Maastricht University

Leerstoel in de Psychodiagnostiek

Citation for published version (APA):

Hurks, P. P. M. (2022). *Leerstoel in de Psychodiagnostiek: Is meten wel zeker(der) weten in de psychologie?* Maastricht University. https://doi.org/10.26481/spe.20220909ph

Document status and date:
Published: 09/09/2022

DOI:
10.26481/spe.20220909ph

Document Version:
Publisher's PDF, also known as Version of record

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy
If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Download date: 27 Apr. 2024

Maastricht University

Inaugurele rede

Prof. dr. Petra P.M. Hurks

Faculty of Psychology and Neuroscience

Leerstoel in de Psychodiagnostiek: Is meten wel echt (beter) weten?

Leerstoel in de Psychodiagnostiek:
Is meten wel echt (beter) weten?

Prof. dr. Petra P.M. Hurks

1.
Geachte rector en leden van de cortège, beste familie, vrienden en collega's,

Wat is het fijn dat u vandaag hier bent. Het verheugt mij een lezing te mogen geven over een heel boeiend vak, te weten de psychodiagnostiek. Het is fantastisch dat de Universiteit Maastricht het grote belang van dit vak voor de psychologie wil uitdragen, met een leerstoel.

Ik hoop u in de komende 45 minuten uit te leggen waarom het vak zo boeiend en belangrijk is. En hoe ik graag wil bijdragen aan dit vak.

Ik kijk hierbij naar de psychodiagnostiek vanuit de psychologie. Echter, psychodiagnostiek is even belangrijk binnen bijvoorbeeld de psychiatrie en de orthopedagogiek.

2.
Op één gebouw van de Universiteit Maastricht pronkte afgelopen voorjaar deze poster. Dit zijn de 17 doelen, die de Verenigde Naties, voor 2030 wil realiseren.

Eén doel is het bevorderen van een goede mentale gezondheid, wereldwijd. Ook Nederland kan op dit gebied nog winst boeken. Dat blijkt onder andere uit cijfers van het CBS. Zij vroeg Nederlanders om enkele vragen te beantwoorden over hun mentale gezondheid. Zo vroeg zij: voelde u zich de laatste 4 weken erg zenuwachtig, neerslachtig, of juist gelukkig? Op basis hiervan schatte het CBS dat medio 2021 circa 15% van de ondervraagden 'mentaal ongezond' was. 15%.

Verder meldde het CBS dat psychische klachten vaak de reden zijn om een arbeidsongeschiktheidsuitkering toe te kennen. In 2020 was dit bijv. bij ca. 4 op de 10 toekenningen het geval. Dat zijn 346 duizend Nederlanders.

Goede psychodiagnostiek is mijns inziens belangrijk als je mentale gezondheid wilt bevorderen. Maar wat is psychodiagnostiek?

3.
Ik ben lid van de landelijke commissie Basisaantekening Psychodiagnostiek. Dit is een commissie van het Nederlands Instituut van Psychologen, het NIP. Deze BAPD-commissie beoogt de kwaliteit van de psychodiagnostiek in Nederland te bevorderen.

Zij definieert psychodiagnostiek als “een onderzoeksproces dat is gericht op het doen van uitspraken over emotie, cognitie en gedrag van personen. Naar aanleiding van dit proces kan een beslissing genomen worden, een behandeling uitgevoerd en/of een advies gegeven”.

Deze definitie kent twee relevante elementen. Zo stelt de definitie dat psychodiagnostiek helpt om een uitspraak over een persoon te doen. Hiermee probeert de psycholoog een antwoord te formuleren op een hulpvraag van een persoon of diens omgeving. Zoals heeft deze persoon een leesstoornis? Of welk onderwijssysteem past het beste bij deze jongere?

Verder stelt de definitie dat psychodiagnostiek een onderzoeksproces is. Hierbij is het doel dat middels onderzoek, op een systematische manier, betrouwbare en valide informatie verkregen wordt over een persoon.

4.
Het doel van de psychodiagnostiek is dus om gefundeerd een uitspraak te doen over personen. Als deze uitspraak fout of minder goed is, kan dit serieuze gevolgen hebben voor iemand’s mentale gezondheid. Bedenk bijvoorbeeld eens wat de consequenties zijn als een jongere niet het juiste schooladvies krijgt en daardoor op een verkeerd niveau onderwijs volgt.

Daarom is psychodiagnostiek zo’n belangrijk vak. We doen al veel goed in Nederland op het gebied van de psychodiagnostiek. In de opleidingen psychologie wordt dit vak onderwezen, wetenschappers verrichten er onderzoek naar, beroepsverenigingen spenderen er tijd aan, etc. maar is dit voldoende?

5.
Er zijn aanwijzingen dat er nog wat ruimte is voor verbetering. Ik bespreek hier een aantal voorbeelden vanuit de geestelijke gezondheidszorg, GGZ. Maar er zijn zeker ook voorbeelden te geven vanuit andere disciplines in de psychologie.

In 2021 stelde bijv. GGZ nieuws dat mensen met vooral complexe psychiatrische problemen soms jaren rondlopen met een verkeerde diagnose of een gemiste hoofddiagnose.

Ook de Nederlandse Vereniging voor Gezondheidszorgpsychologie, de NVGzP, stelt in 2021 dat er nog verbeteringen mogelijk zijn in de kwaliteit van de psychodiagnostiek in bijv. de GGZ.

6.
De redenen hiervan zijn volgens hen divers. Zo is er helaas een gebrek aan geld en tijd binnen de GGZ en zijn er lange wachtlijsten. Om deze wachtlijsten in te korten, vreest de NVGzP dat zorgverzekeraars en managers psychologen gaan aanmoedigen om snel te starten met behandelen en daarmee minder tijd aan de diagnostiek te besteden. Het grote gevaar hierbij is dat als de conclusie van de diagnostiek fout is, iemand een behandeling krijgt die niet of minder passend is en de kosten én de wachttijden potentieel alleen maar oplopen.

