



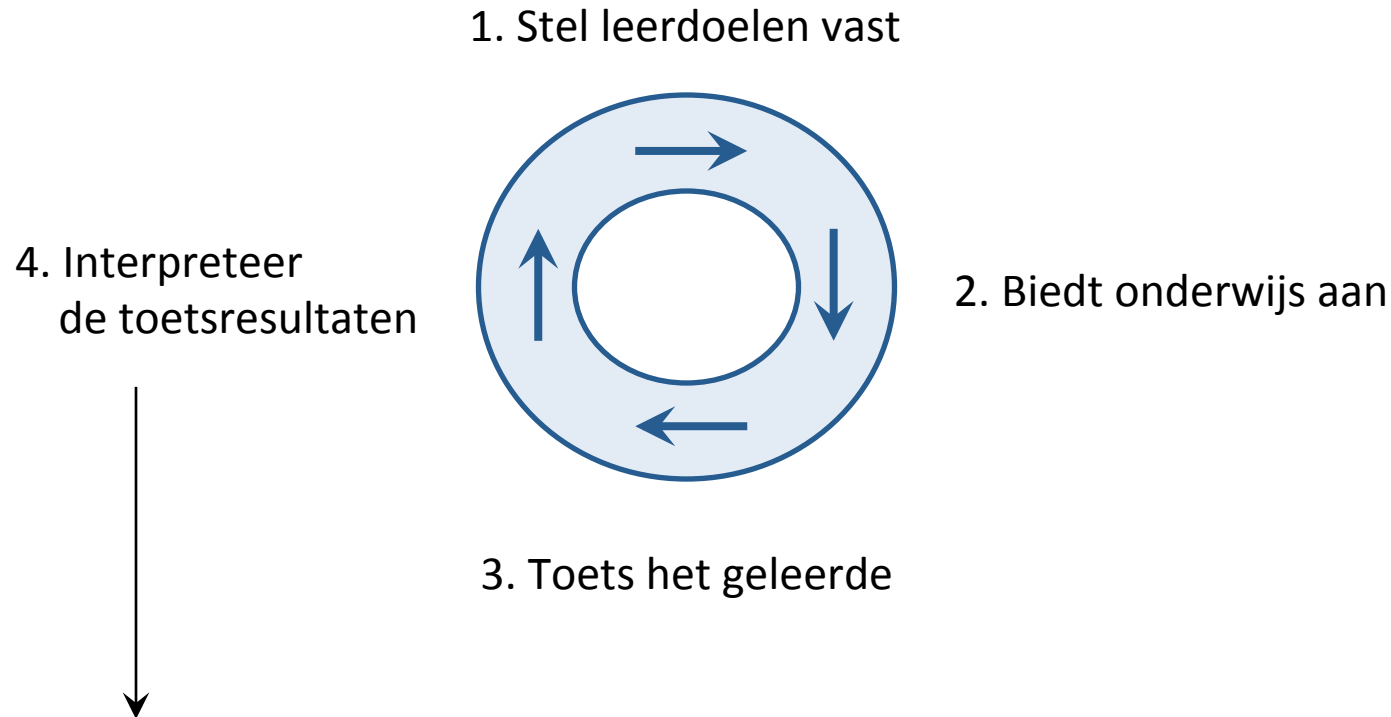
Workshop

Toetsmatrijs & toetsanalyse

Inhoud

- Waarom?
- Wat is het?
- Hoe doe ik het?
- Wat doe ik er mee?
- Hoe kom ik meer te weten?

Waarom?



Student:

Ken/kan ik het? - formatief

Docent:

Kent/kan de student het? - formatief en summatief
Is het aangeboden onderwijs OK? - kwaliteitscontrole

Waarom?

Een toets moet **valide**, **betrouwbaar** (en rechtvaardig) zijn:

Validiteit

Het niveau dat de toets bevraagt komt overeen met het niveau in de leerdoelen
De oefenopdrachten weerspiegelen de toetsopgaven
De gekozen toetsvorm toetst de gewenste vaardigheid
De toets is haalbaar indien je voldoet aan de leerdoelen

Betrouwbaarheid

De toets maakt een individuele beoordeling mogelijk
De manier van toetsen is vooraf bekend gemaakt
De toets is helder en eenduidig geformuleerd
De toets wordt objectief beoordeeld

Waarom?

