

Literatuur

Psychologie | Rijksuniversiteit Groningen

2026 - 2027

RGPS-608-085 | €16,95

Tentamengericht | Overzichtelijke structuur

Sinds 1994 | Beoordeeld met een 8,2



| Slim
Academy

1 maand lang *gratis* samenvattingen ontvangen?

Maak het studeren makkelijker, nu de 1ste maand gratis bij het afsluiten van een jaarabonnement!



De voordelen:

- Beschikbaar op papier en/of digitaal.
- **Goedkoper** dan het bestellen van pakketten of losse producten.
- Voeg **oefenvragen en flashcards** toe voor de goedkoopste prijs.
- Altijd en overal leren met de onze app.
- De **meest complete studiehulp** als eerst beschikbaar voor abonnees!
- Je hoeft je het hele jaar geen zorgen te maken over je materialen.

Hoe werkt het:

- 🌐 Ga naar www.SlimAcademy.nl
- 🔍 Zoek jouw studie in ons keuzemenu
- ⊕ Voeg het abonnement toe aan je winkel mandje
- 🏷️ Voer de code '**SLIMSTART**' in om de eerste maand gratis te krijgen



Slim
Academy

Welkom bij Slim Academy!

Voor je ligt de **literatuur samenvatting** van het vak **Testtheorie en Testgebruik**, dat je volgt in blok 1 van het tweede jaar van de bachelor Psychologie aan de Rijksuniversiteit van Groningen. In dit vak leer je kennis over verschillende soorten testen en hoe je deze verschillende soorten testen kan gebruiken.

Met Slim Academy willen we jou helpen slimmer te studeren. Dat doen we door de kern van de leerstof helder samen te vatten, complexe onderwerpen te versimpelen en je de juiste hulpmiddelen te geven om actief te oefenen. Voor een optimale voorbereiding op je tentamen raden we je aan gebruik te maken van ons volledige aanbod. Voor dit vak zijn de onderstaande producten beschikbaar.

Product	Inhoud	Content	Versies
Samenvatting Literatuur	De belangrijkste begrippen en theorieën met een heldere uitleg die is voorzien van voorbeelden over alle opgegeven tentamenstof.	Week 1 t/m 6	✓ App ✓ Papier
Samenvatting Hoorcolleges	Een overzicht van de belangrijkste punten die tijdens de hoorcolleges zijn behandeld, zodat je niets mist van de stof die aan bod is gekomen.	Week 1 t/m 6	✓ App ✓ Papier
Oefenvragen	Interactief oefenen in de leeromgeving. Zo kun je nagaan of je de stof volledig beheerst en weet je beter wat je kunt verwachten op het tentamen.	• 7 sets; • 250+ vragen	✓ App
Flashcards	Kaarten die je helpen om definities, schema's en kernbegrippen snel te herhalen, zodat je op een actieve manier leert en de stof sneller en beter onthoudt.	• 7 sets; • 150+ kaarten	✓ App

Contact

Heb je vragen over ons aanbod, de levering van je samenvattingen of iets anders? Neem gerust contact met ons op via **06 81 60 19 60** of stuur een e-mail naar **klantenservice@slimacademy.nl**.

Feedback geven

Kan je ons vertellen wat er mis is met de geselecteerde tekst?

Verstuur feedback

Fout gevonden?

Geef het door en wij zorgen dat de samenvatting snel wordt bijgewerkt!

Hoe werkt het?

- Selecteer de tekst
- Klik op het vlaggetje
- Verstuur je feedback

Deze samenvatting is alleen voor persoonlijk gebruik en mag niet worden gekopieerd of verspreid. De tekst is opgesteld naar eigen inzicht van de auteur en bedoeld als ondersteuning bij het studeren. Raadpleeg altijd de originele bronnen.

Digitale en papieren updates

Aan het begin van het blok is nog niet alle vakinformatie beschikbaar, daarom vullen wij de samenvatting zo snel mogelijk aan zodra onze auteurs de nieuwe lesstof hebben ontvangen. Zo blijft de inhoud actueel, betrouwbaar en van hoge kwaliteit. Je hoeft dus niet te wachten tot het einde van het blok: vanaf **21 september 2026** is de volledige samenvatting **digitaal** beschikbaar in de app.

Later in het blok verzenden we de **papieren samenvatting**. Deze wordt op **6 oktober 2026** verstuurd en bevat **week 1 t/m 6**. In het onderstaande overzicht zie je welke stof is opgenomen.