Andere redenen die de NVGzP noemt hebben betrekking op de kennis en vaardigheden van psychologen. Zo stelt zij dat in een deel van de instellingen vooral Master Psychologen de diagnostiek verrichten. Hoewel zij een goede opleiding in psychodiagnostiek hebben genoten, zijn specialisten, zoals Gezondheidszorg psychologen, hier nog beter voor opgeleid. Laatstgenoemden hebben meer diagnostiek onderwijs genoten en langer gewerkt onder supervisie van een ervaren psycholoog.

De NVGzP benadrukt daarbij hoe belangrijk het is dat psychologen na hun opleiding zich een leven lang blijven bijscholen op het gebied van de psychodiagnostiek.

Tenslotte pleit de NVGzP voor meer structurele intercollegiale toetsing en intervisie, als het gaat om de psychodiagnostiek. Het doel hiervan is om van elkaar te leren en de diagnostiek te verbeteren. Ik kom straks nog terug op deze aanbevelingen.

7.
De afgelopen jaren is er ook wetenschappelijk onderzoek verricht naar hoe goed professionals diagnoses stellen. Ik bespreek hier, als voorbeeld van dit soort onderzoek, twee onderzoeken gericht op de diagnostiek van ADHD.

Hier ziet u de DSM-criteria voor ADHD. Deze criteria zijn gebaseerd op een consensus, of overeenstemming, tussen experts en wetenschappelijk onderzoek. Er is van alles te zeggen over de waarde van de DSM, maar voor het verhaal laat ik dat hier achterwege.

De DSM stelt dat ADHD gekenmerkt wordt door aandachtsproblemen, hyperactiviteit en/of impulsiviteit. Daarbij geeft de DSM 9 symptomen van een aandachtsprobleem. Een voorbeeld is: Raakt het kind afgeleid door externe en onbelangrijke informatie? Een kind moet minimaal 6 van de 9 symptomen laten zien, om te spreken van een aandachtsprobleem. Zo is er ook een lijst van symptomen van hyperactiviteit en impulsiviteit en een bijbehorend afkappunt.

Om van ADHD te spreken, moeten deze symptomen voor het 12e levensjaar en in min. 2 settings, zoals op school en thuis, zichtbaar zijn. En dienen de symptomen te leiden tot problemen in het dagelijks leven.

Het oogt eenvoudig om te checken of iemand ADHD heeft met deze DSM. Maar is dat echt zo?

8.
Hier wordt o.a. onderzoek naar gedaan met zogenoemde vignetten. Vignetten zijn casusbeschrijvingen.

Een voorbeeld van een vignet is: Molly is 7 jaar oud. Ouders geven aan dat Molly een slechte slaper is. Ze is overdag druk en vaak onrustig. Op school heeft ze moeite zich te concentreren en wordt ze snel boos als ze haar zin niet krijgt.

Ik bespreek hier de resultaten van twee vignetten-onderzoeken. Bij beide studies kregen proefpersonen 4 vignetten. Slechts één vignet voldeed aan alle ADHD-criteria. De andere vignetten voldeden niet aan alle ADHD-criteria, bijvoorbeeld doordat 2 of 3 DSM-criteria ontbraken. De proefpersoon moest per vignet aangeven of de cliënt ADHD had. Het goede antwoord is dus dat slechts 1 vignet voldoet aan alle ADHD-criteria.

9.
Dit is het resultaat van de 2 studies. Wat bleek: beide studies vinden dat vaak, in gemiddeld 75-82% van de gevallen, de proefpersoon het bij het rechte eind heeft. Dus, de proefpersoon oordeelde vaak correct of iemand wel of geen ADHD had. Dat dit percentage hoog is, viel ook te verwachten, als je bedenkt dat de proefpersonen in hun opleiding les hebben gehad over psychodiagnostiek.

Echter, de studies laten ook zien dat in 18-25% van de gevallen, de proefpersonen onterecht een diagnose ADHD gaven, of juist onterecht zeiden dat er geen sprake was van ADHD. Stelt u zich voor dat hierdoor deze mensen ten onrechte bijv. ADHD-medicatie of juist geen behandeling krijgen.

Ca. 1/5e van de beoordelingen was dus fout in dit voorbeeld.

10.
De auteurs noemen als een mogelijke verklaring hiervoor de representativiteitsheuristiek theorie. Deze theorie veronderstelt dat diagnostici bij een cliënt niet steeds systematisch alle DSM-criteria voor bijv. ADHD nalopen, maar dat zij bij zo’n ADHD-casus vooral checken of de symptomen die zij zelf het meest belangrijk vinden aanwezig zijn.

De data met de vignettenstudies dient vooral gezien te worden als illustratie van het type onderzoek dat gedaan wordt op dit gebied. We moeten voorzichtig zijn om op basis hiervan harde conclusies te trekken over de kwaliteit van de psychodiagnostiek in Nederland.

In de twee vignettenstudies die ik toonde werden daarnaast geen tests afgenomen. Een vraag is of de afname van tests kan helpen om de diagnostiek te verbeteren. Helpt bijvoorbeeld een vragenlijst die systematisch alle ADHD-criteria naloopt, de diagnosticus om beter te bepalen of iemand ADHD heeft?

De vraag is dan ook of de inzet van een test de psycholoog helpt om de kans op fouten te minimaliseren.

11.
De term ‘psychologische test’ verwijst naar “alle gestandaardiseerde instrumenten gebruikt om gedrag of eigenschappen van een persoon te meten” (definitie American Psychological Association). Voorbeelden van deze eigenschappen zijn: emotioneel functioneren, cognitieve functies, als aandacht en redeneren, en persoonlijkheidskenmerken.

Er zijn diverse soorten psychologische tests, waaronder vragenlijsten, praktijkopdrachten en prestatietests.

12.
Hier vind u twee voorbeelden van tests, die beide claimen aandacht te meten. De ene is een vragenlijst. Deze lijst is gebaseerd op de DSM. Een voorbeeld van een vraag is: Hoe vaak wordt u afgeleid door activiteiten of lawaai in uw omgeving? De antwoordopties zijn nooit tot zeer vaak. De som van de vinkjes in de donker gekleurde kolommen geeft een indicatie of iemand symptomen van een aandachtsprobleem vertoont. Overigens vraagt deze lijst ook naar de symptomen van hyperactiviteit en impulsiviteit.