Het opstellen van een toets kent drie punten van aandacht:

- het formuleren van de toetsvragen
- het samenstellen van de toets
- de feitelijke afname van de toets

Formuleren: de vragen moeten overeenkomen met de **leerdoelen** van de cursus

Samenstellen: de vragen moeten goed verdeeld zijn over de **cursusonderwerpen** en de verschillende **cognitieniveaus**
hulpmiddel: **toetsmatrijs**

Afname: met een **psychometrische analyse** inzicht in de **kwaliteit** van de toets

Wat is het?

Een **lijst met onderwerpen** die je wilt toetsen en een **indicatie van het niveau** waarop je die onderwerpen wilt toetsen

Toetsmatrijs theorie-examens voor kandidaat-rijinstructeurs

	Kwaliteitscriteria van autorijden								Totaal
	Veiligheid		Doorstroming		Milieu		Sociaal rijgedrag		
	VE	H	VE	H	VE	H	VE	H	
Vorbereiden en afsluiten van verkeersdeelname	2	0	2	0	1	0	0	0	5
Stromen	4	2	4	1	1	1	1	1	15
Kruisen	7	3	5	1	0	0	3	1	20
Zijdelings verplaatsen	7	3	3	1	1	1	3	1	20
Totaal	20	8	14	3	3	2	7	3	60

Toetsmatrijs examenthema 'Theorie van de rijtaak': aantal waarneem- en voorspel/evalueer opgaven (VE) en aantal handelingskeuzeopgaven (H) per hoofdrijtaak.

Wat is het?

Functies van een toetsmatrijs:

U vermijdt dat teveel opdrachten worden gemaakt die gericht zijn op dezelfde leerstof dan wel op dezelfde vaardigheid

Wanneer de toetsmatrijs een juiste verhouding weergeeft van het aantal vragen in vergelijking met de verschillende onderdelen van het leer- of examenprogramma, is de kans dat de toets een representatieve steekproef vormt van de te toetsen doelstellingen groter

U kunt als u twee toetsen over dezelfde leerstof wilt maken de gelijkwaardigheid tussen die toetsen vergroten door ze beide op te stellen aan de hand van één toetsmatrijs

