

Algemene Standaard Testgebruik NIP (AST-NIP) 2024



Inhoudsopgave

Voorwoord	4
Lijst van begrippen in de context van de psychodiagnostiek	6
Inleiding	12
Psychodiagnostiek in de praktijk	15
1.1 Opdrachtnemer	15
1.1.1 <i>De opdrachtgever</i>	15
1.1.2 <i>Vorbereiding en beoordeling van de vraag en de opdrachtgever</i>	15
1.1.3 <i>Opdrachtacceptatie en -bevestiging door de diagnosticus</i>	15
1.2 Onderzoeksprocedure	16
1.2.1 <i>Uitnodiging van de cliënt</i>	16
1.2.2 <i>Dossiervorming</i>	16
1.2.2.1 <i>Ruwe testgegevens</i>	16
1.2.2.2 <i>Persoonlijke werkaantekeningen</i>	17
1.2.3 <i>Beschrijving van de onderzoeksvraag en de onderzoeksmethode</i>	17
1.2.3.1 <i>Wanneer worden psychodiagnostische instrumenten ingezet</i>	20
1.2.3.2 <i>Principes bij de keuze van psychodiagnostische instrumenten</i>	25
1.2.4 <i>Afname van psychodiagnostische instrumenten</i>	28
1.2.4.1 <i>Verantwoordelijkheid voor afname</i>	30
1.2.4.2 <i>Toezicht bij afname</i>	30
1.2.4.3 <i>Testruimte bij afname</i>	32
1.2.4.4 <i>Kwaliteit testmateriaal bij afname</i>	32
1.2.4.5 <i>Beveiliging van tests</i>	32
1.2.5 <i>Criterium-georiënteerde test scores, norm-georiënteerde test scores en testnormen</i>	32
1.2.6 <i>Toepassing psychodiagnostische instrumenten bij cliënten met specifieke individuele kenmerken</i>	36
1.3 Het psychologisch rapport	38
1.3.1 <i>Onderdelen van het psychologisch rapport</i>	39
1.3.2 <i>Geautomatiseerde rapportages</i>	42
1.3.3 <i>Verantwoording</i>	42
1.3.4 <i>Rechten van de cliënt</i>	42
1.3.5 <i>Verstrekking van gegevens verkregen uit tests aan derden</i>	44
1.4 Dossiervoering	44
Referenties	45
Totstandkoming van de Algemene Standaard Testgebruik NIP 2024	46

Voorwoord

Psychodiagnostiek heeft als doel om op onderbouwde manier vragen over emoties, cognitie en/of gedrag van een persoon te beantwoorden. Hierbij wordt op systematische, professionele manier informatie verzameld en geïnterpreteerd om zo inzicht te krijgen in relevante capaciteiten, eigenschappen en problemen van een persoon. Goede psychodiagnostiek is van belang bij een groot aantal vragen, zoals het identificeren van een specifieke aandoening, het indiceren van de meest passende (psychologische) behandeling, en het verstrekken van relevante adviezen met betrekking tot educatie en loopbaanplanning.

Het uitvoeren van effectieve psychodiagnostiek is een uitdaging. Iedere persoon is uniek, waarbij factoren als achtergrond, persoonlijkheid, en omgeving van invloed zijn op emoties, cognitie en gedrag. Psychodiagnostische instrumenten kunnen van grote waarde zijn voor een nauwkeurige beoordeling van factoren die van belang zijn voor de individuele vraag. De selectie en correcte toepassing van psychodiagnostische instrumenten in de praktijk is niet eenvoudig. Er zijn veel psychodiagnostische instrumenten beschikbaar, en het vereist een hoge mate van expertise om deze correct te selecteren en te gebruiken. Het doel van deze Algemene Standaard Testgebruik NIP (afgekort: AST-NIP) is om diagnostici te ondersteunen bij hun psychodiagnostisch werk, met een specifieke focus op het gebruik van psychodiagnostische instrumenten. Het doel is om de kwaliteit en consistentie van de psychodiagnostiek te behouden en zondig te verbeteren.

Deze AST-NIP is een herziening van de versie uit 2017. De AST-NIP is gebaseerd op de herziene Beroepscode voor psychologen van het NIP, deze is per 1 april 2024 in werking getreden (NIP, 2024). De Beroepscode biedt psychologen een leidraad voor het beroepsmatig handelen en biedt een belangrijke ondersteuning bij het maken van ethische afwegingen in hun werkzaamheden. Het beschrijft de ethische principes van en gedragsregels voor het beroepsmatig handelen en de professionele relatie tussen psychologen en cliënten, externe opdrachtgevers en/of andere betrokkenen. NIP-leden en/of NIP-geregistreerden zijn gebonden aan deze Beroepscode. De AST-NIP kan gezien worden als een uitgewerkte versie van en toelichting op onderdelen van de Beroepscode, met een specifieke focus op het gebruik van psychodiagnostische instrumenten in het kader van psychodiagnostiek (inclusief de evaluatie van psychologische interventies). De AST-NIP beschrijft algemene richtlijnen waaraan een verantwoorde keuze en gebruik van tests moeten voldoen.

Het toezicht op de naleving van de Beroepscode wordt binnen het NIP uitgevoerd door twee onafhankelijke tuchtrechtelijke instanties: het College van Toezicht (CvT) en het College van Beroep (CvB) van het NIP. Personen die bezwaren hebben tegen het professionele handelen van een NIP-psycholoog en/of NIP-geregistreerd psycholoog en zich hierdoor direct benadeeld voelen, kunnen een klacht indienen bij het CvT. Zowel de klager als de psycholoog kunnen beroep aantekenen tegen de uitspraak van dit

college bij het CvB. Organisaties zoals de Stichting Kwaliteitsregister Jeugd (SKJ) hebben hun eigen tucht-rechtprocedures. Ook daar kan het professioneel handelen van geregistreerden worden getoetst aan de Beroepscode van het NIP, indien van toepassing. Tegen BIG-geregistreerde gezondheidszorgpsychologen en psychotherapeuten kan een klacht worden ingediend bij de Regionale Tuchtcolleges voor de Gezondheidszorg en het Centraal Tuchtcollege voor de Gezondheidszorg. Deze tuchtcolleges toetsen aan algemene wettelijke tuchtnormen waarbij de Beroepscode kan worden gebruikt om deze normen nader in te vullen.

We hopen van harte dat de AST-NIP een nuttige informatiebron is voor iedere professional die psychodiagnostiek bedrijft.

De AST-NIP is op 16 september 2024 vastgesteld door het Algemeen Bestuur van het Nederlands Instituut van Psychologen.

November 2024

***Commissie Testaangelegenheden Nederland (COTAN)
Centrale Commissie Beroepsethische Zaken NIP***

Lijst van begrippen in de context van de psychodiagnostiek

Basisaantekening Psychodiagnostiek (BAPD)

een kwaliteitskeurmerk van het NIP dat aangeeft dat een psycholoog (in opleiding) beschikt over basiskennis en basisvaardigheid op het gebied van de algemene psychodiagnostiek.

Beroepscode voor psychologen (2024)

een leidraad of ethisch kompas dat is opgesteld door het NIP, om psychologen een richtsnoer te bieden voor hun beroepsmatig handelen. Het doel van een beroepscode is om ethische overwegingen en verantwoord beroepsmatig handelen van psychologen te bevorderen, en om de belangen van cliënten, patiënten en de samenleving te beschermen. De Beroepscode voor psychologen bevat basisprincipes en richtlijnen met betrekking tot kwesties zoals vertrouwelijkheid, integriteit, respect voor de autonomie van cliënten, het vermijden van belangenconflicten, en professionele competentie. Het niet naleven van de Beroepscode kan leiden tot disciplinaire maatregelen in het kader van het (verenigings)tucht-recht.

Besliskunde

gaat over hoe mensen beslissingen nemen en hoe deze kunnen worden geoptimaliseerd met behulp van een beslissingsanalyse. Dit is een methode om beslissings-situaties te definiëren in termen van goedgestructureerde beslisproblemen die met behulp van wiskundige en statistische benaderingen zijn op te lossen. Deze analyse helpt mensen bij het formuleren van hun voorkeuren, het zoeken naar alternatieven en het afwegen van de voordelen en de nadelen. Dit kan worden toegepast op het nemen van beslissingen met betrekking tot psychologische beoordeling, behandeling en

onderzoek, zoals het helpen bij het kiezen van de meest geschikte behandelingsbenadering of het evalueren van de effectiviteit van interventies.

Betrouwbaarheid van een meetinstrument

de mate waarin de testcores op het meetinstrument consistent zijn over verschillende metingen bij een persoon, waarbij de te meten eigenschap van die persoon niet verandert. Vaak wordt de betrouwbaarheid van een test echter vastgesteld door de test slechts één keer af te nemen en de betrouwbaarheid te berekenen op grond van de samenhang tussen de items. Een hogere betrouwbaarheid betekent een lagere meetfout op de test.

Betrouwbaarheidsinterval

een interval waarbinnen een parameter (zoals een populatiegemiddelde of een ware testscore) met een gegeven zekerheid valt. Het 95%-betrouwbaarheidsinterval van de testscore geeft aan dat de ware testscore (dus zonder meetfout) met 95% zekerheid in dit interval valt. Een smaller 95%-betrouwbaarheidsinterval duidt dus op minder onzekerheid in de waarde van de ware testscore dan een breder interval. Omdat bij de interpretatie van een testscore de onzekerheid als gevolg van meetfout meegewogen moet worden, moet de diagnosticus zich baseren op een betrouwbaarheidsinterval in plaats van alleen de testscore zelf.

Bewaartermijn gegevens

de periode gedurende welke persoonsgegevens in het dossier op grond van de Beroepscode of een wettelijke bewaartermijn dienen te worden opgeslagen en bewaard.

Cliënt

de persoon met wie de psycholoog een professionele relatie aangaat, onderhoudt of onderhouden heeft.

Cliëntstelsysteem

een aantal personen in hun onderling functioneren, met wie de psycholoog een professionele relatie aangaat, onderhoudt of onderhouden heeft.

Autoriteit Persoonsgegevens

de Nederlandse overheidsinstantie die toezicht houdt op de bescherming van persoonsgegevens.

Centrale Commissie Beroepsethische Zaken

een centrale commissie van het NIP, die uitvoering geeft aan de missie van het NIP door het stimuleren van het beroepsethisch bewustzijn en handelen van de leden. De commissie draagt zorg voor een actuele Beroepscode en handreikingen, werkt mee aan richtlijnen en geeft voorlichting aan de leden.

Commissie Testaangelegenheden Nederland (COTAN)

een commissie van het Nederlands Instituut van Psychologen (NIP) die als missie heeft om de kwaliteit van tests en testgebruik in Nederland te bevorderen.

Criterium-georiënteerde testscore

een testscore die bepaald is op basis van criterium-georiënteerde normen; hierbij wordt de door de cliënt behaalde ruwe (sub)testscore vergeleken met een vooraf vastgestelde standaard, zoals bijvoorbeeld cut-off waarden die leiden tot een classificatie als 'laag risico', 'medium risico' of 'hoog risico'.

Deductie en inductie

het proces van redeneren van algemene principes naar specifieke conclusies (deductie) of van specifieke waarnemingen naar algemene conclusies (inductie).

Diagnosticus

een professional die gespecialiseerd is in het stellen van beoordelingen en diagnoses in diverse psychologische contexten. De rol van een diagnosticus omvat het verzamelen van relevante gegevens en het verstrekken van beoordelingen en aanbevelingen die passen bij de specifieke toepassingsgebieden, zoals neuropsychologische diagnostiek, personeelselectie, onderwijs of andere gebieden waar psychologische beoordelingen van toepassing zijn.

Doelgroep

de groep van personen voor wie de test geschikt is om het beoogde begrip te meten. De testontwikkelaar moet de doelgroep zo goed mogelijk definiëren, bijvoorbeeld in de vorm van inclusie- en exclusiecriteria. De doelgroep van de test hoeft niet gelijk te zijn aan de referentiepopulatie van de test. Hierbij valt te denken aan een neuropsychologische test die bedoeld is voor afname bij patiënten met niet-aangeboren hersenletsel (doelpopulatie), waarbij de testscore vergeleken wordt met die van niet-patiënten (referentiepopulatie).