De andere maat is een cognitieve test. De opdracht is om aan te geven hoe vaak de twee bovenste figuren hieronder te zien zijn. Het goede antwoord is 2. Een uitkomstmaat kan zijn hoe snel iemand het goede antwoord vindt.

13.
Toenemend wordt bestudeerd of hersenimaging technieken, als test, bruikbaar zijn binnen de psychodiagnostiek.

Zo wordt gepubliceerd over de meerwaarde van de Elektro-encefalografie (EEG) voor de diagnostiek van ADHD. Op een niet-invasieve wijze, worden via elektrodes in een badmuts hersengolven gemeten. De analyse van deze hersengolven biedt in potentie meer inzicht in bijvoorbeeld de aandachtsregulatie van kinderen.

Ondanks dat de afgelopen jaren op het gebied van hersenimaging belangrijke ontwikkelingen hebben plaatsgevonden, worden dit soort metingen in de meeste instellingen nog niet vaak ingezet in de psychodiagnostiek. Dit gaat mogelijk veranderen in de nabije toekomst, nu apparaten nodig voor hersenregistraties vaak mobieler, goedkoper en gebruikersvriendelijker worden en er meer kennis komt over de betekenis en waarde van dergelijke registraties.

Dit zijn slechts drie voorbeelden van tests. Er zijn wereldwijd veel tests. De Buro’s, een onafhankelijk orgaan dat de kwaliteit van tests in Amerika beoordeelt, stelt bijv. op haar website dat er in Amerika meer dan 3500 psychologische tests zijn. In Nederland ligt dit aantal veel lager, vrees ik.

14.
Ook in Nederland gebruiken psychologen tests vaak. In 2000, 2009 en 2020, hebben we een vragenlijst uitgezet onder Nederlandse psychologen. Hieruit blijkt dat psychologen de afgelopen decennia meer tests zijn gaan gebruiken. Met name computer tests worden tegenwoordig vaker gebruikt. Daarnaast zien psychologen de meerwaarde van tests, met name als de testresultaten gecombineerd worden met andere gegevens.

Maar is deze hoge waardering voor tests terecht?

15.
De titel van mijn lezing vandaag luidde daarom “Is meten wel echt (beter) weten?”. Deze titel is gebaseerd op het gezegde “Meten is weten”. Bij de voorbereiding op vandaag ben ik eerst op zoek gegaan naar waar dit gezegde vandaan komt.

16.
Het internet linkt vaak twee mensen aan dit gezegde, te weten Nobelprijswinnaar Professor Kamerlingh Onnes en Lord Thomson Kelvin. Beiden deden in de 19e eeuw onderzoek naar natuurverschijnselen en stelden dat nauwkeurige metingen kunnen helpen om kennis te vergaren, dus om meer te weten. Kamerlingh Onnes benadrukte de meerwaarde van deze metingen door te stellen: “meten tot weten zou ik gaarne als zinspreuk boven elk fysisch laboratorium willen schrijven”.

Beide ontkennen echter dat meten zeker leidt tot weten, maar dat het eerder zo is dat meten kan bijdragen tot weten. De vraag is of meten ook kan leiden tot minder weten of dat we onterecht denken meer te weten.

Vandaag wil ik met u kijken hoe dit zit in de psychologie en/of er nog iets moet gebeuren qua onderwijs of onderzoek om met nauwkeurige metingen meer te weten.
Echter, voordat ik deze vragen ga beantwoorden, acht ik het nodig om eerst uit te leggen wat het onderzoeksproces van de psychodiagnostiek inhoudt en welke plaats de test daarin inneemt.

17.
Het begint met een hulpvraag. De eerdere genoemde landelijke commissie BAPD maakt onderscheid tussen groeibevorderende hulpvragen, zoals vragen met betrekking tot iemands loopbaan of opleiding, en hulpvragen gericht op een verstoorde of bedreigde ontwikkeling. Hier ziet u een aantal voorbeelden van hulpvragen.

18.
Stel iemand komt met zo’n hulpvraag bij de psycholoog: wat is dan de volgende stap? Hoewel er verschillende uitwerkingen van het diagnostisch onderzoeksproces bestaan, bevatten de meeste uitwerkingen een aantal dezelfde elementen.

19.
Eerder noemde ik Molly, een meisje waarvan de ouders zich afvragen of zij ADHD heeft. Ik zal het diagnostisch onderzoeksproces kort toelichten aan de hand van deze casus.

Het diagnostisch onderzoeksproces start bij Molly met een gesprek met haar ouders.

20.
Na dit gesprek én met kennis vanuit de wetenschap, destilleert de psycholoog een of meerdere aannames, ook wel hypotheses genoemd. Een hypothese kan zijn dat Molly ADHD heeft, omdat bijv. de DSM stelt dat druk gedrag mogelijk kan wijzen op ADHD.

Op dat moment schat de psycholoog in hoe zeker hij is van deze hypothese. Doorgaans is de psycholoog in deze fase nog te onzeker over zijn hypothese om een advies te formuleren, zoals bijv. het wel of niet starten van een ADHD-behandeling.

In een poging zekerder te zijn van zijn hypothese, start de psycholoog dan de volgende fase van de diagnostiek.

21.
Dit doet hij door voor iedere hypothese een toetsingsplan met bijbehorende toetsingscriteria te formuleren. Ook hierbij wordt een link gelegd met de wetenschappelijke literatuur. Zijn hypothese is dus Molly heeft ADHD. Het toetsingscriterium kan zijn dat Molly minimaal 6 van de 9 symptomen van een aandachtsprobleem laat zien, zoals beschreven in de DSM. In deze fase kan een psycholoog besluiten om een test in te zetten om dit criterium te toetsen.

22.
Vervolgens gaat de psycholoog data verzamelen.

Aan de hand van de verkregen data, schat de psycholoog opnieuw in hoe zeker hij is van zijn oordeel. De zekerheid kan zijn toegenomen of juist afgenomen.