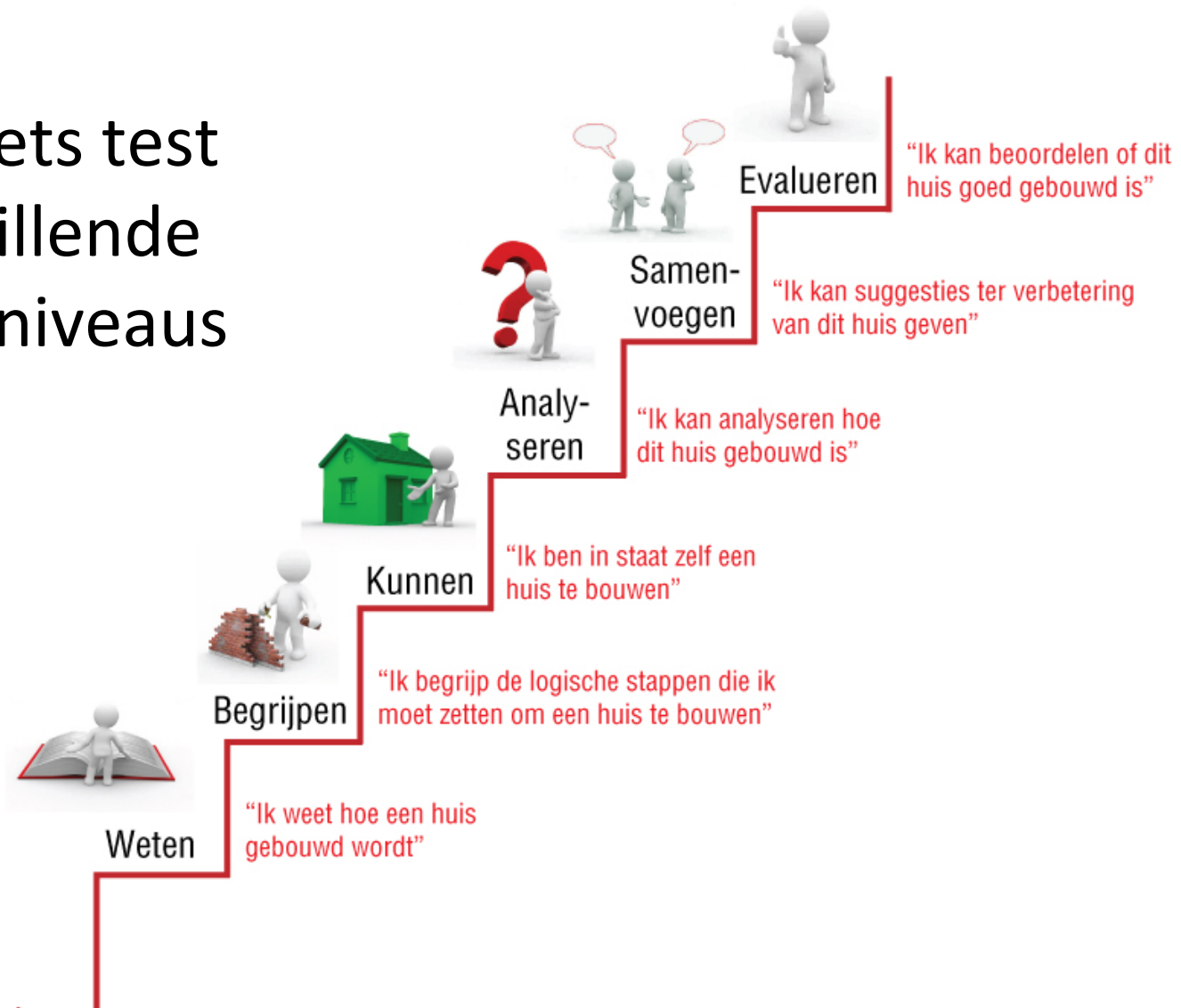
De toetsmatrijs kan dienen als een verantwoording van de inhoud van de toets naar anderen, zoals collega-vakdocenten en examencommissie

Functie: zelfcontrole en verantwoording

Hoe doe ik het?

