

HANDREIKING VRAAGANALYSE VOOR FSW EXAMINATOREN

OPBOUW HANDREIKING

Beknopt stappenplan voor vraaganalyse.....	1
Literatuur.....	2
Praktische informatie.....	2
Vraaganalyse: principes en richtlijnen.....	4
Stap 1. Beoordeel de kwaliteit van de toetsvragen.....	4
Stap 2. Besluit welke vragen behouden, aangepast en/of niet overgezet moeten worden.....	6
Ondersteuning en Contact.....	9

NB: deze handreiking richt zich enkel op de analyse van toetsvragen in Remindo, t.b.v. een goede overstap naar het nieuwe toetsplatform Ans. De handreiking richt zich niet op toetsanalyse, en ook niet op de overige toetsvormen (schriftelijke en mondelinge tentamens en eindwerkstukken).

Een meer uitgebreide handreiking voor examinatoren, inclusief instructie over deze toetsvormen, valt in Q3-Q4 van 2024 te verwachten.

BEKNOPT STAPPENPLAN VOOR VRAAGANALYSE

Hieronder het beknopte stappenplan, zoals in de handleiding uitgebreider uitgelegd.

1. **Identificeer vragen in de Excel/Remindo die evt. aan herziening toe zijn**, o.b.v. oranje en rode p²- en rir-waarden en door studenten gekozen antwoordopties kan je je een oordeel vormen over de kwaliteit van de toetsvragen. Is de vraag (veel) te makkelijk -of juist te moeilijk-, is het antwoordmodel onjuist, betreft het een strikvraag, is de formulering van de vraag en/of de antwoordopties vaag, etc.? Dan is actie van de examinerator gewenst, zie stap 2b-c.
2. **Kies per vraag of je deze wil behouden mét of zonder aanpassingen**, of dat je de vraag niet overgezet wil hebben naar Ans.
 - a. Behoud van de vraag zonder aanpassingen: geen actie, tenzij de vraag een andere status heeft dan 'goedgekeurd'. In dat geval is goedkeuren van de vraag benodigd.

- b. De vraag niet overzetten naar Ans: geef dit aan in het Excel-bestand volgens de instructies in de mail. Evt. kan je (met behulp van je toetsmatrijs¹) direct een nieuwe vraag in Remindo maken. Er wordt echter aangeraden om dit na de migratie in Ans te doen.
- c. Behoud van de vraag mét aanpassingen: afhankelijk van je analyse bij stap 1 zal je het antwoordmodel en/of de toetsvraag en/of de antwoordopties aanpassen in Remindo. Je kan hierbij gebruik maken van de in dit document opgenomen richtlijnen.

LITERATUUR

Bij het opstellen van deze handreiking zijn de volgende bronnen gebruikt:

- Checklist meerkeuzevragen. LLInC.
- Kwaliteitsborging toetsing – Een handreiking voor Examencommissies
- MC Exam Workshop Board of Examiners
- Overzicht p-rir-waarden (LLInC, Universiteit Leiden)
- Paragin informatie over toetsanalyse in Remindo: [Toetsanalyse in RemindoToets - Paragin](#)
- Tips bij toetsen (juli 2016). Geschreven door: Maarten Bergwerff, Marlous Dekker, Floris van Blankenstein, Daan Romein, Marie Klaren. *Faculteit der Geesteswetenschappen*, Universiteit Leiden.
- Toetsen in het hoger onderwijs (2017). Van Berkel, Bax en Joosten. Hoofdstuk 2.4 Psychometrische analyse: het berekenen van indicatoren.
- Toetsing en toetsanalyse (2008). Geschreven door: D.N.M. de Gruijter. *ICLON*, Universiteit Leiden.