Dossier

de op een cliënt of cliëntstelsysteem betrekking hebbende verzameling van alle gegevens, die de psycholoog in zijn beroepsmatig handelen heeft verkregen en die deze bewaart vanwege hun relevantie voor kwaliteit en continuïteit van de professionele relatie. Persoonlijke werkaantekeningen van de psycholoog behoren niet tot het dossier.

European Federation of Psychologists' Associations (EFPA)

de overkoepelende federatie van psychologenverenigingen in Europa.

EFPA Standards for Test Use (EFPA, 2012)

een richtlijn waarin algemene competenties en kwalificaties voor testgebruik worden beschreven. Deze richtlijn is opgesteld door de EFPA.

Equivalentie

de mate waarin een test of meetinstrument vergelijkbare resultaten oplevert bij verschillende groepen of omstandigheden of waarin twee verschillende testversies van een test (bijv. papier vs. digitaal, eerdere vs. nieuwe versie, Nederlands vs. andere taal) vergelijkbare resultaten opleveren.

Fairness

de nauwkeurigheid waarmee de interpretatie van een testscore overeenkomt met het daadwerkelijke niveau van het gemeten begrip, ongeacht de groep waartoe de testgebruiker behoort.

Fairnessmatrijs

een instrument, ontwikkeld door de COTAN, dat informatie biedt over het onderzoek naar de *fairness* van een test.

Geautomatiseerde rapportages

rapporten die automatisch worden gegenereerd met behulp van software of computersystemen.

Hypothesetoetsend onderzoek

onderzoek waarbij hypothesen en toetsingscriteria worden geformuleerd en getest om conclusies te trekken over de waarheid van de hypothesen.

International Test Commission (ITC)

een internationale organisatie die zich bezighoudt met het bevorderen van de kwaliteit van psychologische tests en de kwaliteit van het gebruik van psychologische tests.

Item bias

vertekening in de prestatie op een testitem als gevolg van de invloed van bepaalde kenmerken, zoals geslacht of cultuur.

Kwaliteitsstandaarden

vastgestelde normen en criteria die dienen als referen-

tiepunten voor het meten, evalueren en waarborgen van de kwaliteit van dienstverlening, zorg of producten binnen een specifiek domein. Deze normen zijn gebaseerd op professionele consensus, wetenschappelijke inzichten en vaak ook op ervaringen van patiënten of cliënten.

Meetpretentie

het beoogde doel van een test of datgene wat een test beoogt te meten.

Migratieachtergrond

kenmerk dat weergeeft met welk land een persoon verbonden is op basis van het geboorteland van de ouders of van zichzelf.

Wilsonbekwame cliënt

persoon die die niet in staat is om geïnformeerde beslissingen te nemen over hun eigen behandeling of diensten.

Negatieve predictieve waarde

de waarschijnlijkheid dat een negatief testresultaat daadwerkelijk juist is. Het geeft aan hoeveel procent van de personen met een negatief testresultaat daadwerkelijk de conditie niet heeft. Het wordt beïnvloed door de sensitiviteit, specificiteit en prevalentie van de conditie (net als de positieve predictieve waarde).

Norm-georiënteerde testscore

een testscore die bepaald is op basis van verdelingsgerichte normen, waarbij de door de cliënt behaalde ruwe (sub)testscore wordt omgezet in een gestandaardiseerde score, vergeleken met prestaties in een relevante referentiepopulatie. Een gestandaardiseerde score drukt daarmee uit hoe de cliënt presteert relatief ten opzichte van de referentiepopulatie. Voorbeelden van gestandaardiseerde scores zijn T-scores (gemiddelde = 10, standaarddeviatie (SD) = 3), genormaliseerde IQ-scores (normaal verdeeld, met gemiddelde = 100, SD = 15), en stanines.

Normsteekproef (-populatie)

de steekproef van personen op basis waarvan de testnormen berekend zijn. Deze steekproef moet een goede afspiegeling vormen van de normpopulatie. De normpopulatie is de verzameling van alle referentiepopulaties van de betreffende test.

Onderzoeksvraag

een duidelijke, specifieke vraag die de focus van een psychodiagnostisch onderzoek bepaalt.

Onderzoeksmethode

de procedure of aanpak die wordt gebruikt om onderzoeksvragen te beantwoorden binnen de diagnostiek.

Opdrachtgever

de cliënt of het cliëntstelsel, dan wel de externe opdrachtgever door wie de opdracht wordt gegeven.

Persoonlijke werkaantekeningen

de aantekeningen die de psycholoog maakt ter ondersteuning van diens eigen voorlopige gedachtevorming. Ze zijn tijdelijk van aard, niet noodzakelijk voor een goede beroepsuitoefening en niet bedoeld voor anderen.

Positieve predictieve waarde

de waarschijnlijkheid dat een positief testresultaat daadwerkelijk juist is. Het geeft aan hoeveel procent van de personen met een positief testresultaat daadwerkelijk de conditie heeft. Het wordt beïnvloed door de sensitiviteit, specificiteit en prevalentie van de conditie (net als de negatieve predictieve waarde).

Psychodiagnostiek

het (systematische) proces van het verzamelen, analyseren en interpreteren van gegevens om inzicht te krijgen in het psychologisch functioneren van een persoon.

Psychodiagnostisch instrument

een gestandaardiseerde procedure of instrument dat wordt gebruikt om een of meerdere aspecten van het

psychologische functioneren van een persoon te meten, zoals cognitieve vaardigheden, persoonlijkheidstrekken en emoties. Deze instrumenten helpen psychologen om informatie te verzamelen en te beoordelen over het gedrag, de gedachten en de emoties van een persoon. De resultaten van psychodiagnostische instrumenten bieden inzicht in de persoon en worden gebruikt bij het nemen van beslissingen in verschillende psychologische contexten, zoals klinische behandeling, onderwijs en personeelsselectie. In de AST-NIP worden de termen 'psychodiagnostisch instrument' en 'test' door elkaar gebruikt. Deze termen verwijzen naar verschillende soorten meetinstrumenten, zoals vragenlijsten, observatieschalen, toetsen en gestandaardiseerde interviews.

Psychodiagnostisch proces

een systematische procedure die door psychologen wordt gevolgd om psychologische beoordelingen uit te voeren om één of meer onderzoeksvragen omtrent deze persoon te beantwoorden. Ideaal gezien wordt binnen de psychodiagnostiek gewerkt met een diagnostische cyclus, waarbij expliciet hypothesen worden gesteld, hypothesetoetsingscriteria worden toegepast en gereflecteerd wordt.

Psychologisch rapport

een schriftelijk verslag dat bevindingen en conclusies van psychodiagnostisch onderzoek documenteert.

Psychometrie

het gebied van de psychologie dat zich richt op het ontwikkelen en evalueren van meetinstrumenten en tests.

Psychometrische eigenschappen instrument

kenmerken van een meetinstrument die de kwaliteit van de metingen uitdrukken.

Referentiepopulatie

de specifieke populatie waartegen de testprestatie van een persoon wordt afgezet. Voorbeelden zijn de algemene bevolking van dezelfde leeftijd of de gezonde bevolking

van dezelfde leeftijd, geslacht en opleidingsniveau. De voor een test meest geschikte referentiepopulatie hangt af van de meetpretentie en het doel van de test. Sommige tests hebben meerdere meetpretenties en/of testdoelen, en dan kunnen meerdere referentiepopulaties van belang zijn.

Routine Outcome Monitoring (ROM)

een methodiek voor het systematisch verzamelen en evalueren van gegevens over de effectiviteit van een behandeling.

Ruwe testgegevens

de testvragen of items, samen met de bijbehorende antwoorden en ruwe scores van een cliënt.

Sensitiviteit

het vermogen van een test om ware positieve gevallen correct te identificeren.

Specificiteit

het vermogen van een test om ware negatieve gevallen correct uit te sluiten.

Test

psychodiagnostisch instrument; zie de definitie bij deze term.

Test bias

systematische fouten in testcores of de interpretatie hiervan die voortkomen uit factoren die niet relevant zijn voor de eigenschap die wordt gemeten. Een voorbeeld hiervan is de sociaal-culturele achtergrond van de geteste personen.

Testnormen

de combinatie van alle mogelijke ruwe scores op een test en de bijbehorende gestandaardiseerde scores. Een gestandaardiseerde score drukt uit hoe de persoon scoort ten opzichte van de referentiepopulatie, zoals bijvoorbeeld de algemene bevolking van dezelfde leeftijd

of de gezonde bevolking van dezelfde leeftijd, geslacht en opleidingsniveau.

Unbiased test

een test waarvoor geldt dat de testscore overeenkomt met het werkelijke niveau van het te meten begrip, onafhankelijk van de groep waartoe de testgebruiker behoort. Dit betekent dat als er geen verschil in niveau is tussen groepen, de testcores dan ook gelijk zijn. Als er wel een verschil in niveau is, dan moeten ook de testcores verschillen.

Testuitgevers, testauteurs

organisaties, personen die verantwoordelijk zijn voor het ontwikkelen en publiceren van psychologische tests.

Tuchtcollege, klachtencommissie

een instantie die klachten en geschillen behandelt. Bij tuchtrecht staat het toetsen van het beroepsmatig handelen aan de professionele standaard (waaronder de Beroepscode) centraal. Bij klachtrecht staat een laagdrempelige klachtenbehandeling in de relatie van de cliënt tot de instelling of zelfstandig gevestigde zorgverlener centraal.

(Un)Proctored testafname

het afnemen van een test onder toezicht (*proctored*) of zonder toezicht (*unproctored*).

Validiteit (begripsvaliditeit, criteriumvaliditeit)

de mate waarin een test aan zijn doel beantwoordt; begripsvaliditeit is de mate waarin een test meet wat hij beoogt te meten, en criteriumvaliditeit is in hoeverre de testscore samenhangt met een relevante uitkomstmaat (het criterium) die in het verlengde ligt van het construct dat de test beoogt te meten.

Verwijzer

een persoon, meestal een professional in de gezondheidszorg, zoals een huisarts, psychiater of een andere specialist, die een cliënt doorverwijst naar een psycholoog voor een beoordeling, evaluatie of behandeling.

Ware testscore

de geobserveerde testscore van een persoon minus de meetfout van deze persoon op deze meting. Dit begrip komt uit de klassieke testtheorie, waarbij ervan uitgegaan wordt dat een geobserveerde score gelijk is aan de som van de ware score en de meetfout. Meetfouten moeten zo klein mogelijk zijn, want die zijn willekeurig van meting tot meting en hebben geen informatie over het te meten begrip.

Wet op de Geneeskundige Behandelingsovereenkomst (WGBO)

een wettelijke regeling die de rechten en plichten van zorgverleners en patiënten regelt in Nederland.

Inleiding

De afname van een psychodiagnostisch instrument biedt een gestandaardiseerde benadering voor het verkrijgen van inzicht in een persoon om hiermee één of meer onderzoeksvragen omtrent deze persoon te beantwoorden. Binnen de psychologische professie wordt deze persoon vaak aangeduid als een cliënt. Voorbeelden van cliënten zijn patiënten, leerlingen en sollicitanten. Een psychodiagnostisch instrument is een gestandaardiseerd hulpmiddel dat wordt gebruikt om de eigenschappen van een cliënt te evalueren, met als doel te komen tot een psychologische beoordeling, zowel voor de cliënt zelf als voor anderen, met betrekking tot zaken als behandeling, ontwikkeling, plaatsing of selectie.

Deze context van psychodiagnostiek staat centraal in de AST-NIP. De professional die hierbij betrokken is, is wetenschappelijk geschoold op het gebied van de psychodiagnostiek, zoals een psycholoog of orthopedagoog. Deze AST-NIP duidt deze professional aan met de algemene term 'diagnosticus'.

Wetenschappelijk onderzoek heeft aangetoond dat het gebruik van psychodiagnostische instrumenten belangrijke voordelen biedt ten opzichte van bijvoorbeeld een 'ongewapend oordeel' of de 'klinische blik' die gebaseerd zijn op individuele ervaring, intuïtie, gevoel of inlevingsvermogen, zoals beschreven in **BOX 1** en **BOX 2**. Gegevens verzameld met behulp van psychodiagnostische instrumenten worden doorgaans efficiënter verkregen, zijn betrouwbaarder, meer valide, beter reproduceerbaar en objectiever dan gegevens die zijn verkregen via ongewapende oordelen. Informatie verkregen met psychodiagnostische instrumenten vormt daarom een belangrijke bron in het psychodiagnostische proces. In dit proces kan de diagnosticus ook gebruik

maken van andere bronnen, zoals observaties van de cliënt en informatie verkregen door gesprekken met de cliënt en/of informanten, zoals ouders. De uiteindelijke beoordeling komt tot stand doordat de diagnosticus alle beschikbare informatie integreert.