Een toets test
verschillende
kennisniveaus

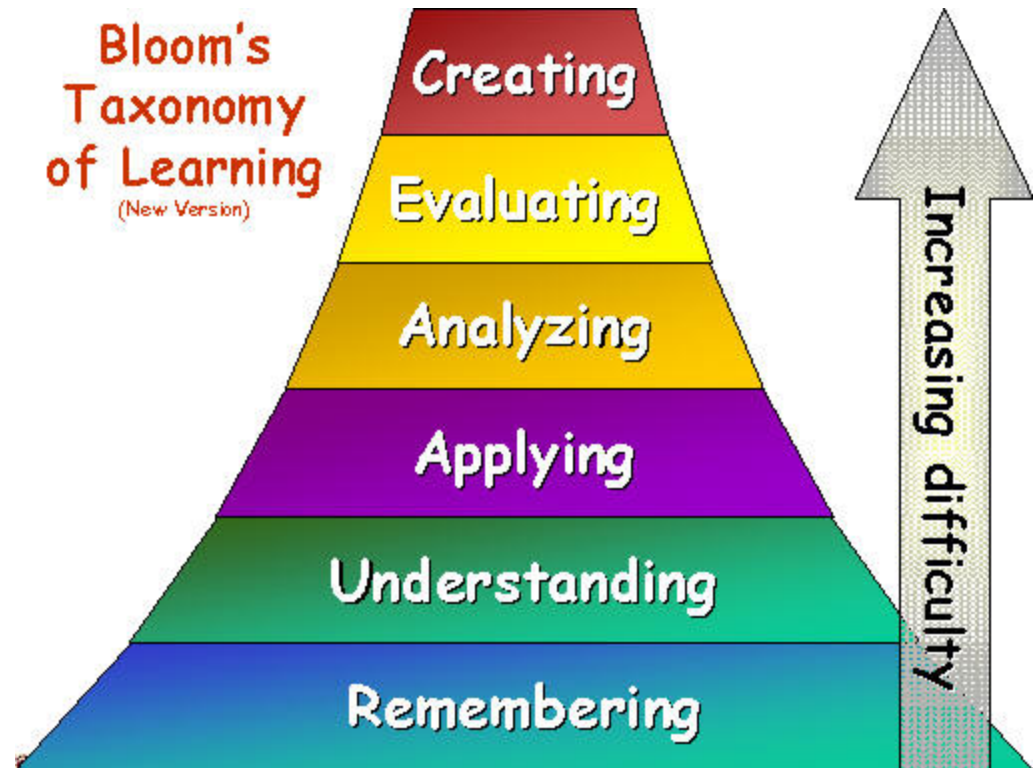
Taxonomie van Bloom



Hoe doe ik het?

Een toets test
verschillende
kennismiveaus

Taxonomie van Bloom



Hoe doe ik het?

Voorbeeld toetsmatrijs

Toetsmatrijs eindtoets FA-103 (Celbiologie)

Januari 2013

Taxonomie van begripsniveau's volgens Bloom

Kennis: Je kunt je iets herinneren, reproductie van feiten

Begrip: Je kunt iets in eigen woorden omschrijven, samenvatten; aantonen dat de kennis een betekenis heeft, bepalen van de gevolgen van de informatie, verder denken dan

Toepassing: Je kunt informatie gebruiken in een nieuwe situatie en een probleem oplossen

Analyse: Je kunt informatie opdelen en de onderdelen met elkaar in verband brengen

Synthese: Je brengt onderdelen samen tot iets nieuws

Evaluatie: Je hebt een beargumenteerd oordeel en een standpunt

Onderwerpen	Open vragen		Gesloten vragen	
	#	niveau subvragen	#	niveau subvragen
Celbiologie algemeen	0		0	
Celmembraan/membraantransport	2	2xB, 1xT	1	1xK
DNA replicatie, regulatie eiwitsynthese	2	2xK, 1xB, 1xA	2	1xK, 1xB
GPCR	1	1xK, 2xB	2	1xK, 1xT
RTK	1	2xK, 1xB	2	2xK
Celcyclus	2	2xK, 1xB, 4xT	1	1xB
	totaal 8	7xK, 7xB, 5xT, 1xA	8	5xK, 2xB, 1xT

Hoe doe ik het?

Voorbeeld toetsmatrijs

		Begripsniveau's volgens Bloom*					
		Kennis	Begrip	Toepassing	Analyse	Synthese	Evaluatie
Ligand binding							
	data-analyse	2a	1b	1d			
	theoretische aspecten	1a	1b 1c	1e 1f			
Enzymkinetiek							
	data-analyse		3e	3c2			
	theoretische aspecten		3c1				
Receptorwerking							
	data-analyse	2d		2f			
	theoretische aspecten		2e				
Thermodynamica							
	data-analyse		2c				
	theoretische aspecten		2b				
Chemisch rekenen		3b					
Algemeen chemisch inzicht (pH, pKa, functionele groepen, zuur-base)							
	data-analyse		3b	3d			
	theoretische aspecten	3a					

Hoe doe ik het?

Werkwoorden bij een vraagconstructie naar cognitieniveau:

Reproductie

Kennis:	‘benoem’ of ‘som op’
Begrip:	‘vertel in eigen woorden’

Inzicht

Toepassing:	‘licht toe’ of ‘leg uit’
Analyse:	‘ontwerp’ of ‘formuleer’

Toepassen

Synthese:	‘verklaar’ of ‘beredeneer’
Evaluatie:	‘bereken’ of ‘demonstreer’

Wat doe ik er mee?

Vóór/tijdens de samenstelling van de toets:

|

zelfcontrole

Na afloop van de toets:

|

verantwoording



Neem je eigen toets en maak een toetsmatrijs
op basis van de taxonomie van Bloom

Wat is het?