Meer informatie over de praktische aspecten van toetsing via Remindo bij de FSW zijn hier te vinden: [Timeline - Digital Assessment with Remindo \(FSW\) | Rise 360 \(articulate.com\)](#)

PRAKTISCHE INFORMATIE

Hieronder een aantal praktische aanwijzingen over het bekijken van vragenbanken:

HOE BEKIJK IK (DE PSYCHOMETRISCHE INFORMATIE OVER) MIJN VRAGEN?

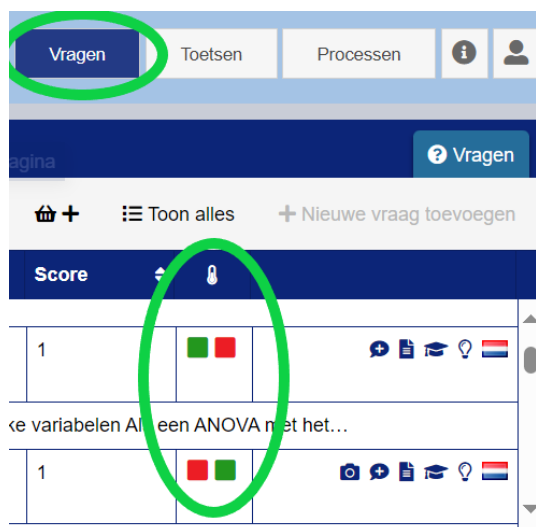
Het Excel-bestand met vragen bevat de benodigde informatie om de kwaliteit van de vragen te beoordelen, je kan hierin ook aangeven als je vragen niet wil overzetten.

¹ Met toetsmatrijs wordt het overzicht/de tabel bedoeld waarin je leerdoelen en leeruitkomsten met elkaar verbindt, incl. weging van de leerdoelen en bijbehorend het aantal vragen dat gesteld wordt per leerdoel. Zie voor een voorbeeld: [voorbeeld-toetsmatrijs.pdf \(ugent.be\)](#)

Als je (op basis van het Excel-bestand) een of meer vragen aan wil passen, dan ga je naar de vragenbank in Remindo:

- Log in bij de Beheeromgeving van Remindo: <https://universiteit leiden.remindotoetsmanager.nl/>.
- Klik rechtsboven op het tabje 'Vragen'.
- Bekijk dan psychometrische informatie bij de 'Indicatoren', te zien als twee vierkantjes (zie afbeelding 1). Als één of beide vierkantjes rood gekleurd zijn, volg dan stap 2 van het stappenplan in deze handleiding.

AFBEELDING 1. INDICATOREN IN DE VRAGENBANK



VRAAGANALYSE: PRINCIPES EN RICHTLIJNEN

Hieronder volgt een stappenplan voor hoe de toetsvragen door examinatoren bekeken kunnen worden tijdens de overstap van Remindo naar Ans. Daarbij kijkt de examiner naar de individuele toetsvragen, daarna bepaalt de examiner welke actie nodig is t.a.v. de individuele toetsvragen.

STAP 1. BEOORDEEL DE KWALITEIT VAN DE TOETSVRAGEN

Eerst bekijk je de individuele toetsvragen o.b.v. de psychometrische informatie in het Excel-bestand of in Remindo. De vragen met de slechtste waarden (negatieve *r_{ir}* of zeer lage *p'*) kunnen als eerste bekeken worden, met de vraag of herziening nodig is in het achterhoofd. Daarbij kan onderstaand stappenplan gevolgd worden.

WELKE VRAGEN ZIJN EVT. AAN HERZIENING TOE?

[RemindoToets](#) markeert een aantal vragen die afwijkende waarden laten zien, deze informatie is ook toegevoegd aan het Excel-bestand. Let bij het bekijken van de informatie op **zeer hoge of lage *p'*-waarden en/of negatieve *r_{ir}*-waarden**.