De resultaten kunnen leiden tot nog altijd te veel onzekerheid of tot nieuwe vragen, waardoor de cyclus zich herhaalt. Dit proces blijft doorgaan totdat de psycholoog voldoende zekerheid heeft om een uitspraak te doen en/of een advies te geven. Het is moeilijk te voorspellen wanneer een psycholoog zegt dat hij voldoende zekerheid heeft. Dit is onder andere afhankelijk van de aard en complexiteit van de hulpvraag.

23.
Diverse studies hebben gekeken naar de meerwaarde van tests binnen deze cyclus. Deze studies geven vaak direct dan wel indirect bewijs voor deze meerwaarde.

Op basis van een literatuur onderzoek stellen Meyer en collega's bijvoorbeeld dat psychologen die uitsluitend vertrouwen op interviews vatbaar zijn voor onvolledige en foutieve inzichten. Zij pleiten daarom voor de inzet van tests in de psychodiagnostiek. Als argument noemen ze onder meer dat de validiteit van de psychologische test sterk en overtuigend is. Het begrip validiteit verwijst naar of de test meet wat hij beoogt te meten. Dus als een test beweert aandacht te meten of deze test dat ook daadwerkelijk meet. Door deze goede validiteit, helpt de inzet van tests volgens Meyer en collega's de psycholoog bij de diagnostiek.

24.
De meerwaarde van instrumenten blijkt ook uit vignettenstudies. Zo gaven bijvoorbeeld Cwik en collega's (2016) proefpersonen drie casussen, of vignetten. Per casus, moesten de proefpersonen aangeven of de casus voldeed aan de criteria van een psychiatrische stoornis, te weten een ernstige depressie, een gegeneraliseerde angststoornis, en/of een Borderline persoonlijkheidsstoornis. 50% van de proefpersonen kreeg alleen de casusbeschrijvingen. De andere helft kreeg aanvullend een checklist met daarop per stoornis een overzicht van de DSM-symptomen van de stoornis.

Wat bleek: de proefpersonen die de checklist kregen, voelden zich zekerder bij het stellen van een diagnose. Daarnaast maakten ze doorgaans minder fouten- hoewel zij nog altijd fouten maakten. Wel stelden de auteurs dat het succes van zo'n checklist afhangt van of diagnostici voldoende training gehad hebben en hoe complex de casus is.

25.
Psychologische tests kunnen dus potentieel de diagnostiek verbeteren. De mate van verbetering hangt daarbij af van diverse factoren. Twee factoren zijn: de kwaliteit van de tests en hoe deze tests meegewogen worden door de psycholoog. Beide punten zal ik hier bespreken.

26.
Ten eerste dient de kwaliteit van de test voldoende te zijn. De Commissie Testaangelegenheden Nederland (de COTAN) is een centrale commissie van het Nederlands instituut van Psychologen, het NIP. Hier ziet u een foto van de COTAN leden. Deze commissie bestaat al meer dan 60 jaar en heeft als doel de kwaliteit van tests en testgebruik in Nederland te bevorderen.

Een manier waarop zij dat doet, is door het opstellen van beoordelingscriteria waarmee de psycholoog de kwaliteit van tests kan beoordelen. Daarnaast beoordeelt de COTAN zelf ook tests, om zo psychologen nog beter te helpen. Tenslotte geeft zij richtlijnen over testgebruik. Ik vind het een eer om voorzitter te zijn van deze commissie met haar gedreven en deskundige leden.

27.
De COTAN beoordeelt dus onder andere de kwaliteit van tests. Zo kijkt ze naar wat de test beoogd te meten, hoe betrouwbaar, of hoe nauwkeurig, de test is en of de test meet wat de test wil meten.

De COTAN geeft geen totaal oordeel. Het is aan de psycholoog om te bepalen of de test inzetbaar is en onder welke voorwaarden. De psycholoog moet daarbij zelf oordelen of een test die als volgt beoordeeld wordt, bruikbaar is bij de vorming en toetsing van hypotheses.

In dit voorbeeld is de validiteit onvoldoende. Het begrip validiteit verwijst dus naar meet de test wat de test beoogt te meten. Stel de testauteur beweert dat de test goed voorspelt of het VWO passend is bij een kind. Met validiteitsonderzoek wordt dan bekeken of de test dit inderdaad goed voorspelt. Als de validiteit laag is, dan voorspelt deze test mogelijk onvoldoende hoe iemand het gaat doen op het VWO.

De Beroepsethische zaken commissie van het NIP stelt dat tests waarvan niet bekend is of de kwaliteit voldoende is, **"**niet gebruikt mogen worden om beslissingen op te baseren. Zij mogen hooguit voor hypothesevorming worden gebruikt".

28.
De waarde van de COTAN maar vooral ook het belang van kwalitatief goede tests wordt mijns inziens door diverse instanties onderschreven. Zo dragen alle universiteiten met een opleiding psychologie in Nederland bij aan het werk van de COTAN. Ook komt de kwaliteit van tests terug in alle psychologie-opleidingen in Nederland. De COTAN beoordelingscriteria komen daarbij vaak aan bod, ook in Maastricht

Testkwaliteit en de COTAN worden daarnaast genoemd in diverse stukken geschreven door ministeries. Ik geef hier een paar voorbeelden. Zo noemt de Arbeidsinspectie het belang van beide in de toelichting op de Wet toezicht gelijke kansen bij werving en selectie en in de checklist arbeidsdiscriminatie.

Ook beoordeelt de COTAN bij wet en in opdracht van het ministerie van Onderwijs, Cultuur en Wetenschap momenteel nog de kwaliteit van de eindtoetsen ingezet in groep 8 van het basisonderwijs.

Verder wordt in diverse documenten op het internet de COTAN genoemd als standaard om de kwaliteit van tests te beoordelen. Dit is bijvoorbeeld het geval in deze landelijke richtlijn ADHD of in diverse kwaliteitsstandaarden opgesteld voor de Akwa GGZ.

