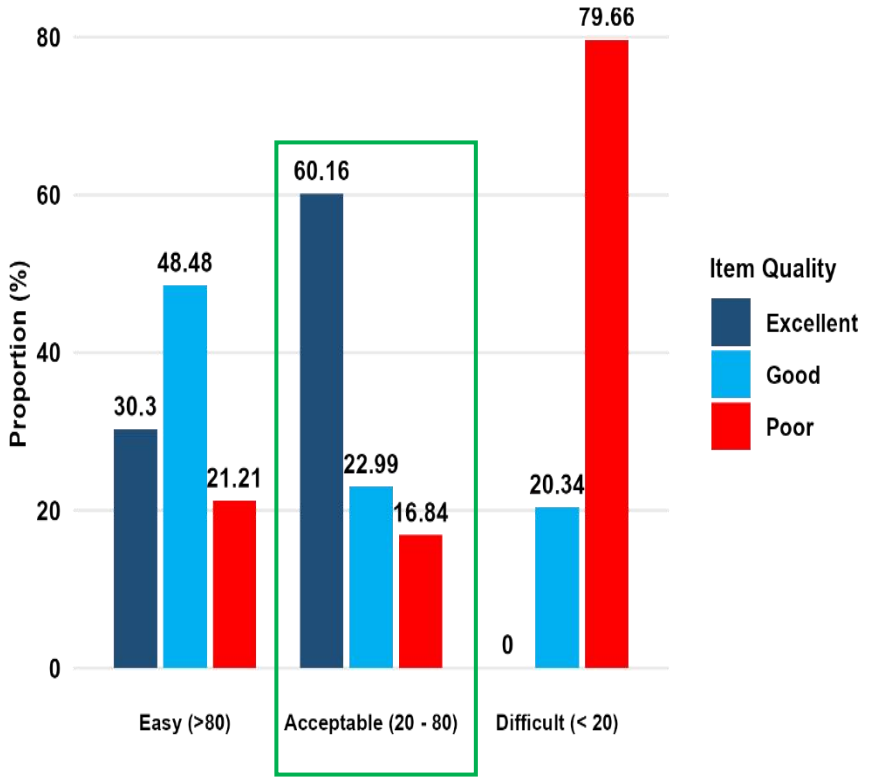
What is the quality of MCQs at SCOLE level in Zambia, as measured by ID, DI, and DE?

Item Difficulty

Subject	<0.2	0.20 – 0.8	>0.8
Geography	17 (8.50%)	182 (91.00%)	1 (0.50%)
Physics	13 (8.13%)	138 (86.25%)	9 (5.63%)
Chemistry	12 (7.50%)	131 (81.88%)	17 (10.63%)
Biology	5 (3.13%)	152 (95.00%)	3 (1.88%)
Principles of Accounts	12 (7.50%)	145 (90.63%)	3 (1.88%)

Discrimination Index

Subject	<0.2	0.2 - 0.35	>0.35
Geography	73 (36.50%)	76 (38.00%)	51 (25.50%)
Physics	35 (21.88%)	27 (16.88%)	98 (61.25%)
Chemistry	24 (15.00%)	27 (16.88%)	109 (68.13%)
Biology	24 (15.00%)	28 (17.50%)	108 (67.50%)
Principles of Accounts	24 (15.00%)	42 (26.25%)	94 (58.75%)