- Je kijkt naar de moeilijkheid van de vraag middels de ***p'*-waarden**: welke vragen zijn zeer goed of slecht gemaakt? De *p'*-waarde geeft de proportie correcte antwoorden voor een vraag aan, deze waarde ligt tussen 0 (alle studenten hadden 0 punten op de vraag) en 1 (alle studenten hadden het maximaal aantal punten op de vraag). Met andere woorden: hoe hoger *p'*, hoe gemakkelijker de vraag². Let op: ook eigenschappen van de student en/of de toets hebben invloed op de *p'*-waarde: vragen aan het eind van een toets leveren vaak een lagere waarde op, ook toetsen met veel herkansers kunnen leiden tot lage *p'*-waardes.
- Het onderscheidend vermogen van de vraag bekijk je middels de ***r_{ir}*-waarde** (item-restcorrelatie). Dit is een correlatie tussen de scores op de vraag en de scores op alle andere vragen binnen een toets.
 - Een **positieve correlatie betekent een goed onderscheidend vermogen**: dan scoren studenten die de betreffende vraag goed hebben gemaakt ook goed op de andere vragen en scoren studenten die de vraag slecht hebben gemaakt ook slecht op de andere vragen. De vraag onderscheidt dus goed tussen goede en slechte studenten.
 - Is de correlatie echter negatief, dan scoren studenten die de vraag goed hebben gemaakt juist slecht op de overige vragen en scoren studenten die de vraag slecht hebben gemaakt

² “Een *p'*-waarde van 0,39 geeft bijvoorbeeld aan dat de kandidaten gemiddeld 39% van het maximum aantal punten voor deze vraag hebben gescoord. Dat zou kunnen duiden op een te moeilijke of niet-relevante vraag. Maar dat hoeft niet zo te zijn. Het is belangrijk dat er genoeg variatie in een examen zit. Een vraag met een *p'*-waarde van 0,39 is bijvoorbeeld prima tussen vragen met een *p'*-waarde van bijvoorbeeld 0,7.” (Paragin, [Toetsanalyse in RemindoToets - Paragin](#))

juist goed op de overige vragen. Een **negatieve correlatie betekent een laag onderscheidend vermogen** van de vraag – dan kan er wat mis zijn met de vraag.³

- Wanneer je kijkt naar de rir-waarden, kun je met behulp van de interactie-analyse van RemindoToets ook direct de a- en rar-waarden bekijken. Zo zie je welke antwoordopties vaak gekozen zijn. En kun je een mogelijke oorzaak van de afwijkende waarden vinden.
- Ten derde probeer je de scores te verklaren, b.v. door jezelf onderstaande vragen te stellen:
 - Toetst de vraag algemene kennis (i.t.t. een specifiek geformuleerd leerdoel)?
 - Is de vraag te makkelijk?
 - Kenden studenten de vraag al?
 - Is de vraag/zijn de antwoordopties voor meerdere interpretaties vatbaar?
 - Betreft het toeval (bij een klein aantal studenten)?
 - Bij een open vraag: voorkomt het antwoordmodel beoordelaarseffecten (voldoende)?
 - Bevat het antwoordmodel een fout? Zie hier instructies om dit aan te passen: [Answer key and changing the response model - Digital Assessment with Remindo \(FSW\) | Rise 360 \(articulate.com\)](#)

Zie ook Tabel 1 om de p'- en rir-waarden te interpreteren⁴. NB dit overzicht is gebaseerd op (vierkeuze) multiple choice-vragen, voor open vragen geldt een ondergrens van p'= 0,25.

TABEL 1. INTERPRETATIE P'- EN RIR-WAARDEN, GEÏNFORMEERD DOOR BERKEL, BAX EN JOOSTEN (2017)