29.
Ondanks het belang van kwalitatief goede tests, geven psychologen op een recente enquête slechts een matig cijfer op de vraag of zij voldoende informatie hebben over de kwaliteit van tests. Ik heb vaker gehoord dat ze het daarnaast vaak lastig vinden om deze kwaliteit zelf te beoordelen ahv. Een handleiding. Een commissie als de COTAN kan de psycholoog helpen, door voorlichtingen te geven over deze kwaliteit van tests. Ook beoordeelt, zoals gezegd, de COTAN hierom zelf de tests.

Echter, op dit moment beoordeelt de COTAN alleen tests die testauteurs zelf aanbieden. Dat betekent dat niet alle tests beoordeeld worden.

Daarnaast zijn er aardig wat beoordeelde tests die min. 1 onvoldoende krijgen van de COTAN.

30.
Zo heb ik gekeken naar een willekeurig elftal cognitieve tests en hoe de COTAN deze tests heeft beoordeeld. Cognitieve tests meten bijvoorbeeld aandacht of geheugen. Ik koos voor cognitieve tests, omdat ik zelf door mijn achtergrond als neuropsycholoog vaker cognitieve tests ingezet heb in het onderwijs en onderzoek. Het valt op dat deze willekeurig gekozen tests vaker onvoldoendes krijgen op een of meerdere beoordelingscriteria, zoals u hier ziet.

In de toelichting op deze testbeoordelingen stelt de COTAN vaker dat zij deze onvoldoende geeft omdat de auteurs niet voldoende onderzoek hebben verricht. Het kan dus best zijn dat deze tests betrouwbaar of valide zijn, maar de COTAN heeft onvoldoende gegevens om er zeker van te zijn dat dit het geval is. Daarom geeft zij een onvoldoende.

Soms vragen mensen zich daarnaast af of de COTAN beoordelingscriteria wel passend zijn bij dit type tests.

31.
Dat cognitieve tests wel kunnen voldoen aan de COTAN eisen blijkt als ik kijk naar de beoordeling van intelligentietests. Daar zien we meer tests die minimaal een voldoende krijgen op alle criteria.

Echter, ik denk wel dat de COTAN veel vraagt van testauteurs. De kwaliteit van tests moet immers goed zijn als u daarmee een uitspraak over een persoon wilt doen.

Het onderzoek nodig om deze kwaliteit te onderzoeken kost veel geld en tijd en er zijn helaas maar beperkt subsidiemogelijkheden voor dit soort werk.

32.
Daarom wil ik hier een lans breken richting overheid, zorgverzekeraars en andere bedrijven om meer gelden vrij te maken voor testontwikkeling en onderzoek naar de kwaliteit van tests in Nederland. Om zo de psycholoog en andere diagnostici te helpen.

Als een goede testkwaliteit zo belangrijk is dat dit in diverse overheidsdocumenten en landelijke richtlijnen wordt genoemd, is het ook noodzaak om middelen vrij te geven om dit werk te kunnen doen. Daarin kan mijns inziens het voorbeeld gevolgd worden van Stichting ABBAS fonds.

Deze stichting wil onderzoek naar en ontwikkeling van psychologische tests vooruithelpen. De stichting stelt daartoe onder andere jaarlijks een subsidie van 30.000 euro beschikbaar. Ieder jaar worden veel hele goede voorstellen ingediend, maar helaas kan daarvan slechts 1 of maximaal 2 voorstellen gesubsidieerd worden, door het budget. Doordat ik in de wetenschappelijke commissie zit van deze stichting ben ik mogelijk bevooroordeeld. Maar mijns inziens helpen deze subsidies de testontwikkeling en het onderzoek naar test kwaliteit. Echter, meer geld is goed bruikbaar opdat we de psychodiagnosticus nog beter zijn werk kunnen laten doen.

33.
De komende jaren wil ik niet alleen via de COTAN en Stichting ABBAS Fonds een bijdrage leveren aan de ontwikkeling van tests. Zo stel ik mijzelf ten doel:
Bestaande tests uit het buitenland te vertalen en/of te bewerken,
Bestaande tests te verbeteren door bijvoorbeeld naar alternatieve scoringsmethodieken te kijken,
En nieuwe tests te ontwikkelen.
Dat doe ik natuurlijk niet alleen.

34.
Hier geef ik u een aantal voorbeelden van tests waar we aan werken. Bij de eerste voorbeelden zetten we technologie in bij de ontwikkeling van tests.

Technologie kan bijdragen aan de diagnostiek in bijv. de zorg, arbeids- en organisatiepsychologie, en het onderwijs. Het belang van technologie wordt bijvoorbeeld genoemd in de recent uitgekomen kennisagenda voor het onderwijs in de toekomst. Zij stelt onder andere dat technologie de leerkracht kan helpen om de ontwikkeling van leerlingen te monitoren. Hiervoor moeten mijns inziens nog wel meer instrumenten ontwikkeld worden. Samen met leerkrachten zijn we dergelijke instrumenten aan het ontwikkelen, opdat we meer informatie over kinderen kunnen verzamelen. Ik geef daar dadelijk ook een voorbeeld van.

35.
Mijn eerste voorbeeld is gebaseerd op de doolhoftest. Dit is een bekende cognitieve test. U kent allemaal het principe van een doolhof - vind zo snel mogelijk uw weg van a naar b. Deze test wordt vaak op papier afgenomen met als uitkomstmaat hoe snel iemand van a naar b kan gaan.

Deze uitkomstmaat zegt iets over hoe iemand presteert op deze taak. Maar het geeft nog weinig inzicht in de redenen waarom iemand bijv. trager is. Misschien is iemand motorisch trager, maar het kan ook zijn dat iemand te vaak een doodlopende weg inloopt. Dit soort informatie is nodig om een genuanceerdere uitspraak te kunnen doen over de persoon.

We denken dat dit soort informatie digitaal gemakkelijker en nauwkeuriger te verzamelen is. Daarom hebben we met experts vanuit diverse disciplines gewerkt aan de intelligent maze.