Om schade aan de cliënt te voorkomen, is het cruciaal dat gegevens over de cliënt op een verantwoorde manier worden verkregen, gebruikt en gedeeld. Bij het gebruik van psychodiagnostische instrumenten kunnen zich twee mogelijke problemen voordoen, die beide kunnen resulteren in een lage kwaliteit van de verkregen informatie over de cliënt:

- Het gebruik van een psychodiagnostische instrument dat niet voldoet aan de vereiste wetenschappelijke criteria, of waarvan niet bekend is of deze hieraan voldoet;
- Het onjuiste gebruik van een psychodiagnostisch instrument, zoals de afname van de test niet conform de handleiding of afname van de test bij een cliënt die niet behoort tot de doelgroep of in een context waarvoor de test niet bedoeld is.

Deze AST-NIP besteedt aandacht aan zowel de wetenschappelijke criteria voor de kwaliteit van de psychodiagnostische instrumenten als het juiste gebruik van de instrumenten. Dit omvat de voorwaarden voor het maken van een goede keuze voor een instrument en het correct toepassen ervan. Het is van groot belang dat een psychodiagnostisch instrument wordt ingezet voor het juiste doel, wordt afgenomen op de juiste manier en dat de resultaten correct worden geïnterpreteerd.

BOX 1 Geen gebruik van psychodiagnostische instrumenten

In uitspraak 14/65 van het College van Toezicht (CvT) van het NIP uit 2015 wordt een klacht van een cliënt over een psychologisch rapport gegrond verklaard.

De inhoud en conclusie van dat psychologisch rapport bleken volgens het College ongefundeerd te zijn, omdat er geen wetenschappelijk onderbouwde psychodiagnostische instrumenten waren gebruikt. Zo was er tijdens het onderzoek door de psycholoog geen intelligentieonderzoek verricht, terwijl in het rapport werd gesteld dat de onderzochte persoon bovengemiddeld intelligent is. Dit werd ook niet door andere observaties of bevindingen onderbouwd.

Het CvT heeft geen uitspraak gedaan over de (on) juistheid van de in het rapport vermelde conclusies. Wel was het volgens het CvT duidelijk dat de conclusies niet waren gefundeerd op (bijvoorbeeld) testresultaten. Het CvT oordeelde dat hiermee door de betrokken psycholoog in strijd is gehandeld met de Beroepscode. Deze dient het beroepsmatig handelen altijd te kunnen verantwoorden in het licht van de stand der wetenschap zoals dat uit de vakliteratuur blijkt (in de vorige beroepscode artikel 106 'Professionele verantwoording van het beroepsmatig handelen' en in de meest recente Beroepscode artikel 107 'De wetenschappelijke basis van het beroepsmatig handelen').

BOX 2 Het rapport geeft geen blijk van een geschikte methode om de vraagstelling te beantwoorden

In uitspraak 21/10 van het College van Toezicht (CvT) van het NIP uit 2021 wordt een klacht over een psychologische rapportage gegrond verklaard en aan de psycholoog een berisping opgelegd.

De psycholoog doet niet alleen uitspraken over de cliënt, de opdrachtgever, maar ook over de ex-partner zonder dat hij haar, klaagster, heeft gesproken en zonder haar toestemming. De cliënt had zich bij de psycholoog gemeld voor een persoonlijkheidsonderzoek. Voor dit doel heeft de psycholoog slechts een volgens het College "niet gebruikelijke" vragenlijst gebruikt en de uitkomsten niet afgezet tegen een normgroep.

Het College stelt vast dat verweerder niet heeft voldaan aan het criterium in artikel 97 van de Beroepscode (2015, artikel 98 Beroepscode 2024) dat het rapport een onderbouwing, of blijk, moet geven van dat bij het onderzoek een geschikte methode van onderzoek is ingezet om de vraagstelling te beantwoorden. Ook heeft de psycholoog uitspraken over de ex-partner gedaan op basis van een niet-geschikte methode en ook de bronnen niet vermeld waarop de bevindingen en conclusies steunen. In dat licht heeft de psycholoog ook in strijd gehandeld met artikel 96 van de Beroepscode (2015, artikel 97 Beroepscode 2024).

Ondeskundig gebruik van een psychodiagnostisch instrument kan schadelijk zijn en kan leiden tot misleiding van de cliënt en/of de opdrachtgever. Daarom is het van groot belang dat de diagnosticus goed is opgeleid in de keuze en het gebruik van psychodiagnostische instrumenten. Voor complexe tests, als intelligentietests, is het essentieel dat de diagnosticus aanvullende training volgt.

Het 'juiste gebruik van een psychodiagnostisch instrument' betekent dat het gebruik van het instrument integraal deel uitmaakt van het psychodiagnostische proces. De diagnosticus moet te allen tijde kunnen verklaren hoe het gekozen instrument verband houdt met de diagnostische vraagstelling. Bovendien moet de diagnosticus kunnen onderbouwen hoe de testresultaten leiden tot de conclusies die de diagnosticus trekt over de cliënt. Tot slot moet de diagnosticus kunnen uitleggen hoe de integratie van alle informatiebronnen, inclusief het psychodiagnostisch instrument, leidt tot het uiteindelijke advies.

De keuze voor de juiste psychodiagnostische instrumenten en het correcte gebruik ervan zijn de verantwoordelijkheid van de diagnosticus. De AST-NIP biedt een leidraad voor verantwoorde keuzes en professioneel gebruik van psychodiagnostische instrumenten. Dit betekent enerzijds dat de AST-NIP als richtlijn fungeert, en anderzijds dat in specifieke gevallen gemotiveerd kan worden afgeweken van de standaard. Als een diagnosticus van plan is om af te wijken van de standaard, is het aan te raden om dit vooraf met collega's te bespreken en de afwijking goed te documenteren (inclusief de onderbouwing) in het cliëntendossier.

Psychodiagnostiek in de praktijk

1.1 Opdrachtnemer

De opdrachtnemer volgens de [Beroepscode voor psychologen](#) is de diagnosticus, en in het bijzonder de gedragswetenschapper die lid is van of geregistreerd is bij het NIP. De diagnosticus is verantwoordelijk voor het gebruik van de tests, het onderzoek en het opstellen van het rapport. Hoewel de Beroepscode dit niet verplicht, raden we aan dat de diagnosticus het psychologisch rapport ondertekent, zodat het voor alle betrokkenen duidelijk is welke diagnosticus de professionele verantwoordelijkheid draagt voor de inhoud van het rapport.

1.1.1 De opdrachtgever

Voordat het psychodiagnostisch onderzoek begint, stelt de diagnosticus vast wie de opdrachtgever is voor het psychodiagnostisch onderzoek (Beroepscode, artikel 1.8 'Opdrachtgever'). Er kan sprake zijn van een externe opdrachtgever (Beroepscode, artikel 1.9 'Externe opdrachtgever'), zoals een werkgever die een psycholoog vraagt om een uitspraak te doen over een sollicitant, een school die een psycholoog om advies en begeleiding vraagt bij een vermoeden van leer- of gedragsproblemen bij een leerling, of een rechter die een psychodiagnostisch onderzoek oplegt.

De cliënt zelf is opdrachtgever (Beroepscode, artikel 1.4 'Cliënt') in een de volgende gevallen:

- als de cliënt direct advies vraagt aan de diagnosticus, bijvoorbeeld een beroepskeuzeadvies;
- als een cliënt zich onder behandeling stelt en de behandelaar een psychodiagnostisch onderzoek uitvoert voor de eigen oordeelsvorming ten bate van de cliënt;
- als een verwijzer (Beroepscode, artikel 1.10 'Verwijzer'),

zoals een arts of een collega-behandelend psycholoog, de cliënt adviseert om naar een psycholoog te gaan voor psychodiagnostisch onderzoek.

Bij minderjarige of wilsonbekwame cliënten worden de rechten die de Beroepscode aan hen toekent, uitgeoefend door hun wettelijk vertegenwoordiger(s) (Beroepscode, artikel 1.11 'Wettelijk vertegenwoordiger(s)'). Deze cliënten moeten echter zo veel mogelijk betrokken worden bij de uitoefening van hun rechten (Beroepscode, artikel 6 'Minderjarige cliënt'; artikel 8 'Meerderjarige wilsonbekwame cliënt').

1.1.2 Voorbereiding en beoordeling van de vraag en de opdrachtgever

De diagnosticus neemt kennis van de vraag van de opdrachtgever en beoordeelt of de vraag in de huidige vorm kan worden beantwoord of dat een herformulering nodig is. Ook beoordeelt de diagnosticus of er geen (beroeps)ethische belemmeringen zijn voor het beantwoorden van de vraag. Voordat het psychodiagnostisch onderzoek begint, heeft de diagnosticus overleg met de cliënt en de eventuele externe opdrachtgever over het doel van de professionele relatie met de cliënt en de wijze van samenwerken (Beroepscode, artikel 63 'Dezelfde informatie voor externe opdrachtgever en cliënt'; artikel 64 'Overleg over invulling van de professionele relatie'). Dit overleg kan mondeling of schriftelijk plaatsvinden.

1.1.3 Opdrachtacceptatie en -bevestiging door de diagnosticus

Op enig moment moeten de diagnosticus, de cliënt en de eventuele externe opdrachtgever overeenstemming bereiken over de opdracht. De diagnosticus formuleert de opdracht in de vorm van een of meer onderzoeksvragen.

De onderzoeksvragen kunnen strikt individueel zijn (bijvoorbeeld bij intake van een individuele cliënt door een psycholoog), maar ook in hoge mate gestandaardiseerd (bijvoorbeeld bij personeelsselectie, als een instrument wordt ingezet om een eerste selectie te kunnen maken uit de potentiële kandidaten). De diagnosticus bevestigt de opdracht schriftelijk (digitaal of op papier) aan alle betrokkenen op dezelfde wijze (Beroepscode, artikel 63 'Dezelfde informatie voor externe opdrachtgever en cliënt'). De betrokkenen zijn de cliënt of het cliëntstelsel en de eventuele externe opdrachtgever. Voorbeelden van een cliëntstelsel zijn een gezin en een groep als een managementteam. De opdrachtbevestiging beschrijft het doel van de professionele relatie, de werkwijze, de rapportagemethode en de naleving van de Beroepscode, met een link naar de Beroepscode (artikel 62 'Informatie bij het aangaan en voortzetten van de professionele relatie'). De diagnosticus kan ervoor kiezen een deel van de opdrachtbevestiging in een standaarddocument vast te leggen en hiernaar te verwijzen. Dit is handig als een diagnosticus regelmatig werkt voor dezelfde opdrachtgever.

1.2 Onderzoeksprocedure

Na de opdrachtacceptatie van een psychodiagnostisch onderzoek met psychodiagnostische instrumenten zijn er de volgende stappen.

1.2.1 Uitnodiging van de cliënt

De cliënt wordt persoonlijk uitgenodigd voor het onderzoek. De uitnodiging is schriftelijk (digitaal of op papier), tenzij dit niet mogelijk is. Dan volgt de uitnodiging op de voor die cliënt meest geschikte manier. Vóór aanvang van het onderzoek wordt de cliënt, bij voorkeur schriftelijk, op de hoogte gebracht van het doel van het onderzoek, de onderzoeksmethode, de rapportagemethode, de wijze waarop inzage wordt verleend, hoe bezwaar kan worden gemaakt tegen het psychologisch rapport, hoe eventuele correcties op het psychologisch rapport kunnen worden verkregen en hoe de eventuele (externe) opdrachtgever of verwijzer op de hoogte wordt gesteld van de onderzoeksresultaten. Pas nadat de cliënt goed geïnformeerd is, kan deze toe-

stemming geven (*informed consent*) voor het uitvoeren van het onderzoek (Beroepscode, artikel 61 'Aangaan en voortzetten van de professionele relatie'; artikel 62 'Informatie bij het aangaan en voortzetten van de professionele relatie').