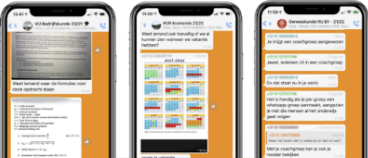
De verplichte literatuur komt uit het boek van

- *Testtheorie: inleiding in de theorie van de psychologische test en zijn toepassingen* (Vierde druk). Houten: Bohn Stafleu Van Loghum. Drenth, P.J.D. & Sijtsma, K. (2006). ISBN: 90 313 4747 7
- Timmerman, M. (2005). *Factor analysis*. Heymans Institute for Psychologie.

Stof	
Week 1	Hoofdstuk 2 en 3: Kenmerken van tests
Week 2	Hoofdstuk 4 en 5: Itemconstructie, analyse, normen en scoring
Week 3	Hoofdstuk 6: Betrouwbaarheid
Week 4	Hoofdstuk 8 en document factoranalyse: Begripsvaliditeit
Week 5	Hoofdstuk 7: Betrouwbaarheid
Week 6	Hoofdstuk 9: Beslissingen in testgebruik

JOIN DE WHATSAPP GROEP

- ✓ Chat met jouw mede-studenten
- ✓ Stel al jouw (studie)vragen aan onze studie-experts
- ✓ Krijg extra oefenvragen om jouw kennis te testen
- ✓ Krijg gratis voorbeeldsamenvattingen


We wensen je veel succes met studeren en het behalen van je tentamens!

Team Slim Academy

Inhoudsopgave

Welkom bij Slim Academy!	1
Digitale en papieren updates	2
Inhoudsopgave	3
Week 1: Kenmerken van tests	4
Hoofdstuk 2: De test	4
Hoofdstuk 3: Begrippen, onderscheidingen en indelingen	11
Week 2: Itemconstructie, analyse, normen en scoring	20
Hoofdstuk 4: Itemconstructie en het kwantificeren van reacties	20
Hoofdstuk 5: Testafname en het verwerken van testgegevens	26
Week 3: Betrouwbaarheid	34
Hoofdstuk 6: Betrouwbaarheid	34
Week 4: Begripsvaliditeit	44
Hoofdstuk 8: Betekenis en validiteit	44
Document: Factoranalyse	55
Week 5: Betrouwbaarheid	61
Hoofdstuk 7: Ontwikkelingen in testtheorie en testconstructie	61
Week 6: Beslissingen in testgebruik	74
Hoofdstuk 9: De test in het besluitvormingsproces	74
Uitgewerkt formuleblad	81

Week 1: Kenmerken van tests

Hoofdstuk 2: De test

Introductie

Dit hoofdstuk gaat over hoofdstuk 2 van het boek *Testtheorie, Inleiding in de theorie van psychologische test en zijn toepassingen*. In dit hoofdstuk staan de volgende onderwerpen centraal: de definitie, toepassingen en kenmerken van de psychologische test.

Definitie van een test

Testonderdelen

Een test bestaat uit testmateriaal, testformulieren en een testhandleiding. Het **testmateriaal** is sterk afhankelijk van het soort test. Het kan gaan om een vragenboekje, puzzels, een papier met inktvlekken of zelfs een discussie. Op de **testformulieren** worden de antwoorden of gegevens verzameld. Dit kan worden geïnterpreteerd, waardoor conclusies getrokken kunnen worden. De testhandleiding kan variëren van een dik boek tot slechts beknopte richtlijnen. Een goede **testhandleiding** bevat de volgende vier onderwerpen:

- De exacte **testinstructie** bevat alles wat betrekking heeft op de gang van zaken tijdens het testonderzoek, van de volgorde van de vragen tot hoe een proefleider waarschuwingen moet geven;
- De **verwerkingsprocedure** bestaat vooral uit richtlijnen voor het toekennen van scores aan de antwoorden;
- Met het gebruik van **normtabellen** kan een score vergeleken worden met de prestaties van normgroepen. Deze vergelijking vormt een voorwaarde voor de verdere interpretatie en evaluatie van de testcores;
- Bespreking van de **wetenschappelijke kwaliteit** van de test: het gaat hierbij om een indicatie van de betrouwbaarheid, de psychologische betekenis en de validiteit (in meerdere vormen) van de test.

Een omschrijving

De bedoeling van het testonderzoek is het opstellen van een voorspelling, beschrijving of classificatie met betrekking tot de onderzochte persoon. Het gaat daarbij vrijwel altijd om een vergelijking met anderen. De soort en de grootte van de vergelijkingsgroep hebben belangrijke gevolgen voor de conclusie over de onderzochte persoon. De psychologische test kan beschreven worden als 'een systematisch onderzoek van gedrag met behulp van speciaal gekozen vragen, met de bedoeling inzicht te krijgen in een psychologisch kenmerk van de onderzochte persoon in vergelijking met anderen'.