36.
Zo ziet de intelligent maze eruit. De opdracht is dat het kind met een pen met magneet een balletje door het doolhof beweegt. Terwijl het kind speelt met het doolhof wordt automatisch informatie opgeslagen over hoe het balletje beweegt als een functie van tijd en plaats.

37.
De eerste data zijn inmiddels verzameld met deze test bij kleuters.

De analyses laten zien dat diverse maten met het instrument berekend kunnen worden, zoals Totale afstand afgelegd, hoe vaak en hoe lang was iemand inactief, gemiddelde snelheid van bewegen, hoe vaak slaat iemand een doodlopende weg in, etc.. Als je deze maten correleert met de originele doolhoftaak, dan blijkt slechts 1 van de 4 maten van onze test te correleren met de score op de originele test. Meer onderzoek met onder andere geavanceerde analysetechnieken is nodig om nog beter te weten wat de meerwaarde is van ons instrument.

38.
Voor het volgende voorbeeld hebben we eerder een innovatie prijs ontvangen. In dit project houden we ons bezig met h*oe robots* kunnen ondersteunen bij het observeren van samenwerkend leren bij kinderen.

In dit project werken we met de zogenoemde EDMO. Deze is goedkoop in de aanschaf en niet gemakkelijk kapot te krijgen.

De opdracht die de kinderen in tweetallen krijgen is om de EDMO zo snel mogelijk van a naar b te laten bewegen. Dit doen ze met 2 tablets, waarop ze diverse variabelen kunnen manipuleren, zoals omvang, positie en relatie. De kinderen wordt niet verteld wat deze variabelen betekenen. Dat moeten ze zelf uitvinden.

39.
Dit zijn voorbeelden van hoe mensen de EDMO laten bewegen.

Net als bij de intelligent maze gaan we bekijken welke data we kunnen verzamelen met de EDMO. Zo kunnen we bijv. analyseren hoe vaak mensen aan bepaalde knoppen zitten en in welke volgorde.

40.
Dit is een weergave van hoe twee kinderen met de EDMO speelden. Op de horizontale as wordt de tijd weergegeven dat de kinderen met de EDMO werkten. Op de verticale as staat welke knoppen er zijn. De streepjes geven weer welke knoppen wanneer zijn ingedrukt. De kinderen in dit groepje lijken alle knoppen te gebruiken. Ook lijkt er een volgorde te zijn in wanneer ze welke knop indrukken. We moeten nog verder uitzoeken wat dit betekent. Verder willen we de verbale uitingen van kinderen tijdens het werken met EDMO digitaal gaan analyseren, zoals hoeveel praten ze, op welke toon et cetera.

Aanvullend zijn we samen met leerkrachten een observatieschaal aan het ontwikkelen, die ingevuld wordt door leerkrachten terwijl de kinderen werken met de EDMO. Door het combineren van al deze data hopen we beter de vaardigheden van kinderen in kaart te brengen.

41.
Hier geef ik nog wat andere voorbeelden van ons werk.

Bij de projecten werken we vaak samen met experts met diverse achtergronden, zoals psychologie, technologie, zorg en onderwijs. Dat maakt het werk juist zo boeiend. Ook denk ik dat we hierdoor verder komen qua testontwikkeling.

Bij al deze tests spenderen we veel tijd aan het onderzoeken van de test kwaliteit. Zelfs als we tests bewerken vanuit een andere taal dient uitgebreid onderzocht te worden wat de kwaliteit van de tests is voor gebruik in Nederland. Zelfs als tests non-verbaal zijn.

42.
De relevantie hiervan blijkt uit o.a. onderzoek van Brouwers en collega’s (2009). die keken naar landenverschillen in testprestaties op de non-verbale Raven Progressive Matrices test. Bij deze test moeten proefpersonen aangeven welk plaatje op de lege plek hoort.

De onderzoekers namen dezelfde non-verbale test af in diverse landen en hier ziet u de grote verschillen in prestaties tussen landen. Zo behaalden de proefpersonen in Nederland ca. 15 punten minder op deze test dan degenen in Frankrijk. Dit is slechts een voorbeeld van de invloed van land en ondersteunt de aanname mijns inziens dat we ook de kwaliteit van een vertaalde versie van de test altijd moeten onderzoeken, qua normen en validiteit bijv.

43.
De komende jaren zullen we verder gaan met testontwikkeling. Daarbij verrichten we ook onderzoek naar factoren die (potentieel) de prestaties van een individu op deze tests beïnvloeden. Zo kijken we naar de invloed van strategie-instructie, itemvolgorde of beloningen op testprestaties.

44.
Deze tests kunnen helpen bij de psychodiagnostiek - als deze van voldoende kwaliteit zijn. Echter, zelfs de beste test is niet perfect.

Ter illustratie: hier ziet u plaatjes van een test die we eerder ontwikkeld hebben. De test meet onder andere cognitief functioneren. Zo moesten kinderen sterren vangen die onder een banaan uit komen.

We hebben de test afgenomen bij jonge kinderen met en zonder ADHD. We hebben bekeken of de test goed kon voorspellen wie in welke groep hoorde. We vonden dat de test 88% (35/40) van de kinderen met ADHD correct indeelde in de ADHD-groep, en 72% (41/57) van de gezonde kinderen in de controle groep. Dit is veel belovend. Echter, deze resultaten tonen ook dat de test 22% van de kinderen niet in de juiste groep indeelde. Hierdoor kunnen we door de test een verkeerd oordeel geven.

Een antwoord op een hulpvraag wordt daarom vrijwel nooit gebaseerd op 1 test. Moons en Van der Graaf (2000) stelden terecht: “Elke bevinding is slechts een onderdeel van het totaal aan informatie dat nodig is om de diagnose te stellen. Elke vorm van diagnostiek is in wisselende mate belastend voor de patiënt, tijdrovend en duur. Diagnostiek moet daarom zo efficiënt mogelijk worden uitgevoerd met inachtneming van acceptabele percentages foute beslissingen”.