Bij minderjarige cliënten (tot 16 jaar) of wilsonbekwame cliënten moeten ook de ouder(s) of andere wettelijke vertegenwoordiger(s) van de cliënt vooraf worden geïnformeerd, bij voorkeur schriftelijk. De nadruk blijft hierbij liggen op de zelfbeschikking van de cliënt (Beroepscode, artikel 59 'Respectvol handelen bij beperkte zelfbeschikking').

1.2.2 Dossiervorming

Voorafgaand of direct na de uitnodiging van de cliënt wordt een dossier van de cliënt aangelegd. Dit dossier bevat alle gegevens die van belang zijn voor de kwaliteit en continuïteit van de professionele relatie, inclusief alle gegevens die van belang zijn voor het psychodiagnostisch onderzoek (Beroepscode, artikel 1.14 'Dossier'; artikel 19 'Volledigheid, noodzakelijkheid en actualiteit van het dossier'). Het dossier kan digitaal en/of op papier zijn.

Ieder dossier moet goed beveiligd zijn tegen onrechtmatige toegang. Het dossier is in principe uitsluitend toegankelijk voor de cliënt, de diagnosticus en de direct onder de diagnosticus' verantwoordelijkheid vallende medewerkers, zoals bijvoorbeeld testassistenten en secretariaatsmedewerkers.

1.2.2.1 Ruwe testgegevens

Onder ruwe testgegevens worden de testvragen of test-items verstaan, samen met de bijbehorende antwoorden en ruwe scores van een cliënt (Frima & Visser, 2008). Hieruit worden meestal genormeerde scores berekend. De ruwe testgegevens en genormeerde scores maken deel uit van het dossier. Als de diagnosticus werkt met een digitaal dossier, zoals het elektronisch patiëntendossier, mag ook (een deel van) de ruwe testgegevens en genormeerde scores in een papieren dossier gearchiveerd worden. Testvragen of items mogen ook buiten het digitale dossier worden bewaard, op voorwaarde dat er in het dossier

expliciet wordt verwezen naar de desbetreffende test en dat de test beschikbaar is voor inzage door de cliënt.

Onder testuitgevers, testauteurs en de beroepsgroep van psychologen is er lange tijd discussie geweest over het verstrekken van ruwe testgegevens aan de cliënt. Volgens de Beroepscode heeft de cliënt het recht om inzage te krijgen in en afschriften te ontvangen van het gehele dossier (artikel 66 'Inzage en afschrift van het dossier'), met uitzondering van persoonlijke werkaantekeningen van de diagnosticus (zie paragraaf 1.2.2.2 'Persoonlijke werkaantekeningen'). Dit is problematisch in het geval van ruwe testgegevens, omdat testvragen auteursrechtelijk beschermd zijn. Bovendien leidt openbaarheid mogelijk tot minder betrouwbare en minder valide metingen en daarmee tot een vermindering van de waarde van psychodiagnostische instrumenten. Op basis van een uitspraak van het College Bescherming Persoonsgegevens (CBP; nu de Autoriteit Persoonsgegevens) is het dossierrecht beperkt door de cliënt alleen inzage (en geen afschrift) in de testvragen te bieden tijdens een gesprek, gecombineerd met een schriftelijk rapport met de resultaten van de test (Autoriteit Persoonsgegevens, 2008). Dit betekent dus dat er geen kopie van de testvragen wordt verstrekt en als de antwoorden inzicht geven in de testvragen ook geen kopie van de antwoorden. Op deze manier behouden psychologische tests hun waarde en blijven zij auteursrechtelijk beschermd.

1.2.2.2 Persoonlijke werkaantekeningen

Persoonlijke werkaantekeningen van de diagnosticus behoren niet tot het dossier (Beroepscode, artikel 1.14 'Dossier'). Deze werkaantekeningen zijn tijdelijke aantekeningen die bedoeld zijn voor de eigen, voorlopige gedachtevorming en die bij de diagnosticus gerezen indrukken, vermoedens of vragen betreffen (Beroepscode, artikel 1.15 'Persoonlijke werkaantekeningen'). Aspecten die van belang zijn voor het diagnostische proces, zoals een weergave van het gesprek en observaties of indrukken met betrekking tot de cliënt, moeten wel in het dossier worden opgeslagen. Persoonlijke werkaantekeningen moeten apart (bijvoor-

beeld digitaal afgeschermd) van het dossier worden bewaard. Het is essentieel dat alleen de diagnosticus zelf de persoonlijke werkaantekeningen kan inzien. Zodra een werkaantekening niet meer nodig is, moet deze vernietigd worden. Dit kan zijn omdat de werkaantekening niet meer van belang is (bijvoorbeeld omdat een vraag in het kader van de voorlopige gedachtevorming inmiddels is beantwoord), of omdat de aantekening is opgenomen in het dossier (omdat ze van belang is in het diagnostische proces).

1.2.3 Beschrijving van de onderzoeksvraag en de onderzoeksmethode

In de opdracht (zie paragraaf 1.1.3 'Opdrachtacceptatie en -bevestiging door de diagnosticus') staan een of meerdere onderzoeksvragen geformuleerd. De diagnosticus bepaalt welke onderzoeksmethode gebruikt wordt om deze onderzoeksvragen te beantwoorden en onderbouwt deze. Uit deze beschrijving moet af te leiden zijn om welke reden(en) en met welk doel de diagnosticus de onderzoeksmethoden, waaronder psychodiagnostische instrumenten, inzet. Op deze manier legt de diagnosticus verantwoording af van diens beroepsmatig handelen (Beroepscode, artikel 33 'Afleggen van verantwoording').

De beschrijving van onderzoeksvraag en onderzoeksmethode omvat:

- De reden voor het onderzoek en de vraagstelling van de opdrachtgever;
- De opdracht van de cliënt en de eventuele opdrachtgever in de vorm van een of meer onderzoeksvragen;
- De onderzoeksvraagstelling(en) en bijbehorende verwachtingen en hypothesen;
- De keuze van methode en instrumenten voor het beantwoorden van de onderzoeksvraagstelling(en), toetsingscriteria en de onderbouwing hiervan (zie ook 1.2.3.2 'Principes bij de keuze van psychodiagnostische instrumenten').

De begrippen ‘aanleiding’, ‘vraagstelling van de opdrachtgever’, ‘opdracht’ en ‘onderzoeksvragen’ wijzen op verschillende fases in het tot stand komen van een overeenkomst tot het uitvoeren van onderzoek. De aanleiding kan bijvoorbeeld zijn ‘een slecht schoolresultaat’. De vraagstelling van de opdrachtgever, zoals een ouder of leerkracht, kan dan zijn ‘Is er sprake van een specifieke leerstoornis, of heeft dit kind onvoldoende cognitieve capaciteiten?’.

Onderzoeksvragen zijn vragen die gerelateerd zijn aan de hulpvraag van de opdrachtgever. De onderzoeksvragen kunnen worden uitgebreid met vragen die de diagnosticus relevant vindt op basis van eigen ervaring, vakliteratuur en verzamelde gegevens, zoals tijdens een intake, observatie of gesprek. De onderzoeksvraag bij bovengenoemde vraagstelling van de opdrachtgever kan dan zijn ‘In hoeverre zijn de intelligentie, de leervaardigheden, motivatie en/of interesse bij de leerling beperkende factoren bij het bereiken van goede schoolresultaten?’. In de ‘methode en instrumenten’ wordt dan bijvoorbeeld aangegeven dat een intelligentietest, een prestatietest en een motivatietest worden afgenomen en een vragenlijst om de wijze te bepalen waarop de leerling omgaat met problemen en stressvolle gebeurtenissen. Dit laatste kan bijvoorbeeld gemotiveerd zijn door de observatie tijdens de intake dat de leerling nogal gelaten reageert op falen of slechte schoolresultaten.

Na het formuleren van de onderzoeksvraag of -vragen en voordat gegevens worden verzameld om deze vraag of vragen te beantwoorden, is het cruciaal dat de diagnosticus een stap terugneemt. Hierbij reflecteert de diagnosticus op diens eigen verwachtingen en hypothesen met betrekking tot de gestelde vragen. Deze overwegingen zijn gebaseerd op de eigen ervaring, de bestaande vakliteratuur, standaarden en de tot dan toe verzamelde gegevens. Het is van essentieel belang dat de diagnosticus op dat moment (en later in het proces) bewust nadenkt over mogelijke vooroordelen of biases die tijdens het diagnostisch proces kunnen optreden. Elke casus kan beïnvloed worden door onbewuste aannames of beperkingen in de waarneming van de diagnosticus. Om deze mogelijke biases tegen te gaan, is

het noodzakelijk dat de diagnosticus een kritische houding aanneemt en probeert om de eigen persoonlijke aannames en vooroordelen te identificeren en te corrigeren. Door open te staan voor alternatieve interpretaties en verschillende perspectieven, kan de diagnosticus een meer evenwichtige en nauwkeurige evaluatie uitvoeren. Dit draagt bij aan het waarborgen van objectiviteit en nauwkeurigheid in het diagnostisch proces, waardoor de uiteindelijke conclusies en aanbevelingen beter onderbouwd zijn.

Voorafgaand aan de testafname(s) moet de diagnosticus de keuze van methode en instrumenten voor het beantwoorden van de onderzoeksvraag en de onderbouwing hiervan expliciet beschrijven. Een voorbeeld van een onderbouwing van de keuze van methode en instrumenten is het beschrijven van het theoretisch kader waaruit is gewerkt en de consequenties van de keuze voor een bepaald instrument voor de inhoudelijke interpretatie. Uit deze beschrijving moet ook de relatie tussen de vraagstelling en de gebruikte psychodiagnostisch instrument(en) duidelijk zijn. Indien er een professionele standaard beschikbaar is en van toepassing is op de onderzoeksvraag, volgt de diagnosticus deze standaard. Hierbij kan men denken aan kwaliteitstandaarden, zoals die gepresenteerd worden door [GGZ Standaarden](#) en Richtlijnen jeugdhulp en jeugdbescherming. Deze bieden algemene aanbevelingen voor het voorkomen, herkennen, diagnosticeren en behandelen van psychische klachten en stoornissen bij kinderen, jongeren en volwassenen.

Bij hypothesetoetsend onderzoek wordt ook expliciet beschreven welke methode(n) de diagnosticus hanteert om de hypothese(n) te toetsen en welke instrumenten daarbij gebruikt worden; zie **BOX 3**. De diagnosticus formuleert voorafgaand aan het onderzoek criteria op basis waarvan de hypothese(n) verworpen of aanvaard worden. Bij gebruik van een instrument met een criterium-georiënteerde testscore, levert het instrument al een classificatie (zoals ‘een vermoeden tot depressie’), die mogelijk direct gebruikt kan worden voor de hypothesetoetsing. Bij gebruik van een instrument met een norm-georiënteerde testscore, kan

BOX 3 Hypothesetoetsingsmodellen

In de literatuur worden diverse modellen beschreven die als leidraad dienen voor het nemen van psychodiagnostische beslissingen. Deze modellen vertonen veel overeenkomsten. Als illustratie wordt een van deze modellen, de diagnostische cyclus van De Bruyn en collega's (2003), hier kort besproken. In de diagnostische cyclus worden vier fasen onderscheiden binnen het psychodiagnostisch proces: de klachtenanalyse, de probleemanalyse, de verklaringsanalyse en de indicatieanalyse.

In de eerste fase, de klachtenanalyse, verzamelt de diagnosticus gegevens over de klachten van de cliënt en/of diens omgeving en bespreekt deze met de cliënt. Indien de beschrijving van de klachten door de betrokkenen vragen of onduidelijkheden oproept, gaat de diagnosticus in gesprek met betrokkene(n) om de klachten te verhelderen. Deze analyse dient te resulteren in een formulering van hulpvragen. De Bruyn en collega's noemen dit ‘de verhelderende diagnostiek’.

In de tweede fase, de probleemanalyse, legt de diagnosticus een verband tussen de gerapporteerde klachten van de cliënt en/of diens omgeving en problemen, en schat de diagnosticus in wat de ernst van de problemen is. De Bruyn en collega's noemen dit ‘onderkende diagnostiek’.

In de derde fase, de verklaringsfase, stelt de diagnosticus hypothesen op en toetst deze aan de hand van vooraf gedefinieerde toetsingscriteria. In deze fase

worden psychodiagnostische instrumenten afgenomen. Op basis van de beschrijving van de cliënt en/of diens omgeving, observaties en testresultaten vormt de diagnosticus een beeld dat met een bepaalde mate van waarschijnlijkheid kan dienen als een verklaring van de problemen die zijn gedefinieerd in de tweede fase. Dit wordt volgens De Bruyn en collega's ‘de verklarende diagnose’ genoemd.