	Rir <0,15/negatief	Rir tussen 0,15-0,25	Rir is hoger dan 0,25
P' lager dan 0,4 (moeilijke vraag)	Kans dat dit een slechte vraag is: Studenten die goed scoren op de toets, scoren juist slecht op deze vraag. <i>Inspecteer vraag en antwoordmodel. Overweeg antwoord-model aan te passen of verwijder vraag als inspectie daar aanleiding toe geeft.</i>	Kans dat dit een slechte vraag is (e.g., mismatch leerdoelen, detailvraag, ambigue vraag, meerdere antwoorden juist) <i>Inspecteer vraag en antwoordmodel. Overweeg antwoordmodel aan te passen of verwijder vraag als inspectie daar aanleiding toe geeft.</i>	Deze vraag maakt waarschijnlijk vooral onderscheid tussen de best scorende studenten. <i>Vraag kan in principe gehandhaafd worden. Let op: wanneer er relatief veel van dit soort vragen in de toets zitten dan is dat mogelijk problematisch.</i>
P' tussen 0,4 en 0,8 (gemiddelde vraag)	Grote kans dat dit een slechte vraag is: Studenten die goed scoren op de toets, scoren juist slecht op deze vraag.	De vraag lijkt slechts een beperkt onderscheidend vermogen te hebben: <i>Inspecteer vraag en antwoordmodel. Overweeg</i>	Dit lijkt een goede vraag te zijn.

³ Verbatim overgenomen uit 'Tips bij toetsen' (FGW)

⁴ Overzicht p-rir-waarden (LLInC) en 'Toetsen in het hoger onderwijs' (2017, van Berkel, Bax & Joosten)

	<i>Inspecteer vraag en antwoordmodel. Overweeg antwoord-model aan te passen of verwijder vraag als inspectie daar aanleiding toe geeft.</i>	<i>antwoord-model aan te passen als inspectie daar aanleiding toe geeft.</i>	<i>Geen aanleiding voor nadere inspectie. De vraag kan gehandhaafd worden.</i>
P' hoger dan 0,8 (makkelijke vraag)	Kans dat dit een slechte vraag is: Studenten die goed scoren op de overall toets, scoren juist slecht op deze relatief makkelijke vraag. <i>Inspecteer vraag. Verwijder de vraag alleen als er bij inspectie duidelijk reden toe is.</i>	Kans dat dit een slechte vraag is: De vraag is makkelijk en discrimineert niet (e.g., vraag toetst algemene kennis, de afleiders zijn niet plausibel). <i>Inspecteer vraag. Verwijder de vraag alleen als er bij inspectie duidelijk reden toe is.</i>	Deze vraag maakt waarschijnlijk vooral onderscheid tussen de slechtst scorende studenten. <i>Vraag kan in principe gehandhaafd worden. Let op: Wanneer er relatief veel van dit soort vragen in de toets zitten dan is dat mogelijk problematisch</i>

DE KWALITEIT VAN EEN VRAAG HANGT MEDE AF VAN DE CONTEXT

Uiteindelijk staat centraal of een vraag vanuit de vooraf opgestelde toetsmatrijs, dus met de leerdoelen in gedachten, voldoende kwaliteit heeft volgens de examinerator. Soms kan het vanuit didactisch oogpunt nuttig zijn om een relatief zeer makkelijke, of juist zeer moeilijke vraag te stellen – en blijft een vraag ondanks een laag onderscheidend vermogen toch liever behouden.

- Wellicht toetst de makkelijke vraag een leerdoel dat herhaaldelijk aan bod is gekomen, en wat studenten moeten hebben behaald als zij zelfstandig hebben gestudeerd. Als de vraag toch té makkelijk lijkt is het mogelijk om het leerdoel op een hoger niveau te toetsen, of om de vraag minder zwaar te laten wegen.
- Mogelijk toetst de moeilijke vraag een aspect van de stof dat is toegevoegd als uitdaging. Overweeg daarbij of de stof voldoende (toegankelijk) is behandeld in het onderwijs.