45.
Veel, maar zeker niet alle tests waar ik me op richt zijn cognitieve tests. Modellen als de Research Domain Criteria (RDoC) beschouwen cognitie als een van de belangrijke domeinen die onderzocht moeten worden als je mentale gezondheid bestudeert.

RDoC stelt daarbij dat de biologie, zoals de hersenen en genen, en de omgeving, zoals culturele factoren en de lesmethoden, van invloed zijn op domeinen als cognitie. Zij moedigt onderzoekers aan om deze relaties te onderzoeken in de context van de mentale gezondheid.

Ook hieraan wil ik de komende jaren een bijdrage leveren middels wetenschappelijk onderzoek, ook omdat deze kennis in te zetten is bij de hypothesevorming en besluitvorming binnen de psychodiagnostiek.

Kortom: naast testontwikkeling verrichten we daarom onderzoek naar de relaties tussen biologische en omgevingsfactoren op de diverse domeinen genoemd door RDoC. Ik zal hiervan een paar voorbeelden geven.

46.
Zwangerschapsvergiftiging, of pre-eclampsie, komt voor bij ca. 5% van alle zwangerschappen wereldwijd (Abalos et al., 2013). Zwangerschapsvergiftiging lijkt samen te gaan met een verhoogd risico op allerlei problemen.

Zo laat onderzoek van Alers en collega's met onder andere dit plaatje zien dat mensen na een zwangerschapsvergiftiging een hogere kans hebben op cognitieve klachten dan mensen met een gewone zwangerschap. Het verschil neemt af naarmate de zwangerschap langer geleden is, echter, pas na 20 jaar lijkt dit verschil te verdwijnen.

De oorzaken van deze klachten zijn onbekend. Om de diagnostiek goed te kunnen doen, wil je graag deze oorzaken weten. Zijn deze klachten bijv. biologisch van aard? In Maastricht onderzoeken we dit vraagstuk. Zo zijn er hersenscans gemaakt van mensen met en zonder zwangerschapsvergiftiging.

47.
Zo vonden we op groepsniveau diverse verschillen in de hersenen en de functionele hersenorganisatie van mensen met en zonder een historie van zwangerschapsvergiftiging. Dat geeft een eerste aanwijzing dat er mogelijk biologische factoren een rol spelen bij de verklaring van de cognitieve klachten die mensen kunnen ervaren na zwangerschapsvergiftiging.

De komende jaren kunnen we hopelijk verder werken aan dit onderzoek. Hierbij werken we samen met experts vanuit diverse disciplines zoals de gynaecologie, neurowetenschappen en cardiologie.

48.
Hierbij geef ik u nog een aantal voorbeelden van projecten waar we aanwerken met experts uit diverse disciplines.

49.
Dus ik wil graag bijdragen aan testontwikkeling en aan wetenschappelijke kennis over mentale gezondheid en factoren die dit beïnvloeden.
Het derde doel is om bij te dragen aan hoe we effectief kennis over psychodiagnostiek overdragen aan bijvoorbeeld psychologen en psychologiestudenten. Dit licht ik toe middels een voorbeeld.

50.
Eerder stelde ik dat de kwaliteit van de diagnostiek afhangt van een aantal zaken: waaronder de kwaliteit van tests, maar ook hoe de testprestaties meegewogen worden door de psycholoog. Ik zal nu ingaan op het laatste.

51.
Zelfs met tests, is het voorspellen van menselijk gedrag moeilijk.
Stelt u zich voor dat 3 personen solliciteren op een baan als docent bij de Universiteit Maastricht. De commissie heeft besloten om deze instrumenten in te zetten bij de selectie. Zo vindt er een interview plaats, we nemen een capaciteitentest en een persoonlijkheidsvragenlijst af, en we doen een rollenspel. De selectiecommissie kiest kwalitatief goede instrumenten hiervoor.

Na afname heeft de selectiecommissie per kandidaat een heleboel data. De vraag is dan: hoe combineer je deze gegevens om te bepalen of een kandidaat geschikt is voor de baan als docent? En hoe kies je dan vervolgens één kandidaat die je de baan aanbiedt?

52.
Recent hebben we professionals, die vaker dit soort beslissingen moeten nemen, gevraagd hoe zij deze info zouden combineren. Verre weg de meesten gaven aan dat ze daarbij de holistische methode gebruiken. Bij deze methode combineert de professional alle verkregen informatie in zijn hoofd en komt hij zo tot een oordeel. Hierbij gebruikt hij zijn kennis, ervaring en intuïtie.

Echter, al in 1954, stelde Meehl dat deze aanpak nadelen heeft. Zo zijn mensen niet goed in het consistent wegen van verschillende gegevens in hun hoofd. In ons voorbeeld, kan dit betekenen dat de professional bij de ene sollicitant de intelligentietest het zwaarst laat wegen en bij de andere persoon de persoonlijkheidstest.

53.
Onder andere Meehl stelt als alternatief de mechanische aanpak voor. Hierbij stelt de sollicitatiecommissie bij de start van de selectieprocedure een algoritme of beslisregel op over hoe zij de informatie gaat combineren en met name gaat wegen.

54.
Dit is een voorbeeld van zo'n algoritme, met daarbij de gekozen weging per instrument. De hoogte van de weeggewichten is door de beoordelaars gezamenlijk bepaald bij de start van de selectieprocedure. In het voorbeeld weegt het rollenspel het zwaarste en het interview het lichtst. Hier ziet u de scores die ieder individu heeft gehaald op de verschillende onderdelen. In dit voorbeeld kan op ieder onderdeel max. 10 punten gehaald worden. Kandidaat 1 haalt een hoge score op het rollenspel, en kandidaat 3 een hoge scores op de capaciteitentest. Als we alle resultaten in de formule stoppen dan blijkt dat kandidaat 3 de hoogste totaal score behaalt en daarmee volgens het algoritme de meest geschikte kandidaat is voor de baan.

Hoe het algoritme eruit ziet, bepaalt de commissie zelf. De expertise van bijv. een psycholoog is ook bij deze methode essentieel, bijvoorbeeld bij de keuze van de instrumenten, het opstellen van het algoritme etc.