Toetsanalyse zegt iets over

betrouwbaarheid

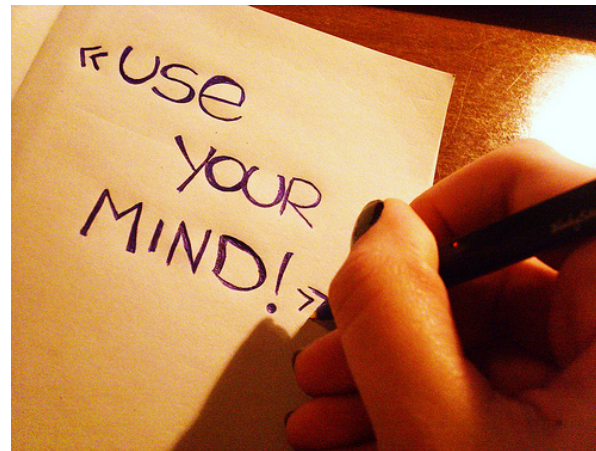
bij herhaalde afname onder dezelfde omstandigheden eenzelfde resultaat

validiteit

meet de toets wat de toets moet meten

Psychometrische analyse
op itemniveau
op toetsniveau

p-waarde R_{it} R_{ir} KR20 / Cronbach's α



Hoe doe ik het?

Excel spreadsheet van Jaap Milius en Andries Koster

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB	AC			
1	 Universiteit Utrecht					TOETSANALYSE van SCORES																										
2																																
3																																
4																																
5	Departement:																															
6	Cursus:																															
7	Toets:																															
8	Datum:																															
9																																
10	aantal deelnemers:						N.B. maximaal 300																									
11	aantal vragen:						N.B. maximaal 40																									
12																																
13						In de gele velden kunnen de scores per vraag worden ingevuld. Vermeld hierboven het overeenkomstige aantal deelnemers en aantal vragen.																										
14																																
15																																
16																																
17																																
18																																
19	Nr.	totaal	score/vr	respnr.	naam																											
20	1																															
21																																
22																																
23																																
24																																
25																																
26																																
27																																
28																																
29																																
30																																
31																																
32																																
33																																
34																																
35																																
36																																

Hoe doe ik het?

Excel spreadsheet van Jaap Milius en Andries Koster

[illegible]

Hoe doe ik het?

Belangrijke tips:

Voer alle (deel)vragen in:

dus **1a-f, 2a-d, 3a-c, 4a-b** zijn 15 vragen, i.p.v. 4

Voer alleen die studenten in die verschijnen op het tentamen

dus niet alle studenten die ingeschreven staan voor het tentamen

[illegible]

Wat doe ik er mee?

(1) Item-analyse Moeilijkheid

p-waarde = moeilijkheidsgraad v/d vraag

- $p = \frac{\text{aantal studenten dat de vraag goed heeft}}{\text{totaal aantal studenten}}$
- waarde tussen 0 en 1 (0 = moeilijk, 1 = makkelijk)

Actie: vergelijk bedoeling v/d vraag met de p-waarde, op basis hiervan vraag al dan niet eruit halen

p-waarde < kan betekenen: (te) hoge moeilijkheidsgraad
tijd tekort (bij laatste vragen in tentamen)
onduidelijk geformuleerde vraag
een hiaat in het geboden onderwijs

Wat doe ik er mee?

(2) Item-analyse Correlatie

R_{it} - en R_{ir} -waarde = onderscheidend vermogen v/d vraag

- Verband vraagscore en toetsscore van student
- Hebben hoogscorende studenten 'm goed gemaakt en slecht scorende studenten 'm slecht gemaakt?
- Waarde tussen -1 en 1
- R_{ir} is gecorrigeerde R_{it} : eigen score telt niet mee
- een negatieve R_{ir} : wees alert, hier kan iets mis zijn! (goede studenten denken: strikvraag!)

Rir waarde	Kwalificatie (Ebel/Friskie)
0,35 en hoger	Zeer goed
0,25 tot 0,35	Goed
0,15 tot 0,25	Voldoende
< 0,15	Slecht

Actie: vraag kritisch bekijken, op basis hiervan vraag al dan niet eruit halen

Wat doe ik er mee?