STAP 2. BESLUIT WELKE VRAGEN BEHOUDEN, AANGEPAST EN/OF NIET OVERGEZET MOETEN WORDEN

Bepaal, na de analyse van de toetsvragen, welke van de onderstaande besluiten van toepassing is:

- a. **Als je een toetsvraag (met status 'goedgekeurd') zonder aanpassing wil behouden hoef je verder niets te doen.** De vraag wordt door het DT-team overgezet naar Ans.
 - NB vragen met een andere status (b.v. Need repair/concept) zijn niet opgenomen in het Excel-bestand. Het kan lonen om in de beheeromgeving van Remindo naar deze vragen te kijken. Als je de vraag over wil zetten naar Ans is het nodig om de vraagstatus aan te passen naar 'goedgekeurd', zie: [Managing your question bank - Digital Assessment with Remindo \(FSW\) | Rise 360 \(articulate.com\)](#)

- b. **Als je een/meerdere toetsvra(a)g(en) niet overgezet wil hebben naar Ans**, geef dit aan in het Excel-bestand dat je via de mail ontvangt van het DT-team. Het DT-team zorgt ervoor dat deze vra(a)g(en) niet overgezet wordt/worden naar je nieuwe vragenbank in Ans. T.z.t. betekent dit mogelijk dat je (een) nieuwe tentamenvra(a)g(en) zal moeten maken met behulp van je toetsmatrijs.
- c. **Als je een/meerdere toetsvra(a)g(en) aan wil passen**, zie de suggesties bij ‘Kwaliteit van een multiple choice vraag verbeteren’ en/of ‘Kwaliteit van een open vraag verbeteren’. Hier worden suggesties gegeven om de formulering van de vraag en/of antwoordopties direct in Remindo aan te passen.
 - o **Als je het antwoordmodel aan wil passen en/of direct wijzigingen door wil voeren in Remindo**, zie [Adding question\(s\), approving and editing - Digital Assessment with Remindo \(FSW\) | Rise 360 \(articulate.com\)](#).

KWALITEIT VAN EEN MULTIPLE CHOICE VRAAG VERBETEREN

Een aantal richtlijnen voor het beoordelen van de kwaliteit van een multiple choice vraag zijn hieronder beschreven. Er is gebruik gemaakt van de ‘Checklist meerkeuzevragen’ (LLInC), de MC training voor Examencommissies (Politieke Wetenschap) en van de Tips voor toetsing (FGW).

Het schrijven van de stam (=context en de vraag zelf)

- Geef duidelijke aanwijzingen in de stam (waarom, wanneer, wat, hoe, enz.). De student zou de vraag ook zonder de antwoordopties moeten kunnen beantwoorden.
- Beschrijf de context en het centrale idee van de vraag.
- Houd de stam kort en bondig.
- Gebruik positieve bewoordingen, vermijd zoveel mogelijk woorden als NIET, GEEN of BEHALVE. Als het noodzakelijk is om negatieve woorden te gebruiken, zorg dan dat deze woorden opvallen door bijvoorbeeld hoofdletters te gebruiken of deze te onderstrepen.

Het schrijven van de alternatieven (=de antwoordopties)

- Zorg dat één antwoord echt juist is – vermijd overlappingen in de alternatieven. Vermijd ook de formulering ‘geen (of alle) van voorgaande alternatieven’.
- Houd antwoordopties onafhankelijk; er moet geen overlap tussen antwoordopties zijn.
- Zorg dat antwoordopties inhoudelijk en grammaticaal vergelijkbaar zijn.
 - o Vermijd dubbele ontkenningen.
 - o Laat alternatieven grammaticaal aansluiten bij de stam
 - o Gebruik consistent dezelfde terminologie
 - o Vermijd afkortingen
- Zorg dat de lengte van antwoordopties ongeveer hetzelfde is.
- Wees terughoudend in het gebruik van bewoordingen als ‘geen enkele van bovenstaande...’ of ‘alle bovenstaande...’
- Formuleer antwoordopties positief; vermijd negatieve formuleringen zoals NIET.
- Geef geen aanwijzingen richting het juiste antwoord, zoals:
 - o Specifieke aanduidingen (altijd, nooit, helemaal, absoluut)