Testkenmerken

Er zijn zes kenmerken waarop een test meestal beter is dan een voorwetenschappelijk oordeel (een oordeel in de dagelijkse praktijk), namelijk:

- Standaardisatie;
- Efficiëntie;
- Normering;
- Validiteit
- Betrouwbaarheid;
- Objectiviteit.

Om een goede uitspraak over een testscore te kunnen doen, moet deze vergelijkbaar zijn met de testcores van anderen. De vergelijkbaarheid van testcores kan worden bereikt door het zoveel mogelijk standaardiseren van invloeden die op de testprestatie kunnen inwerken.

Standaardiseren is het uniform en gelijk houden van de testafname, de instructies, de hulpmiddelen en de scoringsprocedure voor iedereen, zodat verschillen in scores uitsluitend toe te schrijven zijn aan verschillen in de gemeten eigenschap. Standaardisatie is een ideaal en kan nagestreefd worden, maar nooit geheel bereikt worden, door onder andere een goede testhandleiding. Standaardisatie is noodzakelijk om te kunnen spreken over een 'test'.

Een test is **efficiënt**, omdat het speciaal is samengesteld om een schatting te maken van een eigenschap, bijvoorbeeld intelligentie. Men is niet afhankelijk van de beschikbaarheid van situaties waarin intelligent gedrag voorkomt. Ook worden storende invloeden zoveel mogelijk verwijderd bij het maken van een test.

De verschillen tussen testcores moeten zo exact mogelijk vastgesteld worden, daarom is de beschikbaarheid van normen een noodzakelijke voorwaarde voor een test. De generaliseerbaarheid van een norm wordt groter naarmate de representativiteit van de normgroep beter is. Het bepalen van de rangorde is de simpelste vorm van **normering**. Een enkele keer is er geen normgroep vereist, bijvoorbeeld wanneer er slechts een rangorde binnen een groep moet worden vastgesteld. Het gebruik van normering kan relevant zijn voor bijvoorbeeld de toelating voor een bepaalde opleiding. In dit geval werkt normering beter en nauwkeuriger dan gewoon een persoonlijke inschatting over of iemand geschikt is of niet. Bij het kijken naar hoe verschillende rangcores zich verhouden tot elkaar, kan men gebruikmaken van bijvoorbeeld een normale verdeling (ervan uitgaande dat de scores afkomstig zijn uit een normaal verdeelde populatie).

De psychologische betekenis (naar welke eigenschap verwijst de test) en de **validiteit** vormen de belangrijkste kenmerken van de test. Van een goede test zijn de validiteit en psychologische betekenis bekend. In hoofdstuk 7 worden de verschillende vormen van validiteit behandeld.

De **betrouwbaarheid** van een test gaat over de mate waarin de resultaten van een test bij een herhaalde meting gelijk zijn aan de resultaten van de eerste meting. Het is vereist, omdat het niet uit zou mogen maken wannéér iemand getest wordt. Het bepalen van de betrouwbaarheid bij psychologische tests brengt problemen met zich mee. Het is namelijk vaak lastig om twee onafhankelijke metingen te verkrijgen. In hoofdstuk 5 wordt de betrouwbaarheid verder behandeld.

Objectiviteit is de onafhankelijkheid van storende invloeden vanuit de beoordelaar. Dit is belangrijk voor een goede interpretatie van de testcores. Als er sprake van objectiviteit is, maakt het niet uit wie de beoordelaar van de test is. Dit heet het **intersubjectiviteitsprincipe**.

Daarnaast zorgt objectiviteit voor reproduceerbaarheid van de test en de evaluatieprocedure. De mate van overeenstemming tussen beoordelaars geeft een indicatie voor de mate van objectiviteit. Dit wordt de **interbeoordelaarsbetrouwbaarheid** genoemd. Deze interbeoordelaarsbetrouwbaarheid van twee beoordelaars kan worden uitgedrukt in een statistische maat, afhankelijk van de soorten testgegevens. Bij rangordes is de interbeoordelaarsbetrouwbaarheid een rangcorrelatie, uitgedrukt in bijvoorbeeld Kendalls tau of Spearman's Rho. Er bestaan ook methoden waarmee de overeenstemming tussen meerdere beoordelaars kan worden bepaald.

Wanneer de beoordelaars geen rangorde bepalen, maar mensen of objecten in categorieën indelen op de nominale schaal, dan wordt vaak Cohens kappa gebruikt om de interbeoordelaarsbetrouwbaarheid te berekenen.

$$\text{Kappa} = \frac{P_0 - P_t}{1 - P_t}$$

In het geval van twee beoordelaars, die beide mensen/objecten in categorieën indelen, wordt de (onderdelen van) kappa als volgt berekend.