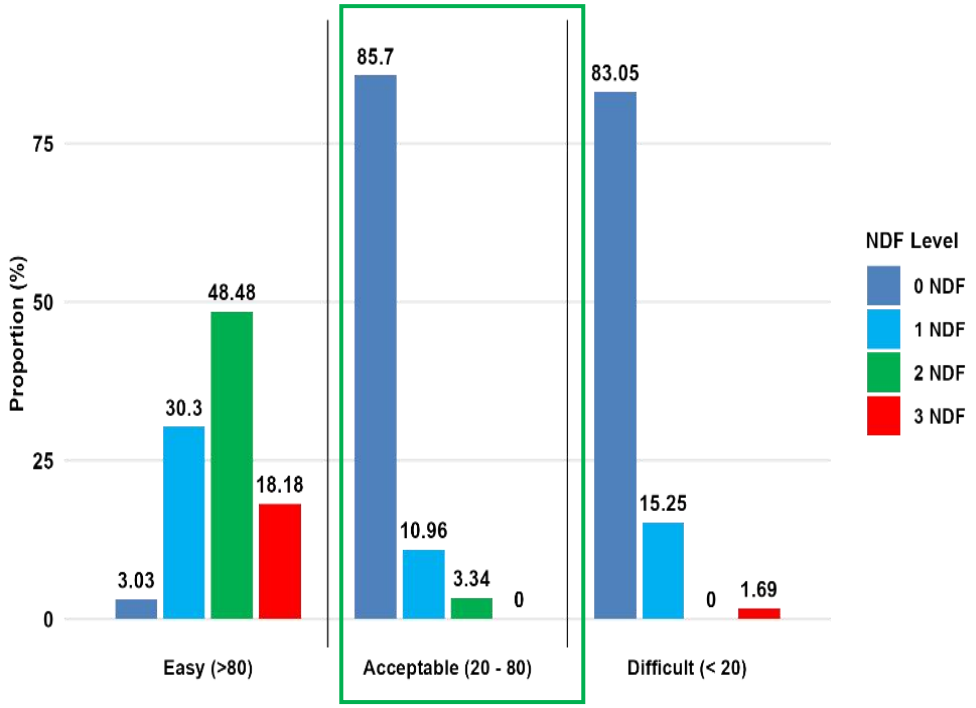


FINDINGS (Cont..)

What is the quality of MCQs at SCOLE level in Zambia, as measured by ID, DI, and DE?

Distractor Efficiency

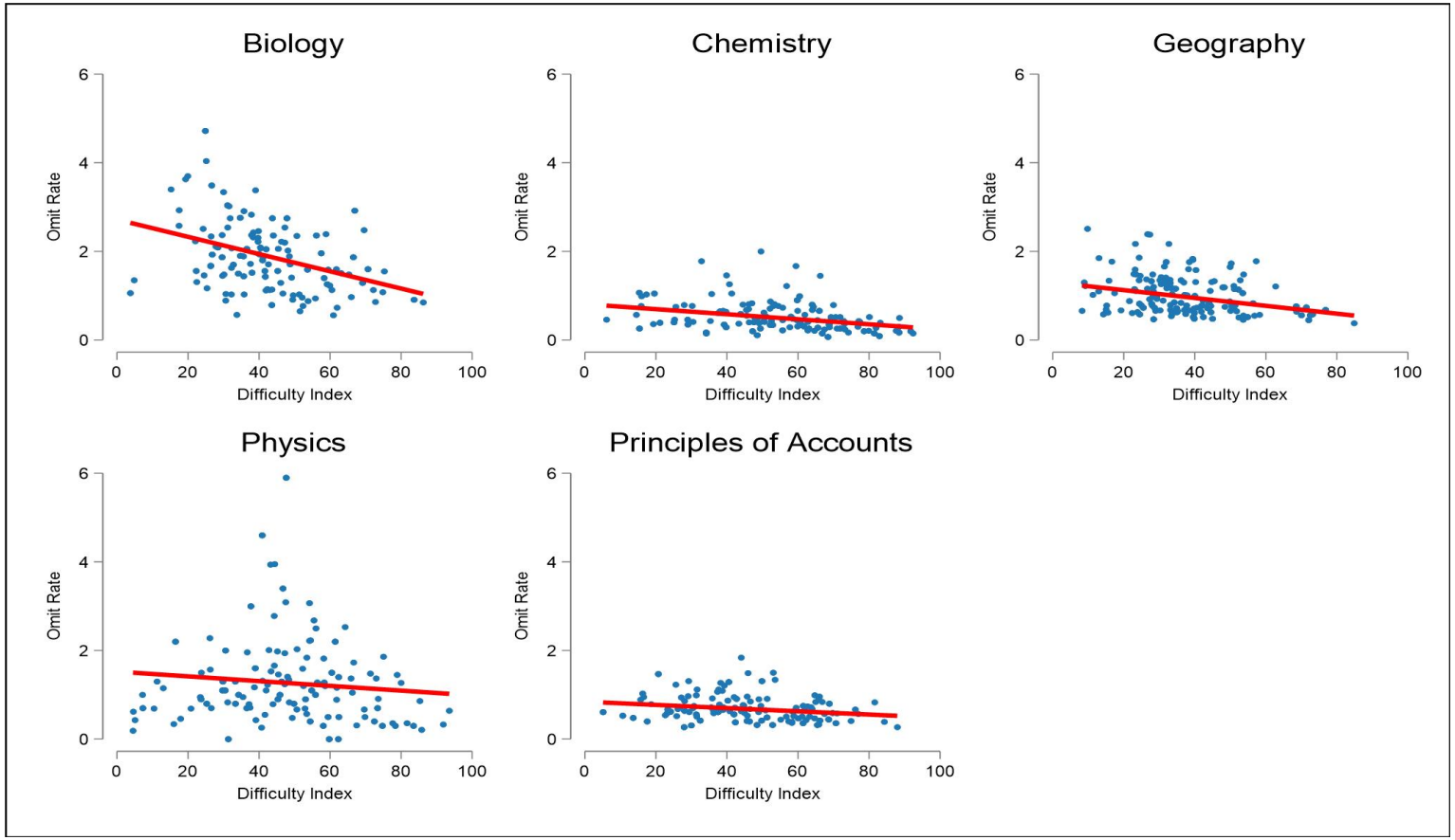
Subject	0 NFDs	1 NFDs	2 NFDs	3 NFDs
Geography	196 (98.00%)	3 (1.5%)	1 (0.5%)	0 (0.00%)
Physics	110 (68.75%)	35 (21.88%)	12 (7.50%)	3 (1.88%)
Chemistry	107 (66.88%)	33 (20.63%)	16 (10.00%)	4 (2.50%)
Biology	141 (88.13%)	15 (9.38%)	4 (2.5%)	0 (0.00%)
Principles of Accounts	137 (85.63%)	15 (9.38%)	8 (5.00%)	0 (0.00%)
	Over 67%			**2 %**