- Associaties met tekst in de stam; woorden die identiek zijn aan woorden in de stam of die lijken op woorden in de stam
 - Grammaticale inconsistenties tussen antwoordopties
 - Meerdere samenhangende antwoordopties
 - Opvallend afwijkende antwoordoptie
 - Absurde antwoordoptie
- Zorg dat alle afleiders in principe plausibel zijn.
 - Zorg dat de vraag niet te beantwoorden is met algemene kennis, maar echt is gerelateerd aan de in het onderwijs behandelde stof. Voeg waar verwezen wordt naar een opinie, of waar een citaat wordt gebruikt, de referentie naar de bron/het artikel/het boek toe.
 - Gebruik typische denkfouten die studenten maken om afleiders te schrijven (die kun je verzamelen tijdens het onderwijs...)
 - Varieer met de plek van het juiste antwoord of plaats antwoorden altijd in dezelfde logische volgorde (bijvoorbeeld alfabetisch).

KWALITEIT VAN EEN OPEN VRAAG VERBETEREN⁵

Een aantal richtlijnen voor het beoordelen van de kwaliteit van een open vraag zijn hieronder beschreven.

Taalgebruik

- Is de vraag grammaticaal juist geformuleerd?
- Bevat de vraag ingewikkelde zinsconstructies?
- Bevat de vraag een dubbele ontkenning?
- Bevat de vraag onnodige tussenvoegsels?
- Is de vraag onnodig negatief gesteld?
- Kan de formulering van de vraag aanleiding geven tot misverstanden?
- Kan de vraag door een verschuiving van de klemtoon een andere betekenis krijgen?

Informatie

- Bevat de vraag voldoende informatie voor beantwoording?
- Geeft de vraag voldoende informatie over de gewenste lengte en vorm van het antwoord?
- Is duidelijk dat een antwoord moet worden verklaard/toegelicht?
- Zijn informatie en probleemstelling duidelijk gescheiden?

Relevantie

- Kan de vraag met gebruik van andere dan de beoogde kennis worden beantwoord?
- Suggereert de vraag een probleem dat niet aan de orde is?

⁵ Verbatim overgenomen uit 'Tips bij toetsen' (FGW), gebruikmakend van de bron: CITO - Het construeren van open vragen http://www.toetswijzer.nl/html/toetsenopschool/hoofdstuk7_herzien.pdf

- Bevat de vraag onbedoelde hints voor de beantwoording?
- Is de moeilijkheidsgraad van de vraag respectievelijk van de gehele toets acceptabel?
- Wordt de moeilijkheidsgraad van de vraag onnodig verhoogd door irrelevante gegevens?

Context

- Is het gebruik van context (tekeningen, grafieken, teksten, afbeeldingen) functioneel?
- Is de context duidelijk en correct weergegeven?
- Bevat de context irrelevante informatie?

Presentatie

- Zijn de vragen en de vraagonderdelen duidelijk van elkaar te onderscheiden?
- Is de nummering van de vragen logisch en overzichtelijk?
- Zijn de conventies wat betreft spelling, symboolgebruik, interpunctie, e.d. in acht genomen?
- Zijn de verwijzingen in de vraag naar teksten, tekeningen, e.d. juist?

ONDERSTEUNING EN CONTACT

Bij technische of inhoudelijke vragen of opmerkingen over toetsing kan je contact opnemen met het Digital Testing team via digitaltesting@fsw.leidenuniv.nl.

Ook worden er ter ondersteuning van examinatoren workshops georganiseerd door SOLO en LLInC:

- 4 april 11.00-12.00 How to improve your exam (questions) with psychometric data
- 4 april 12.00-13.00 Mastering the art of test question design
- 4 juni 11.00-12.00 How to improve your exam (questions) with psychometric data
- 4 juni 12.00-13.00 Mastering the art of test question design

Je kan je hiervoor inschrijven via: <https://www.medewerkers.universiteitleidennl/vr/sociale-wetenschappen/solo/informatie-voor-docenten/develop-your-teaching-skills--at--fsw/meld-je-aan>