Al in 1954 stelde Meehl dat het gebruik van algoritmes vaak leidt tot betere en transparantere besluiten of voorspellingen dan de holistische aanpak. Later werd dit herhaaldelijk aangetoond met onderzoek, in diverse disciplines binnen de psychologie.

55.
Als deze methode dan beter werkt, en we weten dat al heel lang, waarom worden algoritmen dan nog relatief weinig gebruikt in de psychodiagnostiek? Dit hebben we in ons recente onderzoek gevraagd aan psychologen. Zij gaven een aantal redenen. Deze hebben vooral betrekking op misverstanden over algoritmen en behoefte van de psychologen aan meer praktische kennis hierbij. Ze noemden voorbeelden als hoe stel je de algoritmen op, wat doe je met tests die niet van voldoende psychometrische kwaliteit zijn, kun je ook bij complexe casuïstiek uit de voeten hiermee en hoe etc. Ook gaven ze aan dat er bij andere belanghebbenden, zoals opdrachtgevers, ook meer kennis nodig is hierover. Er zouden bij deze opdrachtgevers vooroordelen bestaan over algoritmen die de acceptatie hiervan in de weg staat.

In ons onderzoek noemden de proefpersonen zelf een aantal manieren om de acceptatie van algoritmen te bevorderen:

56.
Zoals geef trainingen of cursussen tijdens de bachelor, master opleidingen, maar ook daarna. Daarnaast noemden proefpersonen het belang van intervisie en samenwerken, waarbij psychologen, binnen één instelling of tussen instellingen moeten overleggen over hoe ze in een casus beslist hebben en hoe ze een algoritme hebben opgesteld.

Ook stelden de proefspersonen in dit onderzoek dat psychologen koppig zijn en dat zij denken dat ze beter kunnen voorspellen dan algoritmen. Hun werkwijze veranderen zal een uitdaging worden, volgens de proefpersonen.

57.
Deze adviezen sluiten aan bij de eerdergenoemde NVGZP die stelde dat psychodiagnostiek In Nederland belang heeft bij goede scholing, ook na de master of post-master opleidingen, en bij intercollegiale toetsing en intervisie.

Deze adviezen zijn van toepassing op alle disciplines in de psychologie waar beslissingen genomen worden.

In ons onderzoek meldde proefpersonen verder dat het kan helpen om landelijke richtlijnen op te stellen.

58.
Een voorbeeld van zo'n richtlijn is de checklist arbeidsdiscriminatie bedoeld voor werving- en selectiegesprekken. Deze is opgesteld door de arbeidsinspectie en stelt onder andere: maak gebruik van vooraf opgestelde selectiecriteria en richt het selectieproces zo in, dat je altijd kunt nagaan op basis waarvan beslissingen zijn genomen.

Algoritmen zijn hierbij van meerwaarde. Ik denk persoonlijk dat het daarbij nodig is om gebruikers van de checklist meer kennis te bieden over hoe zij bijv. de algoritmen kunnen opstellen en inzetten. Dit sluit aan bij de resultaten van ons eigen onderzoek, zojuist genoemd.

Een doel dat ik mijzelf dan ook stel is om te onderzoeken hoe we dit soort kennis over algoritmen, maar ook over tests, testgebruik en besluitvorming goed over kunnen dragen aan professionals, als psychologen en psychologiestudenten. Opdat we een kwaliteitsslag kunnen maken. Dat kan mijns inziens alleen maar samen, met professionals, opleidingen, beroepsverenigingen als het NIP en commissies als de COTAN en de BAPD-commissie. Hoewel psychologen eigenwijs zijn, heb ik er alle vertrouwen in dat we samen, met alle kennis aanwezig in Nederland, veel kunnen bereiken om de al goede psychodiagnostiek nog beter kunnen maken.

59.
Samenvattend, heb ik mijzelf onder andere de volgende doelen gesteld:
Tests ontwikkelen en de kwaliteit hiervan onderzoeken;
Kennis vergaren over factoren die testprestaties beïnvloeden;
Bijdragen aan onderzoek naar mentale gezondheid en factoren die dit beïnvloeden,
En onderzoeken hoe we effectief kennis over psychodiagnostiek kunnen overdragen.

60.
Tenslotte een woord van dank. 18 jaar geleden promoveerde ik in deze zaal. Collega's gaven mij toen een boek met hun adviezen voor mijn carriere. Mijn promotor adviseerde: "Eigenwijs wordt eigen wijze, dat zou goed zijn. Zit er wel in, komt goed". Dat ben ik nooit vergeten. Ik vrees alleen dat ik nog altijd wat eigenwijs ben, zoals klaarblijkelijk meer psychologen. Maar ik luister, denk ik, inmiddels beter, want wat een mooie verhalen en interessante en stimulerende opmerkingen heb ik in die jaren gehoord, in de samenwerking met zo veel mensen binnen en buiten de universiteit.

Het liefst bedank ik al deze mensen persoonlijk. Ik denk hierbij aan collega's en studenten op de universiteiten, en aan de benoemingsadviescommissie en de begeleidingscommissie opgericht in het kader van deze leerstoel. Ook denk ik zeker aan mensen buiten de universiteiten waarmee ik mag samenwerken, zoals leerkrachten, gedragsdeskundigen, proefpersonen. Deze samenwerkingen brengen onderzoek, onderwijs, en valorisatie zoveel verder. Het zorgt voor nieuwe ideeën en inzichten. Bedankt allemaal!

Ik wil de Universiteit Maastricht en in het bijzonder het college van bestuur, het college van decanen, de faculteit der psychologie en neurowetenschappen, bedanken voor het vertrouwen dat zij mij geven.

En natuurlijk zijn er dan mijn vrienden en familie. Bedankt dat jullie er altijd zijn, voor mij en Bart. Jullie zijn geweldig!

Tenslotte mijn partner, ouders, zus, en schoonouders.

61.
Hoe spijtig is het dat 3 van jullie vandaag niet aanwezig kunnen zijn. Weet dat ik jullie allemaal altijd in mijn hart meedraag. Ik houd van jullie.

Ik heb gesproken.