In de vierde fase, de indicatieanalyse, formuleert de diagnosticus een of meerdere behandelings- of begeleidingsadviezen, ook wel ‘indicaties’ genoemd. Deze adviezen zijn gebaseerd op gegevens die zijn verzameld in de eerdere fasen en dienen besproken te worden met de cliënt. De Bruyn en collega's noemen dit ‘de indicerende diagnose’.

Aan het einde van iedere fase toetst de diagnosticus diens hypothesen (of veronderstellingen) en theorieën aan de hand van de verkregen gegevens.

een grenswaarde van de testscore gebruikt worden om tot een classificatie te komen. Het is dan wel belangrijk dat de diagnosticus zich realiseert dat de precieze specificiteit en sensitiviteit van de betreffende classificatie niet bekend zijn, omdat deze niet zo gemakkelijk te berekenen is. Een indicatie geeft de breedte van het 90% betrouwbaarheidsinterval (90%BHI): Hoe smaller het 90%BHI, hoe preciezer, dus hoe hoger de specificiteit en sensitiviteit.

Het gebruik van een grenswaarde met een norm-georiënteerde testscore kan als volgt gedaan worden. Stel dat een diagnosticus bij een onderzoek naar een mogelijke ontwikkelingsachterstand een ontwikkelingstest afneemt bij een kind van 38 maanden. De test geeft als score een Ontwikkelingsindex, die in de referentiepopulatie (i.e., zich op de gebruikelijke wijze ontwikkelende kinderen van dezelfde leeftijd en geslacht) normaal verdeeld is met een gemiddelde heeft van 100, en een standaarddeviatie van 15. Als grens voor een vermoeden van een ontwikkelingsachterstand hanteert de diagnosticus een ware score van 81 of lager (dit komt overeen met de 10% laagst scorende kinderen). Een kind kan dan gevlagd worden als 'vermoeden tot een ontwikkelingsachterstand' als het 90% betrouwbaarheidsinterval (BHI) de waarde 81 of lager bevat. Dus, een 90%BHI van 72 tot 79, en een 90%BHI van 75 tot 83 zouden geïnterpreteerd worden als 'vermoeden tot een ontwikkelingsachterstand', en een 90%BHI van 82 tot 93 niet.

Ook als niet gewerkt wordt met hypothese(n), moet de diagnosticus expliciet aangeven hoe alle verzamelde gegevens geïntegreerd zullen worden om tot een conclusie te komen en de onderzoeksvragen te beantwoorden.

De wijze waarop de integratie van verzamelde gegevens tot de conclusie zal plaatsvinden wordt niet vooraf met de cliënt gedeeld, omdat dit mogelijk tot een ongewenste beïnvloeding leidt.

1.2.3.1 Wanneer worden psychodiagnostische instrumenten ingezet

De afname van een of meerdere psychodiagnostische instrumenten is doorgaans belastend voor de cliënt en mogelijk voor anderen in diens omgeving, zoals ouders, een partner of leerkracht. Om personen niet onnodig te belasten, moet het gebruik van een instrument goed kunnen worden onderbouwd. Dit betekent dat de diagnosticus een instrument alleen moet inzetten als dit naar verwachting zinvolle informatie oplevert om de onderzoeksvraag (of –vragen) te beantwoorden, en daarmee de kwaliteit van de uitspraken of adviezen over een persoon duidelijk verhoogt. De diagnosticus moet dus kunnen uitleggen waarom het zinvol is om deze test(s) af te nemen bij deze specifieke cliënt. Het feit dat een instelling beleid heeft om altijd dezelfde test(s) af te nemen bij elke aangemelde cliënt, is bijvoorbeeld op zichzelf geen voldoende rechtvaardiging. Tegelijkertijd moet de diagnosticus kunnen verantwoorden waarom die ervoor kiest om bepaalde psychodiagnostische instrumenten niet in te zetten, zie ook **BOX 1** en **BOX 2**.

De beslissing om al dan niet een instrument in te zetten, moet door de diagnosticus worden onderbouwd met verwijzingen naar relevante literatuur of deducties op basis van deze literatuur. De kwaliteitstoename van het diagnostisch proces door het gebruik van psychodiagnostische instrumenten kan blijken uit een toename van validiteit, betrouwbaarheid en/of efficiency in vergelijking met andere (vaak niet-gestandaardiseerde) informatiebronnen, zoals schoolcijfers, werkprestaties, anamnesegegevens, praktijk-simulatie of deskundigenoordelen. In situaties waarin geen (gestandaardiseerde) psychodiagnostische instrumenten beschikbaar zijn die aan de gebruikelijke kwaliteitscriteria voldoen, moet de diagnosticus nagaan of er een informatiebron beschikbaar is die bruikbaar is om de benodigde informatie te verkrijgen. Voorbeelden hiervan zijn observaties van de cliënt en informatie verkregen door gesprekken met de cliënt en/of informanten, zoals ouders. Het kan voorkomen dat de conclusie moet zijn dat de gestelde onderzoeksvraag niet beantwoord kan worden omwille van een gebrek aan betrouwbare en valide informatie.

BOX 4

Afname van hetzelfde meetinstrument binnen kort tijdsinterval

Cliënt is een jongen van 7 jaar, waarbij de verwijzer aan psycholoog X vraagt om te onderzoeken of zijn gedrag passend is bij de classificatie ADHD. Psycholoog X kiest er voor de cliënt te onderzoeken door het afnemen van verschillende gedragsvragenlijsten en intelligentietest Y. Ouders van de cliënt uiten twijfels over de noodzaak van het afnemen van intelligentietest Y. Ze vragen waarom een intelligentietest nodig is en waarom juist deze specifieke intelligentietest wordt gebruikt. Bovendien merken de ouders op dat dezelfde test zes maanden geleden al is afgenomen door een geregistreerd diagnosticus, die destijds in zijn rapport aangaf dat de test op een betrouwbare manier is afgenomen.

Toelichting BOX 4

Conform de Beroepscode is de herhaalde afname van de intelligentietest Y in deze casus alleen geoorloofd als psycholoog X kan onderbouwen dat deze herhaalde afname aanvullende gegevens oplevert die bijdragen aan de diagnostiek, oftewel in dit geval waardevol zijn voor de classificatie ADHD.

Psycholoog X moet ook verantwoorden hoe de psychometrische kwaliteit van intelligentietest Y bij de herhaalde afname wordt gehandhaafd, aangezien er inderdaad leer- en procedure-effecten kunnen optreden als de intelligentietest binnen een kort tijdsbestek herhaaldelijk wordt afgenomen. Er moet rekening gehouden worden met de mogelijkheid dat de norm-

gegevens, gebruikt om de ruwe scores om te zetten naar standaardcores, wellicht niet geschikt zijn voor herhaalde metingen. Dit komt doordat deze normgegevens meestal zijn gebaseerd op een eenmalige afname. Onderzoek naar cognitieve tests, zoals de prestaties van mensen over het algemeen verbeteren bij herhaalde afname van cognitieve tests. Deze verbetering is het meest uitgesproken wanneer bij de eerste en de tweede meting hetzelfde instrument ingezet wordt (Collie et al., 2003).

BOX 5

Gebruik instrument waarvan de psychometrische eigenschappen nog niet zijn aangetoond

Psycholoog X weet dat universiteit A een aandachts-test aan het ontwikkelen is. De universiteit is bezig met het onderzoeken van de psychometrische eigenschappen van dit nieuwe instrument. Psycholoog X besluit na overleg met de universiteit om dit instrument af te nemen bij cliënt C, zodat de psycholoog het kan gebruiken als additionele maat om de aandachtsregulatie van deze cliënt in kaart te brengen; bovendien kan de psycholoog zo de universiteit onderzoeksgegevens verschaffen. Ondanks dat psycholoog X denkt dat het instrument veel potentie heeft, vraagt die zich wel af of de belasting van de afname van dit instrument en het niet hebben van zekerheid dat dit een psychometrisch goed instrument is, wel te verantwoorden is.

Toelichting BOX 5

Een vraag die bij deze casus onder meer gesteld moet worden is of aannemelijk te maken is dat de afname van de aandachtstest voldoende bijdraagt aan de klinische zorg voor cliënt C, of dat het instrument bij cliënt C primair in het kader van wetenschappelijk onderzoek wordt afgenomen. Gezien het feit dat de test nog in ontwikkeling is en de psychometrische eigenschappen van de test nog niet zijn vastgesteld, is de afname niet gerechtvaardigd in het kader van de klinische zorg. Als de psychometrische eigenschappen al bekend zouden zijn en hieruit zou blijken dat de test passend is gezien de onderzoeksvraag van cliënt C, dan was afname wel gerechtvaardigd. De Beroepscode (artikel 85 'Informatieverstrekking voor

wetenschappelijk onderzoek') stelt dat de psycholoog desgevraagd aan een derde gegevens over de cliënt mag verstrekken ten behoeve van onderzoek. Dit mag zonder toestemming van de cliënt als de gegevens geanonimiseerd zijn, dat wil zeggen dat herkenbaarheid van de persoon uitgesloten is. Als dit niet mogelijk is dan mogen de gegevens alleen met toestemming van de cliënt worden verstrekt (Beroepscode, artikel 86 'Gegevens voor publicaties, onderwijs, kwaliteitszorg, supervisie en intervisie').

Er zijn verschillende situaties te onderscheiden waarin een cliënt mogelijk onnodig wordt belast en afname van het meetinstrument daarmee niet gerechtvaardigd is. De eerste situatie is dat recentelijk een instrument met een vergelijkbare meetpretentie is afgenomen en het niet kan worden onderbouwd dat herhaling of aanvullend onderzoek noodzakelijk is, zie ook **BOX 4**. Voorbeelden van situaties waarbij herhaling of aanvullend onderzoek wel gerechtvaardigd is, zijn bij evaluatie tijdens of na een interventie en als er behoefte is aan verdiepend onderzoek op een specifiek domein. Bij herhaling moet wel worden overwogen hoe procedure- en leereffecten bij de cliënt worden beperkt, zoals beschreven in **BOX 4**.

De tweede situatie is dat de psychometrische kwaliteit van het instrument onvoldoende bekend is (zie ook 1.2.3.2 'Principes bij de keuze van psychodiagnostische instrumenten'). Dit kan zijn omdat het een nieuw instrument betreft waarvan de psychometrische eigenschappen nog niet zijn vastgesteld, zie ook **BOX 5**, of omdat deze eigenschappen niet vastgesteld zijn voor de doelgroep waartoe de cliënt behoort. Als de diagnosticus niet aannemelijk kan maken dat de inzet van het instrument te verantwoorden is, moet dit instrument niet ingezet worden.

De derde situatie is dat de individuele kenmerken van de cliënt maken dat het psychodiagnostische instrument niet geschikt is voor gebruik bij deze cliënt. Voorbeelden van dergelijke individuele kenmerken zijn fysieke of verstandelijke beperkingen, niet-Nederlandstaligheid en een niet-Nederlandse culturele achtergrond van de cliënt. Ieder psychodiagnostisch instrument heeft een bepaalde doelgroep waarvoor het instrument is ontwikkeld en onderzocht op de psychometrische eigenschappen. Buiten deze doelgroep is het instrument niet zonder meer toepasbaar. Een instrument dat bijvoorbeeld is ontwikkeld en onderzocht voor het Amerikaanse cultuur- en taalgebied, is na een letterlijke vertaling naar het Nederlands niet zonder meer toepasbaar binnen Nederland. Voor meer informatie over de vertaling en bewerking van een instrument in een andere taal, zie de 'ITC Guidelines for Translating and

Adapting Tests' (International Test Commission, 2017) en Van de Vijver en Leung (2021). Een instrument geschikt voor de algemene Nederlandse bevolking is daarnaast mogelijk ongeschikt of minder geschikt voor gebruik bij een persoon met een migratieachtergrond. In sommige gevallen is de ongeschiktheid overduidelijk, bijvoorbeeld als de cliënt de Nederlandse taal onvoldoende machtig is om de opdrachten in een test te begrijpen, maar in andere gevallen kan de geschiktheid moeilijk te beoordelen zijn. De diagnosticus moet dan zo goed mogelijk onderbouwd de keuze maken (zie ook 1.2.6 'Toepassing psychodiagnostische instrumenten bij cliënten met specifieke individuele kenmerken').