De kans op overeenstemming (P_0) wordt berekend door het aantal identiek gemaakte beoordelingen te delen door het totaal aantal gemaakte beoordelingen. Stel dat beoordelaars A en B het 25 keer van de totale 40 beoordelingen met elkaar eens zijn, dan is de P_0 0,625.

Dan wordt de toevalskans P_t berekend. De grootte van deze kans geeft aan hoe vaak beoordelaars het met elkaar eens zouden zijn wanneer hun beoordelingen slechts afhankelijk zijn van een persoonlijke frequentieverdeling. Deze frequentieverdeling geeft aan hoe de verhoudingen tussen de ingedeelde categorieën zijn en verschilt per beoordelaar. Door ervan uit te gaan dat beoordelaars altijd volgens dezelfde verhoudingen indelen, kan uitgerekend worden hoe groot de overeenstemming tussen beoordelaars is als ze verder geen andere informatie meenemen in hun beoordelingen.

Dit wordt statistische onafhankelijkheid genoemd, omdat het alleen afhankelijk is van de frequentieverdelingen (f). De frequentieverdeling per beoordelaar kan worden vastgesteld door te tellen hoeveel mensen/objecten die beoordelaar in de verschillende categorieën heeft ingedeeld. Stel $f(A)$: 12, 16, 12; en $f(B)$: 12, 12, 16 bij een totaal aantal van 40 beoordelingen. Het aantal beoordelingen waarover de beoordelaars A en B het onafhankelijk eens zijn, is $(12 \times 12)/40 = 3.6$; $(16 \times 12)/40 = 4.8$; $(12 \times 16)/40 = 4.8$. De toevalskans is dan: $(3.6 + 4.8 + 4.8) / 40 = 0.33$.

De mate van overeenstemming tussen A en B is dan het verschil tussen P_0 en P_t . In bovenstaand voorbeeld is dat $0.625 - 0.33 = 0.295$. Dit kan nog worden afgezet tegen het maximaal mogelijke verschil tussen de beoordelaars. Dit is $1 - P_t$ en in bovenstaand voorbeeld dus $1 - 0.33 = 0.67$. Hiermee kan Cohens kappa berekend worden, zie eerdere formule.

Een negatieve waarde van kappa betekent dat de waargenomen overeenstemming nog slechter is dan verwacht zou kunnen worden volgens statistische onafhankelijkheid of toeval. De maximumwaarde van kappa is 1.

Cohens kappa kan worden gezien als een correctie voor toevalsovereenstemming. Wanneer twee beoordelaars onafhankelijk van elkaar personen of objecten indelen, zullen zij het over een deel van de gevallen al puur op basis van toeval met elkaar eens zijn. De waargenomen overeenstemming (P_0) geeft hierdoor vaak een overschatting van de werkelijke overeenstemming. De formule berekent welk deel van de mogelijke overeenstemming bovenop het toeval daadwerkelijk is behaald. De teller ($P_0 - P_t$) geeft de werkelijke overeenstemming boven de toevalskans weer, terwijl de noemer ($1 - P_t$) de maximaal haalbare overeenstemming boven toeval representeert. Een kappa van 0 betekent dat de overeenstemming niet groter is dan op grond van toeval verwacht mag worden, terwijl een kappa van 1 duidt op een perfecte overeenstemming tussen de beoordelaars.

Objectiviteit is een begrip dat zo veel mogelijk wordt nagestreefd bij testen. Of en in hoeverre het bereikt kan worden, hangt onder andere af van de soort test. Testscores die ook worden bepaald door de interpretatie en evaluatie (bijvoorbeeld een essay) van de proefleider zijn minder objectief dan testscores waarbij interpretatie van de proefleider geen rol speelt (bijvoorbeeld meerkeuzevragen). Indien men de interpretatie wel wil laten meewegen, is een exact voorgeschreven verwerkingsprocedure vereist.

Meting van eigenschappen aan de hand van test

Metten en testen worden vaak door elkaar heen gebruikt. In theorie is een test een middel om iets te meten. Deze paragraaf kijkt naar hoe een test verschillende psychologische eigenschappen kan meten.

Meetniveaus

Volgens Stevens (1951) is er sprake van meten op het moment dat getallen op een systematische wijze aan objecten worden toegekend. Dit leidt tot vier meetniveaus:

1. Nominale schaal;
2. Ordinale schaal;
3. Intervalschaal;
4. Verhoudingsschaal.