(3) Toets-analyse Betrouwbaarheid

KR20 / Cronbach's α = betrouwbaarheid v/d toets

- Interne consistentie (homogeniteit) toetsvragen
- Meten vragen even goed hetzelfde?
- Waarde tussen 0 en 1

Cronbach's α is een manier om vast te stellen of meerdere items samen één schaal mogen vormen. Dit wordt getoetst op basis van de onderlinge correlatie van de verschillende items

1. hoe gelukkig ben jij in de ochtend?	}	hoe gelukkig ben jij?
2. hoe gelukkig ben jij in de middag?		
3. hoe gelukkig ben jij in de avond?		

Wat doe ik er mee?

(3) Toets-analyse Betrouwbaarheid

KR20 / Cronbach's α = betrouwbaarheid v/d toets

- Interne consistentie (homogeniteit) toetsvragen
- Meten vragen even goed hetzelfde?
- Waarde tussen 0 en 1
- Oorzaak lage α : geclusterde vragen, heterogene vragen, relatief weinig vragen

KR20/ α -waarde	Kwalificatie (summatief)
0,90 en hoger	Zeer goed
0,80 – 0,90	Goed
0,70 – 0,80	Voldoende
< 0,70	Slecht / Middelmatig

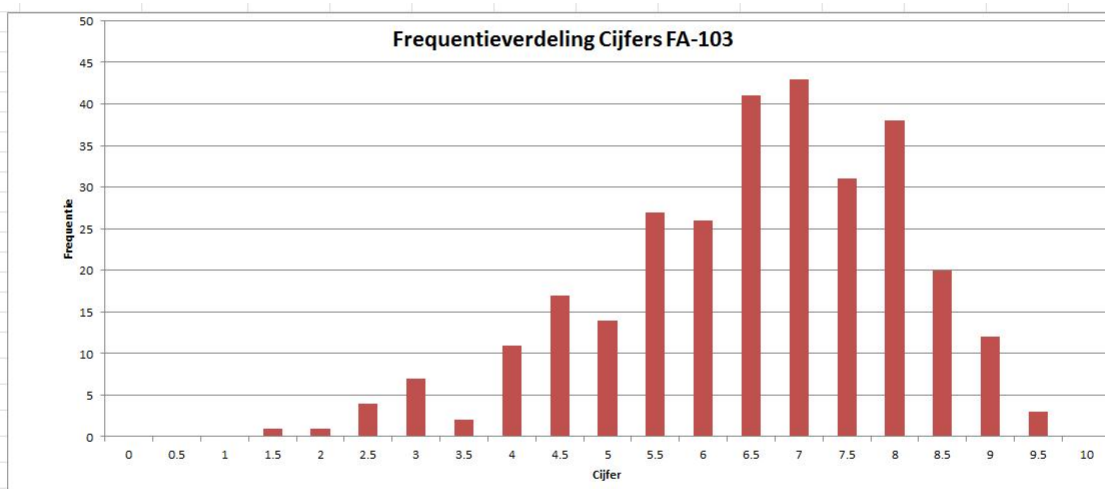
Actie: toets kritisch bekijken, en dan.....

Wat doe ik er mee?

(4) Toets-analyse Slagingspercentages en frequentieverdeling

298	Totaal aantal studenten		
237	voldoende	80	%
61	onvoldoende	20	%
10	is hoogste punt		
15	lager dan 4.0 (geen aanv. toetsing)		
14	tussen 5,0 en 5,5 voor eindtoets		

0	0	0
0.5	0	0
1	0	0
1.5	1	1
2	2	1
2.5	6	4
3	13	7
3.5	15	2
4	26	11
4.5	43	17
5	57	14
5.5	84	27
6	110	26
6.5	151	41
7	194	43
7.5	225	31
8	263	38
8.5	283	20
9	295	12
9.5	298	3
10	298	0



Wat doe ik er mee?