BOX 6

Zelfbeschikkingsrecht van de cliënt versus professionele verantwoordelijkheid van de psycholoog

Ouders van cliënt C (8 jaar) willen graag een onderzoek naar de intelligentie van cliënt, omdat zij cliënt graag willen inschrijven op school Z. School Z heeft als toelatingscriterium dat de score op een passende IQ test een 90% betrouwbaarheidsinterval heeft waarin de score 130 of hoger valt. Psycholoog X heeft voorafgaand aan het onderzoek een gesprek met de ouders. De ouders geven tijdens dit gesprek aan dat zij graag meerdere opties van intelligentietests voorgelegd krijgen, zodat zij samen met de psycholoog een keuze kunnen maken met welk instrument hun kind wordt onderzocht om het intelligentieniveau te bepalen. Ze hebben namelijk op een internetforum voor ouders gelezen dat een specifieke intelligentietest, test Y, meer geschikt is voor afname bij kinderen die vooral verbaal sterk zijn. Ook was moeder in het verleden 2 jaar ingeschreven als student aan een bachelor opleiding in de orthopedagogiek.

Toelichting BOX 6

Als een uitwerking van het basisprincipe 'Respect' legt de Beroepscode de psycholoog de professionele verantwoordelijkheid op om de kennis, het inzicht en de ervaring van de betrokkenen, in dit geval de ouders, te respecteren (artikel 55 'Respect voor kennis, inzicht en ervaring'). De Beroepscode stelt dat uiteindelijk de psycholoog beslist, op basis van diens professionele verantwoordelijkheid, of het verantwoord is om een bepaald psychodiagnostisch instrument af te nemen (artikel 9 'Verantwoordelijkheid').

Gezien het toelatingscriterium van de school (IQ score hoger dan 130) is het belangrijk bij de keuze van de intelligentietest te zorgen voor een hoge testbetrouwbaarheid, voor kinderen van, in dit geval, 8 jaar met een hoge intelligentie. Deze hoge betrouwbaarheid komt tot uiting in een smal 95% betrouwbaarheidsinterval rondom de IQ score van 130. Het respect voor de autonomie en zelfbeschikking van de cliënt komt in deze casus tot uiting in het recht van de cliënt en diens ouders om wel of niet een professionele relatie met psycholoog X aan te gaan (Beroepscode, artikel 58 'Respect voor autonomie en zelfbeschikking'; artikel 60 'Toestemming bij het aangaan of voortzetten van de professionele relatie' en artikel 61 'Aangaan en voortzetten van de professionele relatie'), maar niet in de beslissing welke intelligentietest gebruikt gaat worden. Het staat de cliënt of (in dit geval zijn ouders namens hem) in deze casus vrij om, op grond van zorgvuldige informatieverstrekking door de psycholoog te besluiten om al dan niet verder te gaan met de betreffende psycholoog.

Overeenkomstig de Beroepscode moet de diagnosticus bij het aangaan en voortzetten van de professionele relatie de cliënt voorzien van informatie over de beschikbare methoden van onderzoek of behandeling, de verwachtingen en het soort gegevens dat verzameld wordt. Bij voorkeur wordt deze informatie schriftelijk gegeven en waar mogelijk mondeling toegelicht (Beroepscode, artikel 62 'Informatie bij het aangaan en voortzetten van de professionele relatie'). De Beroepscode (in het bijzonder artikel 58 'Respect voor autonomie en zelfbeschikking'; artikel 62 'Informatie bij het aangaan en voortzetten van de professionele relatie' en artikel 64 'Overleg over invulling van de professionele relatie') biedt de cliënt de mogelijkheid om diens voorkeuren en meningen mee te laten wegen bij de keuze voor en het gebruik van instrumenten voor de beoordeling van specifieke eigenschappen (bijvoorbeeld intelligentie of prestatiemotivatie) en de diepgang van het onderzoek. Zie de casus beschreven in **BOX 6**. Dit weerspiegelt het spanningsveld tussen het zelfbeschikkingsrecht van de cliënt en de professionele verantwoordelijkheid van de diagnosticus, een kwestie die onder andere ook speelt in de geneeskunde.

1.2.3.2 Principes bij de keuze van psychodiagnostische instrumenten

Als de diagnosticus overweegt om een psychodiagnostisch instrument in te zetten, is het van belang dat die zich informeert over de instrumenten die beschikbaar zijn voor het beoogde doel. In veel gevallen zijn meerdere instrumenten met een vergelijkbare meetpretentie beschikbaar. De kwaliteit van deze instrumenten is dan een belangrijke overweging bij het maken van een keuze. De informatie die een testuitgever in de handleiding verstrekt, kan hierbij behulpzaam zijn. Daarnaast kan de kwaliteitsbeoordeling van de [Commissie Testaangelegenheden Nederland \(COTAN\)](#) over een instrument (indien beschikbaar) een nuttig hulpmiddel zijn. Verder moet de diagnosticus voldoende gekwalificeerd zijn om het specifieke gekozen instrument te gebruiken (Beroepscode, artikel 106 'Kwalificatie').

De COTAN beoordeelt de kwaliteit van psychodiagnostische instrumenten aan de hand van een beoordelingssysteem (Evers et al., 2009), zoals toegelicht in **BOX 7**. Het COTAN beoordelingssysteem is genuanceerd: er is geen algemeen keurmerk van "goedgekeurd psychodiagnostisch instrument" of COTAN-gecertificeerde test.

De COTAN beoordeelt een instrument op zeven criteria (zie **BOX 7**) als 'onvoldoende', 'voldoende' of 'goed', met een bijbehorende toelichting. De toelichting bevat belangrijke beschrijvende informatie, samengevatte onderzoeksgegevens en de redenering van de beoordelaars voor het oordeel. Het benadrukt eventuele beperkingen bij het gebruik van het instrument. Het is daarom essentieel dat de diagnosticus niet alleen naar de beoordeling kijkt, maar ook de toelichting leest.

Bij de keuze van een instrument beoordeelt de diagnosticus of een instrument geschikt is voor het beoogde doel en voor deze cliënt. Vergewis je ervan dat de cliënt tot de doelgroep behoort en dat het instrument het beoogde construct goed kan meten gezien de individuele eigenschappen van de cliënt, zoals leeftijd, gender, taal of culturele

achtergrond van de cliënt. Zie ook 1.2.6, 'Toepassing psychodiagnostische instrumenten bij cliënten met specifieke individuele kenmerken'.

Als de diagnosticus van mening is dat het instrument geschikt is voor het beoogde doel en deze cliënt, is het nuttig om naast de handleiding van het instrument ook de COTAN beoordeling van het instrument (indien beschikbaar) te raadplegen. Bij de keuze voor een instrument wordt de voorkeur gegeven aan een instrument met een 'voldoende' of 'goede' COTAN-beoordeling op alle criteria. Echter, als de COTAN een 'onvoldoende' heeft gegeven voor een bepaald criterium, is het belangrijk om de toelichting op de reden voor deze 'onvoldoende' zorgvuldig te bestuderen. Het kan bijvoorbeeld zijn dat er onvoldoende onderzoek is gedaan naar het instrument voor alle specifieke doelen en doelgroepen zoals vermeld in de handleiding. Dit kan echter voor het specifieke beoogde doel bij deze cliënt niet relevant zijn. Het kan echter ook betekenen dat het instrument duidelijk achterblijft bij de stand van de wetenschap. In dat geval moet de diagnosticus, met inachtneming van artikel 15, 'Professionele standaard', van de Beroepscode overwegen of de diagnosticus het betreffende instrument kan en wil inzetten. Als er een alternatief, goed instrument beschikbaar is met een vergelijkbare meetpretentie, maakt de diagnosticus gebruik van dit instrument.

De verantwoordelijkheid voor het selecteren van instrumenten van de hoogst mogelijke kwaliteit die passen bij de vraagstelling, blijft bij de diagnosticus. Artikel 102, 'Gebruik van doeltreffende en doelmatige methoden', van de Beroepscode benadrukt de professionele verantwoordelijkheid van de diagnosticus om methoden te kiezen waarmee het beoogde doel zo effectief en efficiënt mogelijk wordt bereikt; dit geldt ook voor de keuze van testinstrumenten. Bij een 'onvoldoende' of afwijken van de professionele standaard dient de diagnosticus deze keuze of afweging zorgvuldig te onderbouwen.

BOX 7

Beoordelingscriteria van de COTAN

1. **Uitgangspunten van de testconstructie:** worden het te meten construct en de theoretische achtergrond ervan, het gebruiksdoel van het instrument en de doelgroep duidelijk beschreven? Ook de operationalisatie moet worden verantwoord: hoe zijn de items tot stand gekomen en is aannemelijk dat zij voortvloeien uit de definitie van het construct? Als bij een instrument bijvoorbeeld het theoretisch model dat eraan ten grondslag ligt niet wordt beschreven of als uitleg over het proces van operationalisatie ontbreekt, kan een onvoldoende op dit criterium worden gegeven.
2. **Kwaliteit van het testmateriaal:** zijn de instructie, de items en de scoring gestandaardiseerd en worden er voldoende aanwijzingen gegeven voor de geteste? Een onvoldoende op dit criterium kan gegeven worden als bijvoorbeeld de items multi-interpretabel of onnodig ingewikkeld geformuleerd zijn of als er geen duidelijke aanwijzingen worden gegeven over hoe de antwoorden moeten worden gescoord (wat bij een mondeling en individueel af te nemen instrument het meest speelt). Bij een instrument dat via de computer wordt afgenomen worden aanvullende kwaliteitseisen gesteld, onder andere aan de schermkwaliteit.
3. **Kwaliteit van de handleiding:** wordt er informatie gegeven ter ondersteuning van de testgebruiker bij afname en interpretatie van het instrument? Als bijvoorbeeld een verzameling wetenschappelijke artikelen wordt aangeleverd als handleiding, dan krijgt het instrument een onvoldoende beoordeling

op dit criterium. De handleiding moet namelijk het gebruik van het instrument in de praktijk op een toegankelijke wijze ondersteunen, onder meer door het bespreken van één of meer gevalsoverbeschrijvingen en het wijzen op andere soorten informatie die bij de interpretatie van belang kunnen zijn.

4. **Normen:** wat is de kwaliteit van de normen en de informatie die daarover wordt verschaft? Zijn de normgroepen van voldoende omvang en vooral: zijn ze representatief voor de beoogde doelgroep en gelet op het doel van het instrument? Als van de steekproef die is gebruikt om de normgegevens te verzamelen geen duidelijke beschrijving wordt gegeven of als de omvang of de representativiteit tekortschiet, zal dit tot een onvoldoende beoordeling op dit criterium leiden.
5. **Betrouwbaarheid:** in welke mate heeft de testscore geen last van willekeurige meetfouten? De meetnauwkeurigheid is uiteraard altijd van belang en bij veel doelen (bijvoorbeeld toelating tot speciaal onderwijs) is het voor de belanghebbenden extra belangrijk zo nauwkeurig mogelijke uitkomsten te verkrijgen. Voor het beoordelen van de betrouwbaarheid kunnen gegevens over verschillende betrouwbaarheidscoëfficiënten worden gerapporteerd. De data waarop deze zijn berekend moeten overeenkomen met de doelgroep en het doel van het instrument. Als er geen beschrijvende informatie wordt gegeven over het betrouwbaarheidsonderzoek of als de gerapporteerde waarden onder de maat zijn, dan wordt een onvoldoende gegeven op dit criterium.

6. **Begripsvaliditeit:** meet het psychodiagnostisch instrument inderdaad de eigenschap die wordt verondersteld? Ook hier geldt dat de data overeen moeten komen met de doelgroep en het doel van het instrument. Veelgebruikte methoden voor het aantonen van de begripsvaliditeit zijn factoranalyse, het vergelijken van gemiddelde scores van groepen waarvan men mag verwachten dat ze verschillen zullen vertonen op de betreffende eigenschap en het berekenen van correlaties met instrumenten die hetzelfde begrip beogen te meten (soortgenoten). Deze onderzoeken zijn elk op zichzelf nog geen aanleiding tot een voldoende beoordeling op dit criterium. Meerdere onderzoeken naar zowel de interne structuur als de verbanden met externe variabelen kunnen leiden tot de beoordeling 'voldoende' of 'goed'.

