

SUPPLEMENTARY MATERIAL

Hybrid Framework for Pre-Assessing Arabic Multiple-Choice Questions Using ChatGPT and Traditional Item Analysis

Omar Dahham Sarhan

Email: omar.sarhan@uoanbar.edu.iq

Additional content is provided as further examples of the prompting style used in the AI-based before-assessment phase. The extra content is a sample prompt given to ChatGPT to predict the overall difficulty of the test, the discrimination function of each item, and the effectiveness of each distractor before the test was administered.

Supplementary Prompt Example:

Original Arabic Prompt

"بالاعتماد على خوارزميات الذكاء الاصطناعي قم بالتنبأ بفاعلية الأسئلة بنفس أسلوب التقييم التقليدي، بما يشمل معامل التمييز، معامل الصعوبة، وفاعلية المشتتات. قم بتنظيم النتائج في جدول بحيث تكون قابلة للمقارنة مع التقييم البعدي بعد تطبيق الاختبار على الطلاب."

English Translation

"Utilizing AI-enabled analytical thinking, forecast the success of each of the next MCQs based on the commonly used psychometric measures (item difficulty measure, discrimination index, and distractor effectiveness) used in MCQ assessment. Present your predicted results in a table to facilitate comparison with student-based evaluation results after the post-test."

Supplementary Table S1. Response Distribution Matrix for MCQ Items

Supplementary Table S1 presents the frequency distribution of students' responses across all MCQ items. The table specifies the number of students who selected each answer option for every question. Granting this additional information is a deliberate step to enhance transparency and ensure that the traditional psychometric analysis methods can be checked.

Question	A	B	C	D	No Response	Correct Answer
Q1	7	46	7	1	4	7
Q2	2	1	46	5	11	46
Q3	4	37	0	1	23	37
Q4	48	4	3	1	9	48
Q5	13	2	0	42	8	42
Q6	1	52	1	1	10	52
Q7	1	34	25	3	2	34
Q8	13	3	34	9	6	34
Q9	3	10	24	15	13	24
Q10	9	47	5	1	3	47
Q11	2	46	2	9	6	46
Q12	3	4	7	19	32	19
Q13	24	6	3	13	19	24
Q14	1	3	2	58	1	58
Q15	3	54	4	2	2	54
Q16	20	1	20	15	9	20
Q17	22	6	7	8	22	6
Q18	3	5	41	5	11	41