Voorbeelden (Huid en Dermatica):



Universiteit Utrecht

0
0
FA 304
jan 2011

Aantal deelneme	121	Cronbach's alfa	0,499
Aantal vragen	18	Standaardmeetfout	6,21

	min	max	variantie	standaard deviatie	moeilijkheidsgraad	item-test correlatie	item-rest correlatie
toets	22	68	76,90	8,77	0,55	(R _{it})	(R _{ir})
vr 1a	3	5	0,21	0,46	0,92	0,46	0,42
vr 1b	0	5	1,64	1,28	0,60	0,37	0,23
vr 1c	0	5	3,19	1,79	0,19	0,33	0,13
vr 1d	0	5	2,41	1,55	0,47	0,34	0,17
vr 2b	0	5	2,54	1,59	0,23	0,29	0,11
vr 2c	0	5	2,45	1,57	0,79	0,41	0,25
vr 3a	0	8	2,28	1,51	0,52	0,40	0,24
vr 3b	0	3	1,00	1,00	0,75	0,29	0,19
vr 3c	0	1	0,21	0,46	0,69	0,31	0,26
vr 3d	0	3	1,84	1,36	0,42	0,34	0,19
vr 3e	0	5	1,72	1,31	0,66	0,35	0,21
vr 4a	0	5	4,28	2,07	0,23	0,21	-0,03
vr 4c	0	5	1,68	1,29	0,65	0,24	0,10
vr 4d	0	5	2,26	1,50	0,42	0,24	0,07
vr 5a	2	6	0,92	0,96	0,80	0,48	0,39
vr 5b	0	4	2,86	1,69	0,43	0,18	-0,01
vr 5c	0	2	0,73	0,85	0,53	0,08	-0,02
vr 5d	0	8	8,42	2,90	0,54	0,64	0,38

Wat valt op:

Lage Cronbach's alfa

1a: makkelijk echter R_{ir} is goed

1c en 2b: moeilijk en lage R_{ir}

4a-c en 5b-c: zeer lage R_{ir}

Toetsmatrijs:

Niet alle leerdoelen komen aan bod

Meeste vragen op slechts

één begripsniveau

(tentamen januari 2011)

Wat doe ik er mee?



Universiteit Utrecht

farmaceutische Wetenschappen

FA 304

40909

31-1-2012

Aantal deelneme 120

Aantal vragen 24

Cronbach's alfa 0,701

Standaardmeetfout 5,99

	min	max	variantie	standaard deviatie	moeilijk- heidsgraad	item-test correlatie	item-rest correlatie
toets	29	88	119,87	10,95	0,60	(Rit)	(Rir)
vr 1a	0	5	2,25	1,50	0,62	0,34	0,21
vr 1b	0	8	2,89	1,70	0,70	0,41	0,27
vr 1c	0	3	1,27	1,13	0,68	0,44	0,35
vr 1d	0	4	2,18	1,48	0,77	0,46	0,35
vr 2a	1	4	0,56	0,75	0,78	0,35	0,29
vr 2b	0	4	2,59	1,61	0,46	0,48	0,36
vr 2c	0	4	1,00	1,00	0,75	0,32	0,24
vr 2d	0	4	0,53	0,73	0,85	0,26	0,20
vr 2e	0	4	3,11	1,76	0,62	0,37	0,22
vr 3a	0	4	0,73	0,85	0,25	0,22	0,15
vr 3b	0	5	0,99	1,00	0,66	0,26	0,18
vr 3c	0	7	3,02	1,74	0,62	0,30	0,14
vr 3d	0	3	1,23	1,11	0,44	0,36	0,27
vr 4a	0	4	2,60	1,61	0,46	0,55	0,44
vr 4b	0	3	0,96	0,98	0,46	0,40	0,32
vr 4c	0	3	2,06	1,43	0,44	0,45	0,34
vr 4d	3	3	0,00	0,00	1,00	#DEEL/0!	#DEEL/0!
vr 4e	1	4	0,51	0,72	0,75	0,30	0,24
vr 4f	0	3	1,23	1,11	0,18	0,29	0,19
vr 5a	0	4	2,89	1,70	0,38	0,44	0,31
vr 5b	0	4	2,20	1,48	0,68	0,34	0,21
vr 5c	0	4	1,37	1,17	0,54	0,39	0,30
vr 5d	0	4	0,76	0,87	0,58	0,28	0,21
vr 5e	0	4	2,45	1,57	0,81	0,35	0,22

Aanpassing:

Herformulering leerdoelen

Taxonomie van Bloom is consequent doorgevoerd

Toetsmatrijs:

Betere spreiding van vragen over diverse leerdoelen

Beter spreiding over de begripsniveaus.