7. **Criteriumvaliditeit:** in hoeverre hangt de testscore samen met relevant niet-testgedrag, met gedrag in de praktijk? Hierbij kan men, bijvoorbeeld bij een psychodiagnostisch instrument dat wordt gebruikt voor personeelsselectie, denken aan de voorspellende waarde van dat instrument voor het latere functioneren van de werknemers. Een onvoldoende beoordeling op dit criterium kan worden gegeven als het valideringsonderzoek onvoldoende bewijs levert voor de samenhang met de uitkomstmaat. Ook kan een onvoldoende worden gegeven als het onderzoek met heel kleine steekproeven is uitgevoerd of als de uitkomstmaten weinig relevant of van geringe kwaliteit zijn.

Het COTAN Beoordelingssysteem is uitgebreid met drie addenda met betrekking tot unproctored (zonder toezicht) gegevensverzameling, actualisatie instrumenten en fairness.

We geven drie voorbeelden die duidelijk maken waarom het belangrijk is om kennis te nemen van de toelichting op de COTAN-beoordeling van een instrument. Het eerste voorbeeld betreft instrumenten waarvan een digitale versie bestaat die ontwikkeld is op basis van een papieren versie. Deze twee versies verschillen vaak in instructies, voorbeelden en de presentatie van de testitems. Gelijkheid van beide versies in termen van meeteigenschappen is niet gegarandeerd als de testauteurs of -uitgevers geen aanvullend onderzoek hebben gedaan om te beoordelen of de eigenschappen van de papieren en de digitale versies equivalent zijn. In de toelichting bij de COTAN-beoordeling wordt hier vaak op ingegaan, wat de diagnosticus extra informatie biedt om een weloverwogen keuze te maken bij de selectie van een instrument.

Het tweede voorbeeld betreft het gebruik van voetnoten bij het samenvattende oordeel per criterium. Voor bepaalde criteria, zoals *normen*, *betrouwbaarheid* en *begripsvaliditeit*, wordt een voetnoot toegevoegd wanneer de beoordeling voor verschillende subgroepen of subschalen verschilt. Bijvoorbeeld, voor de betrouwbaarheid kunnen de gevonden coëfficiënten voor verschillende leeftijdsgroepen aanzienlijk variëren, resulterend in een beoordeling over leeftijdsgroepen variërend van 'onvoldoende' tot 'goed'. De toelichting biedt dan een volledige en genuanceerde uitleg van deze variatie, waardoor het lezen van de toelichting essentieel is bij het maken van een verantwoorde keuze voor een instrument.

Het derde voorbeeld betreft het gebruik van herhaalde metingen bij een cliënt, uitgevoerd om de behandeling te volgen en evalueren. Dit speelt bijvoorbeeld bij toepassing van Routine Outcome Monitoring (ROM). De psychometrische eigenschappen van een instrument voor herhaalde metingen kunnen verschillen van de eigenschappen van datzelfde instrument wanneer het eenmalig wordt afgenomen. Bij toepassing van een instrument voor herhaalde metingen moet dus bekeken worden of de COTAN-beoordeling (ook) betrekking heeft op herhaalde metingen.

1.2.4 Afname van psychodiagnostische instrumenten

Psychodiagnostiek, en met name het gebruik van tests, is een professionele toepassing van de psychologie die aanzienlijke invloed kan hebben op het leven van personen. Daarom is het essentieel dat diagnostici voldoende theoretische en praktische scholing hebben gehad om psychodiagnostiek verantwoord uit te oefenen. Een diagnosticus moet actief ontwikkelingen in de psychodiagnostiek en psychometrie volgen en zich zo nodig bijscholen, zoals vermeld in artikel 101 'In stand houden en ontwikkelen professionele deskundigheid' van de Beroepscode. Het voert te ver om in de AST-NIP een uitputtende lijst van alle mogelijke kwalificaties op het gebied van de psychodiagnostiek in Nederland te geven. Wel worden twee voorbeelden genoemd van regelingen die tot doel hebben om goed testgebruik te stimuleren.

Het eerste voorbeeld is een richtlijn opgesteld door de European Federation of Psychologists' Associations (EFPA), bekend als The EFPA Standards for Test Use (EFPA, 2012), waarin algemene competenties en kwalificaties voor testgebruik worden beschreven. De richtlijn classificeert testgebruikers in drie categorieën, met toenemend kennis- en vaardigheidsniveau: de assistent-testgebruiker, de testgebruiker en de specialist in tests en testgebruik. Ter illustratie is in **BOX 8** een Nederlandse samenvatting opgenomen van de beschrijving van deze competentieniveaus in de EFPA Standards for Test Use. Veel uitgevers van tests gebruiken kwalificatie-eisen om te bepalen wie bevoegd is om een specifiek instrument aan te schaffen, soms met de aanvullende eis van het volgen van een training voor het gebruik van het instrument.

BOX 8 Drie competentieniveaus voor testgebruik in de EFPA Standards for Test Use

Assistent Testgebruiker – niveau 1 (EFPA Level 1)

Iemand die specifieke psychodiagnostische instrumenten gebruikt in een afgebakende context, onder toezicht van een meer ervaren testgebruiker, bijvoorbeeld een persoon opgeleid als psychologisch testassistent. Ze werken binnen de richtlijnen van een organisatie voor psychodiagnostische instrumenten en testgebruik. Ze hebben geen zeggenschap over de keuze van de instrumenten en de wijze waarop ze worden gebruikt. Samengevat, de persoon:

- kan specifieke instrumenten gebruiken in beperkte omstandigheden onder toezicht van een hoger gekwalificeerde persoon;
- heeft geen bevoegdheid om instrumentkeuzes te maken of testcores verder te interpreteren dan de standaardrapportages;
- heeft bewustzijn van bredere kwesties met betrekking tot psychodiagnostische instrumenten en testgebruik, en wanneer deskundige hulp nodig is.

Testgebruiker - niveau 2 (EFPA Level 2)

Iemand die specifieke psychodiagnostische instrumenten in een afgebakende context gebruikt. Voorbeelden zijn medewerkers op HR-afdelingen, uitzendbureaus, of adviesbureaus met testdiensten. Ze kunnen betrokken zijn bij het testen van kandidaten voor personeelsselectie, personeelsontwikkeling of loopbaanbegeleiding en advies, etc. Samengevat, de persoon:

- begrijpt de technische/psychometrische eigenschappen van psychodiagnostische instrumenten en kan ze zelfstandig gebruiken in afgebakende contexten;
- kan scores interpreteren op specifieke psychodiagnostische instrumenten;
- heeft geen bevoegdheid om keuzes te maken over het gebruik van psychodiagnostische instrumenten buiten hun expertise.

Specialist in Tests en Testgebruik – niveau 3 (EFPA Level 3)

Dit zijn doorgaans ervaren psychologen, die zich hebben gespecialiseerd in specifieke toepassinggebieden van testgebruik. Ze kunnen diensten aanbieden als advies en training in testgebruik, testconstructie en getuige-deskundige met betrekking tot testgebruik in rechtszaken. Hoewel niveau 3 geen expertise in testconstructie vereist, vallen personen met deze expertise ook onder dit niveau.

N.B. De voorbeelden zijn bedoeld ter illustratie van de rollen op elk niveau en zijn niet bedoeld als een uitputtende lijst.

Het tweede voorbeeld is een keurmerk dat aangeeft dat de professional beschikt over basiskennis en -vaardigheden op het gebied van de psychodiagnostiek. [De Basisaantekening Psychodiagnostiek \(BAPD\)](#) is een kwaliteitskeurmerk van het NIP. Om de BAPD te behalen, moeten psychologen voldoen aan verschillende criteria, waaronder minimaal 200 uur werkervaring in de psychodiagnostiek onder supervisie, het schrijven van drie casusverslagen volgens het BAPD-format en het voldoen aan specifieke theoretische eisen met betrekking tot psychodiagnostiek, psychometrie en besliskunde, psychodiagnostische instrumenten en werkwijzen, praktijkvaardigheid met betrekking tot psychodiagnostische instrumenten, gespreksvoering, observatie en beslissen en communicatievaardigheden.

1.2.4.1 Verantwoordelijkheid voor afname

Als bij een psychodiagnostisch onderzoek gebruik gemaakt wordt van een psychodiagnostisch instrument, dan is de diagnosticus er verantwoordelijk voor dat het instrument correct wordt afgenomen, gescoord en geïnterpreteerd. Als anderen in opdracht en onder directe begeleiding van de diagnosticus specifieke taken uitvoeren, blijft de diagnosticus verantwoordelijk voor de kwaliteit van deze werkzaamheden (Beroepscode, artikel 30 'Verantwoordelijkheid voor de kwaliteit van medewerkers'). Dit betekent bijvoorbeeld dat de diagnosticus de kwaliteit van het werk van eventuele testassistenten waarborgt. Dit vereist een duidelijke instructie en het faciliteren van een correcte testafname, zoals nader beschreven in **BOX 9**. De taakverdeling over de verschillende professionals moet duidelijk worden vastgelegd in het dossier en moet in ieder geval aangeven wie zorg draagt voor:

- Het formuleren van de onderzoeksvraag en het volgen van de onderzoeksprocedure;
- Het uitvoeren van de verschillende testafnames;
- Het scoren van de tests en het gebruik van scoringsprogramma's;
- Het interpreteren van de resultaten;
- Het opstellen van rapporten voor de opdrachtgever en het uitvoeren van nabespreking(en).

Bij de testafname moet de afnameprocedure zoals beschreven in de testhandleiding worden gevolgd. Dit omvat onder andere de voorgeschreven methoden en omstandigheden voor de afname, instructies aan de geteste personen, de volgorde van afname, de volledigheid van de afname, het eventuele gebruik van een computer, de mogelijkheid tot tussentijdse gesprekken en praktijkoefeningen en andere specifieke eisen die in de handleiding worden vermeld. Hieronder worden enkele van deze aspecten verder toegelicht.

1.2.4.2 Toezicht bij afname

Bij veel tests is toezicht tijdens de afname (proctored testing) volgens de betreffende testhandleiding essentieel. De voordelen hiervan zijn dat een professional beschikbaar is om vragen van de cliënt te beantwoorden en dat de gegevens van de cliënt worden verzameld onder omstandigheden die vergelijkbaar zijn met die waarin het instrument genormeerd is. Dit draagt bij aan de betrouwbaarheid en validiteit van het testresultaat.

Wanneer er geen toezicht aanwezig is (unproctored testing), bijvoorbeeld wanneer de cliënt zelfstandig online of op papier een vragenlijst invult, is er het evidente voordeel dat het instrument op afstand kan worden afgenomen. De nadelen zijn het ontbreken van de zekerheid over de identiteit van de geteste persoon, de mogelijkheid om te controleren of de persoon de testitems onafhankelijk heeft beantwoord en het vermogen van de persoon om tijdens de afname om verduidelijking te vragen. Dit maakt unproctored afname gevoeliger voor fouten en mogelijk misbruik. Online unproctored afname kan worden beveiligd met behulp van wachtwoorden, gebruikersnamen en proctoringmethoden die gebruik maken van webcams en screencapture, waarmee toezicht wordt gehouden op de persoon, de omgeving en het scherm van de persoon die de test maakt. Echter, de diagnosticus moet zich er bewust van zijn dat dit geen 100% zekerheid biedt dat de beoogde geteste persoon ook degene is die de test gemaakt heeft. Zie voor meer informatie het [addendum](#) op het COTAN Beoordelingssysteem met betrekking tot unproctored afname.

BOX 9

Verantwoordelijkheid voor afnameprocedure

Psycholoog X onderzoekt een meisje van 10 jaar en wil haar intelligentie meten met een intelligentietest. De ouders vragen of de test thuis kan worden afgenomen. Psycholoog X laat de afname over aan zijn stagiair, een vierdejaars student met weinig ervaring met afname van deze test. De stagiair neemt de test thuis af in de woonkamer, terwijl de moeder soms lekkers brengt en de oudere broer van het meisje aanwezig is in de woonkamer en het meisje soms helpt bij moeilijke vragen. Na afloop rapporteert de stagiair de scores aan Psycholoog X, die de resultaten interpreteert en concludeert dat het meisje gemiddeld intelligent is.