FINDINGS (Cont..)

What is the quality of MCQs at SCOLE level in Zambia, as measured by ID, DI, and DE?

Correlation between Item Difficulty and Omission Rate



FINDINGS (Cont..)

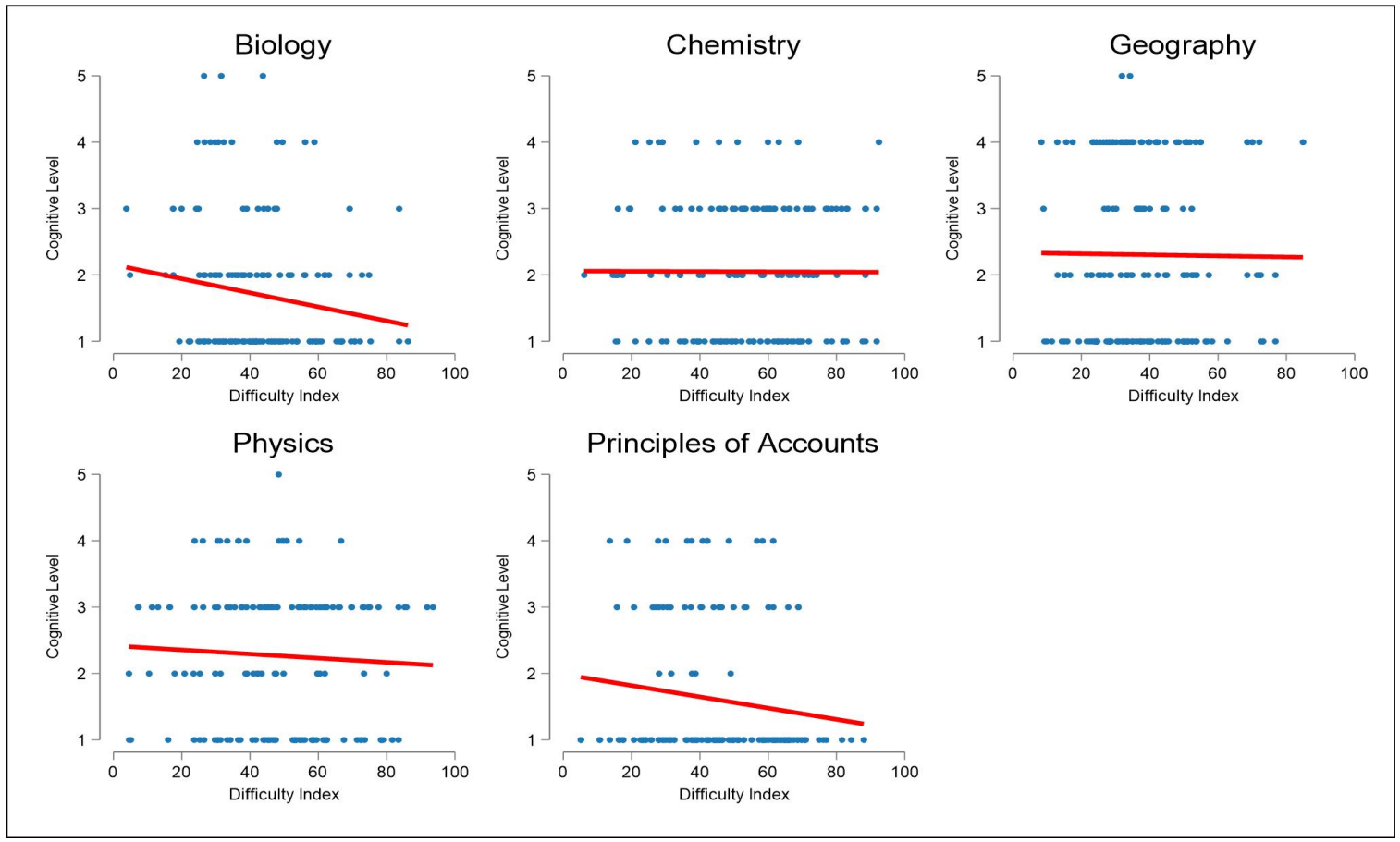
To what extent do the MCQs align with the TSTs?

Subject		Remember (%)	Understand (%)	Apply (%)	Analyse (%)	Evaluate (%)	Create (%)	Total (%)
Geography	Test Spec	36.00	32.00	20.00	12.00	0.00	0.00	100
	Inter Corder	41.00	19.00	9.50	29.50	1.00	0.00	100
Physics	Test Spec	12.50	30.00	25.00	25.00	5.00	2.50	100
	Inter Corder	34.38	15.00	40.63	9.38	0.63	0.00	100
Chemistry	Test Spec	12.50	35.00	17.50	25.00	5.00	5.00	100
	Inter Corder	40.63	21.25	30.63	7.50	0.00	0.00	100
Biology	Test Spec	7.50	45.00	15.00	15.00	10.00	7.50	100
	Inter Corder	58.13	24.38	8.75	6.88	1.88	0.00	100
Principles of Accounts	Test Spec	15.00	70.00	15.00	0.00	0.00	0.00	100
	Inter Corder	72.50	3.13	16.25	8.13	0.00	0.00	100

FINDINGS (Cont..)

Does the ID of test items correspond appropriately to the cognitive skill levels outlined in Bloom's Taxonomy?

Correlation between Item Difficulty and Cognitive Demand



Subject	Correlation Coef.	p-Value
Geography	-0.01	0.893
Physics	-0.057	0.473
Chemistry	-0.004	0.965
Biology	-0.162	0.041
Principles of Accounts	-0.144	0.068

DISCUSSION & CONCLUSION

- MCQs in Zambia generally show sound psychometric properties
- Many items had appropriate difficulty and discrimination
- Many items in acceptable ID range exhibited the most effective distractor functioning
- Misalignment between items and TST
- Overrepresentation of lower-order thinking skills
- Weak association between ID and cognitive demand

RECOMMENDATIONS

- Capacity build item writers:
  1. Development of plausible distractors
  2. Pitching items at appropriate cognitive demand
- Conduct pretesting of items before live exams

THANK YOU FOR YOUR ATTENTION