

Nama : Alfiantika Putri
NPM : 2413031095
Kelas : 2024C
Mata Kuliah : Evaluasi Pembelajaran Ekonomi

MENGANALISIS KUALITAS TES DAN BUTIR – BUTIR SOAL

A. Pendahuluan

Analisis kualitas tes merupakan kegiatan yang dilakukan untuk mengevaluasi kelayakan instrumen yang telah di ujicobakan kepada peserta didik. Melalui analisis ini, dapat diketahui tingkat validitas dan reliabilitas instrumen sehingga dapat ditentukan kelayakan soal sebagai alat ukur hasil belajar pada materi Perpajakan dalam mata pelajaran ekonomi.

B. Penyajian Data Hasil Uji Coba

[illegible]

[illegible]

C. Analisis Tingkat Kesukaran Soal

Analisis tingkat kesukaran soal dilakukan untuk mengetahui tingkat kesulitan setiap butir soal yang diujikan kepada peserta didik, melalui analisis ini dapat diketahui apakah suatu soal termasuk dalam kategori mudah, sedang, atau sukar. Hasil analisis tingkat kesukaran dapat digunakan sebagai bahan evaluasi untuk memperbaiki dan menyempurnakan kualitas soal yang akan digunakan pada kegiatan pembelajaran berikutnya.

$$\text{Rumus} = P = B/J$$

Keterangan:

P = indeks kesukaran soal

B = Jumlah peserta didik yang menjawab benar

J = Jumlah seluruh peserta didik

Nilai	Kategori
0,00 - 0,29	Sukar
0,30 - 0,69	Sedang
0,70 - 1,00	Mudah

No Soal	Indeks Kesukaran	Kategori
1	0,48	Sedang
2	0,94	Mudah
3	0,67	Sedang
4	0,76	Mudah
5	1,00	Mudah

6	0,85	Mudah
7	0,91	Mudah
8	0,82	Mudah
9	1,00	Mudah
10	0,94	Mudah
11	0,97	Mudah
12	0,97	Mudah
13	1,00	Mudah
14	0,97	Mudah
15	0,94	Mudah
16	0,94	Mudah
17	0,94	Mudah
18	1,00	Mudah
19	0,76	Mudah
20	0,94	Mudah

D. Uji Validitas

[illegible]

[illegible]

S1 8	Pearson Correlati on	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b	. ^b
	Sig. (2- tailed)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
	N	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
S1 9	Pearson Correlati on	-.159	.153	.050	.010	. ^b	.155	.313	.100	. ^b	.153	.313	-.100	. ^b	.313	.153	.153	.449**	. ^b	1	-.144	.429*
	Sig. (2- tailed)	.378	.396	.782	.956	.	.388	.076	.580	.	.396	.077	.580	.	.077	.396	.396	.009	.		.425	.013
	N	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
S2 0	Pearson Correlati on	-.008	.468**	.090	.153	. ^b	.247	-.080	-.120	. ^b	-.065	-.045	.696**	. ^b	-.045	-.065	-.065	-.065	. ^b	-.144	1	.224
	Sig. (2- tailed)	.966	.006	.619	.396	.	.166	.657	.507	.	.721	.804	.000	.	.804	.721	.721	.721	.	.425		.210
	N	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33
TO TA L	Pearson Correlati on	.341	.349*	.434*	.499**	. ^b	.581**	.435*	.377*	. ^b	.475**	.505**	.418*	. ^b	.505**	.475**	.099	.600**	. ^b	.429*	.224	1
	Sig. (2- tailed)	.052	.046	.012	.003	.	.000	.011	.030	.	.005	.003	.016	.	.003	.005	.585	.000	.	.013	.210	
	N	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33	33

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

b. Cannot be computed because at least one of the variables is constant.

Analisis: Validitas butir soal dapat dilihat dari nilai korelasi antara skor setiap butir soal dengan skor total (*item-total correlation*). Suatu butir soal dinyatakan valid apabila nilai korelasi (*r* hitung) lebih besar dari nilai *r* tabel. Dengan jumlah sampel 33 orang pada taraf signifikansi 5%, nilai *r* tabel sebesar sekitar 0,344. Berdasarkan hasil uji, dari 20 butir soal yang diuji terdapat 12 soal yang valid dan 8 soal yang tidak valid. Soal yang valid menunjukkan bahwa butir soal tersebut mampu mengukur apa yang seharusnya diukur sehingga layak digunakan dalam instrumen tes. Sementara itu, soal yang tidak valid menunjukkan bahwa butir soal kurang mampu mengukur tujuan yang diharapkan, sehingga perlu direvisi atau diganti.

E. Uji Reliabilitas

Case Processing Summary

		N	%
Cases	Valid	33	100,0
	Excluded ^a	0	,0
	Total	33	100,0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
,635	20

Analisis: Berdasarkan hasil uji reliabilitas, diperoleh nilai Cronbach's Alpha sebesar 0,635. Nilai tersebut menunjukkan bahwa instrumen memiliki tingkat reliabilitas yang cukup baik, sehingga hasil pengukuran yang diperoleh konsisten. Oleh karena itu, instrumen tes dapat dikatakan cukup reliabel dan layak digunakan untuk mengukur hasil belajar.