Gevolg:

Cronbach's alfa: 0.701

 R_{ir} -waarden zijn goed

Wat doe ik er mee?

Voorbeelden (Farmacotherapie van het CZS):

Onderwerp	# deelvragen	Docent
anti-epileptica	2	A
dementie	2	A
interacties, switchen AD	2	A
migraine	2	B
slaap, benzo's	2	B
depressie: therapieresist.	2	B
Parkinson	2	C
antidepressiva	2	C
depressie: bipolair	1	C
antipsychotica	3	D
angststoornissen	2	D

Wat doe ik er mee?

	min	max	variantie	standaard deviatie	moeilijk- heidsgraad	item-test correlatie	item-rest correlatie
toets	0,3	74,5	333,82	18,27	0,53	(Rit)	(Rir)
vr 1	0	5	4,62	2,15	0,63	0,41	0,31
vr 2	1	4	0,69	0,83	0,72	0,43	0,39
vr 3	1	5	1,71	1,31	0,60	0,59	0,54
vr 4	2	4	0,43	0,65	0,60	0,42	0,39
vr 5	1	5	1,24	1,11	0,53	0,37	0,31
vr 6	0	5	2,35	1,53	0,65	0,18	0,10
vr 7	0	3	0,67	0,82	0,48	0,48	0,44
vr 8	0	3	1,26	1,12	0,37	0,60	0,56
vr 9	0	6	3,64	1,91	0,38	0,49	0,40
vr 10	0	5	3,27	1,81	0,51	0,26	0,16
vr 11	3	6	0,80	0,90	0,74	0,20	0,16
vr 12	1	6	2,27	1,51	0,39	0,05	-0,03
vr 13	0	6,5	3,49	1,87	0,42	0,61	0,54
vr 14	1	3	0,50	0,70	0,74	0,33	0,30
vr 15	0	3	0,52	0,72	0,68	0,66	0,64
vr 16	1	5	2,14	1,46	0,50	0,75	0,71
vr 17	0	2	0,69	0,83	0,70	-0,06	-0,10
vr 18	0	2	0,52	0,72	0,17	-0,04	-0,08
vr 19	0	3	0,60	0,77	0,60	0,28	0,24
vr 20	0	3	1,27	1,13	0,62	0,62	0,58
vr 21	0	3	1,35	1,16	0,42	-0,02	-0,08
vr 22	0	3	1,21	1,10	0,58	0,70	0,67
vr 23	0	3	1,40	1,18	0,47	0,46	0,40
vr 24	0	5	2,03	1,42	0,44	0,17	0,09
vr 25	0	5	3,41	1,85	0,37	0,40	0,31



Vraag 18: deze vraag was niet goed geformuleerd

Hoe kom ik meer te weten?



Boeken

- Suskie, L. – Assessing student learning 2nd ed (2009)
- De Gruijter, D.N.M. – Toetsing en toetsanalyse (2008)
<http://media.leidenuniv.nl/legacy/toetsing-en-toetsanalyse.pdf>
- Molkenboer, H., Soeting, J. – Toetsontwikkeling in de praktijk.
Hoe maak ik goede vragen en toetsen? (2004)
- Milius, J. – Schriftelijk tentamineren. Een draaiboek voor docenten in
het hoger onderwijs (2007)
- Magnee, M. – Toetsen en beoordelen (1997)

Websites

- Toetsen op school (website met interessante documenten en links)
http://www.cito.nl/onderzoek%20en%20wetenschap/toetsen_op_school.aspx
- Toetswijzer
<http://toetswijzer.kennisnet.nl/>

Intercollegiaal contact

- Toetsadviescommissie Farmaceutische Wetenschappen

Dank voor uw aandacht
en “happy testing”





Neem je eigen toets en doe een toetsanalyse in Excel:

- Bepaal de p-waarde, R_{ir} , R_{it} en Crohnbach's α .
- Bepaal het slagingspercentage en de frequentieverdeling.