Toelichting BOX 9:

Afname van een psychodiagnostisch instrument bij de cliënt thuis is alleen toegestaan als de ruimte en omstandigheden voldoen aan de voorwaarden in de testhandleiding. Dit betekent dat de ruimte geschikt moet zijn voor het onderzoek, storende invloeden minimaal moeten zijn en de testomstandigheden vergelijkbaar moeten zijn met die van de normgroep. De diagnosticus moet, in overleg met de cliënt, zorgen dat aan deze voorwaarden wordt voldaan en draagt de verantwoordelijkheid om te beoordelen of thuisafname verantwoord is.

Daarnaast moet de diagnosticus erop toezien dat de cliënt zelfstandig reageert op de items, zonder hulp van anderen. Dit laatste is ook een aandachtspunt bij unproctored afname (zonder toezicht) van digitaal (online) testmateriaal. De beroepscode legt de diagnosticus de professionele verantwoordelijkheid

op om methoden te kiezen die effectief en doelmatig zijn, rekening houdend met de beperkingen van die methoden (artikel 102 'Gebruik van doeltreffende en doelmatige methoden'). Als de voorwaarden niet worden nageleefd, kunnen testcores vertekend zijn en is de betrouwbaarheid en de validiteit in het geding.

Deze casus benadrukt het belang van duidelijke instructies voor testassistenten en een goede nabespreking om onregelmatigheden te signaleren. De diagnosticus blijft verantwoordelijk voor de kwaliteit van het werk van medewerkers onder zijn toezicht, zoals de stagiair (artikel 30 'Verantwoordelijkheid voor de kwaliteit van medewerkers'), en moet hen ondersteunen om het werk op een ethisch verantwoorde manier uit te voeren (artikel 31 'Hulp en steun aan collega's, studenten en supervisanten').

De diagnosticus moet de wijze van afname van een instrument kunnen verantwoorden, in overeenstemming met vooral artikel 107 'De wetenschappelijke basis van het beroepsmatig handelen' van de Beroepscode. Wijzigingen in de oorspronkelijke afnameprocedure moeten worden gedocumenteerd en bij voorkeur worden onderbouwd met literatuur of onderzoeksgegevens, zie ook de casus beschreven in **BOX 9**.

1.2.4.3 Testruimte bij afname

De testruimte moet geschikt zijn voor het specifieke psychodiagnostisch onderzoek dat wordt uitgevoerd. Dit betekent dat de ruimte adequaat verlicht, geluidsarm en schoon moet zijn. Zowel de stoel als de tafel moeten comfortabel zijn voor de cliënt, zodat deze zonder problemen gedurende langere tijd aan de test(s) kan werken. Cliënten moet een omgeving worden geboden waarin zij niet worden waargenomen door derden tijdens de professionele activiteiten, tenzij de cliënt ermee instemt dat derden die activiteiten kunnen waarnemen (Beroepscode, artikel 71 'Ruimtelijke privacy'). Storende invloeden tijdens de testafname moeten tot een minimum worden beperkt om afwijkingen van de standardsituatie, waarin het psychodiagnostisch instrument is genormeerd, te voorkomen. Afwijkingen in de testruimte, zoals te weinig ruimte, afleidende elementen of een benauwde omgeving, kunnen de testprestatie negatief beïnvloeden en resulteren in onbruikbare testcores volgens de normen van het psychodiagnostisch instrument, zie ook de casus beschreven in **BOX 9**.

1.2.4.4 Kwaliteit testmateriaal bij afname

Bij de afname van een psychodiagnostisch instrument is het belangrijk dat het instrument in de originele of bedoelde staat wordt gebruikt. Als het instrument in een slechtere staat wordt gebruikt dan bedoeld, zoals bijvoorbeeld met slecht leesbare kopieën van een vragenlijst, kan dit de testresultaten negatief beïnvloeden. De test kan bijvoorbeeld dan ook meten hoe goed een cliënt om kan gaan met de verminderde leesbaarheid van de vragen, wat uiteraard niet de bedoeling is. Bij afname van een instrument via een computer, zoals bij prestatietaken, moet de scherm-

helderheid, scherpte en schermgrootte in orde zijn. De testomstandigheden moeten zoveel mogelijk overeenkomen met de omstandigheden waarin het instrument genormeerd is, om te voorkomen dat verschillen in testprestaties worden veroorzaakt door afwijkingen in het gebruikte materiaal. Door gestandaardiseerde testafname wordt de nauwkeurigheid van de testresultaten gewaarborgd. Zie ook de casus beschreven in **BOX 10**.

1.2.4.5 Beveiliging van tests

De beveiliging van psychodiagnostische instrumenten omvat twee aspecten. Het eerste aspect betreft de beveiliging van testgegevens van een individuele cliënt. De diagnosticus moet waarborgen dat alleen rechtmatig betrokkenen toegang hebben tot deze gegevens (Beroepscode, artikel 80 'Beveiliging van het dossier') en dat misbruik wordt voorkomen, zoals diefstal of het ongeoorloofd aanbrengen van wijzigingen in de resultaten. Verder moeten testgegevens vertrouwelijk behandeld worden vanwege de geheimhoudingsplicht (Beroepscode, artikel 70 'Geheimhouding' e.v.).

Het tweede aspect betreft de beveiliging van testmateriaal. Het risico van het bekend worden van de testinhoud onder (potentiële) cliënten moet geminimaliseerd worden om de validiteit en betrouwbaarheid van tests te waarborgen. Dit betekent dat alleen geautoriseerd personeel toegang heeft tot het testmateriaal. Cliënten moeten geen onnodige toegang hebben tot items en moeten geen toegang hebben tot scoringsvoorschriften.

1.2.5 Criterium-georiënteerde test scores, norm-georiënteerde test scores en testnormen

Een test bestaat uit verschillende items, zoals testvragen, waarbij de antwoorden of, voor observatie-instrumenten, reacties van een cliënt de basis vormen voor de ruwe score per item. Ruwe scores kunnen variëren, zoals goed/fout (gescoord als 1/0), een score op een 5-punts Likertschaal (gescoord als 1, 2, 3, 4, 5), of een reactietijd (in seconden). Tests bevatten vaak meerdere subtests, zoals 'logisch redeneren', 'verbaal redeneren' en 'ruimtelijk redeneren'. De ruwe testscore, vaak berekend als de ongewogen som

BOX 10 Kwaliteit testmateriaal

Schoolpsycholoog X heeft een leerling met aandachtsproblemen onder zijn hoede. De leerling heeft slecht gepresteerd op de toets aan het einde van groep 7. De vraag is of de leerling naar regulier vervolgonderwijs kan doorstromen of dat hij in groep 8 een aangepast traject gaat volgen. Ouders van de leerling twijfelen aan de geldigheid van het toetsresultaat; zij eisen dat de leerling de toets opnieuw mag maken. Als argument voeren zij aan dat gewerkt is met hergebruikte toetsboekjes, waardoor hun zoon in de war raakte. Bij navraag door psycholoog X bleek dat de toetsboekjes inderdaad al in een andere klas gebruikt waren. De meerkeuzevragen waren daardoor al door andere leerlingen ingevuld. De klassendocent had dit "opgelost" door aan de kinderen de instructie te geven om niet te letten op wat er al ingevuld was en hun eigen antwoorden op een apart antwoordformulier in te vullen.

Toelichting BOX 10:

In de instructie van leerlingvolgtoetsen staat vermeld dat leerkrachten ervoor moeten zorgdragen dat leerlingen de toets zoveel mogelijk onder dezelfde omstandigheden maken, omdat alleen dan hun toetsresultaten vergelijkbaar zijn met die van andere kinderen. In het geval van de casus is het testmateriaal van ondeugdelijke kwaliteit, waardoor de betrouwbaarheid en validiteit van de scores van de leerling in het geding zijn. Doordat er al antwoorden ingevuld zijn, meet de toets niet meer alleen de beoogde kennis en/of vaardigheden, maar ook in hoeverre een leerling in

staat is de storende informatie te negeren. Het lijkt aannemelijk dat het voor een leerling met aandachtsproblematiek of een onzeker kind des te moeilijker is om zich af te sluiten voor storende invloeden. Schoolpsychologen dienen toezicht te houden op het gebruik van deugdelijke instrumenten en al het mogelijke te doen om een juiste testafname te bevorderen wanneer deze uitgevoerd wordt door docenten (Beroepscode, artikel 29 'Medeverantwoordelijkheid voor de kwaliteit van het team'). Het is aan te bevelen om afnameprocedures met docenten te bespreken en indien mogelijk te oefenen, zie ook artikel 31 'Hulp en steun aan collega's, studenten en superviseren' van de Beroepscode.

Vergelijkbare situaties zijn het afnemen van gekopieerde psychologische vragenlijsten (met name als door het kopiëren de leesbaarheid is verslechterd), of het onzorgvuldig omzetten van een papier-en-potlood test naar een digitale variant. In het geval dat deze psychodiagnostische instrumenten door of onder verantwoordelijkheid van de diagnosticus worden afgenomen moet deze zich kunnen verantwoorden voor de keuze en het gebruik van het testmateriaal (Beroepscode, artikel 102 'Gebruik van doeltreffende en doelmatige methoden'). Vanzelfsprekend dient de diagnosticus niet te werken met ongeautoriseerde kopieën van tests of vragenlijsten (Beroepscode, artikel 100 'Kennis van wettelijke bepalingen').

van de ruwe scores op de bijbehorende items, biedt op zichzelf vaak onvoldoende informatie over het te meten begrip.

Om deze reden wordt de ruwe testscore omgezet in een interpreteerbare score, afhankelijk van het type test. Bij criterium-georiënteerde test scores wordt de ruwe test score vergeleken met een vooraf vastgestelde standaard, bijvoorbeeld cut-off waarden die leiden tot een classificatie zoals 'laag risico', 'medium risico' of 'hoog risico'.

Psychometrisch onderzoek richt zich op de kwaliteit van de classificatie, in termen van begrippen als de sensitiviteit, specificiteit, positieve predictieve waarde (PPW) en negatieve predictieve waarde (NPW), in een relevante steekproef (of steekproeven) uit populatie(s) die overeenkomen met de doelgroep van de test. Deze begrippen zijn onderling afhankelijk van elkaar, waarbij PPW en NPW sterk worden beïnvloed door de prevalentie van een conditie/aandoening in een populatie. Merk verder op dat een screeningsinstrument vaak een hoge sensitiviteit en een lage specificiteit heeft. Dit betekent dat als dit instrument ingezet wordt voor een diagnostisch doel, de kans op een misclassificatie van personen zonder de conditie/aandoening (zogenoemde controls) vrij groot is.

Bij een norm-georiënteerde testscore wordt de ruwe test score omgezet in een gestandaardiseerde score op basis van de normen van de test. De gestandaardiseerde score drukt een vergelijking uit met prestaties in een relevante referentiep populatie, zoals bijvoorbeeld de algemene bevolking van dezelfde leeftijd. Een gestandaardiseerde score drukt daarmee uit hoe de cliënt presteert relatief ten opzichte van de referentiep populatie. Voorbeelden van gestandaardiseerde scores zijn T-scores (gemiddelde = 10, standaarddeviatie (SD) = 3), genormaliseerde IQ-scores (normaal verdeeld, met gemiddelde = 100, SD = 15), en stanines (gemiddelde = 5, SD = 2). De normen van een test worden berekend door de testontwikkelaar op basis van scores in een (grote) normsteekproef, die getrokken wordt uit alle referentiep populaties van de betreffende test (zoals bijvoorbeeld de algemene bevolking in de leeftijd 20-80 jaar).

De keuze van de referentiep populatie hangt af van de meetpretentie en het gebruiksdoel van de test. Een veel gebruikte referentiep populatie, bijvoorbeeld in intelligentietests, is de algemene bevolking van dezelfde leeftijd. Een gestandaardiseerde intelligentiescore drukt daarmee uit hoe goed de cliënt presteert ten opzichte van de algemene bevolking met dezelfde leeftijd. Dit is een zinvolle referentiep populatie in het kader van, bijvoorbeeld, selectie voor een baan of opleiding. Voor klinische tests, zoals die toegepast worden in de klinische (neuro)psychologie, is de referentiep populatie vaak de gezonde bevolking van dezelfde leeftijd, en soms mensen met een specifieke aandoening van dezelfde leeftijd.

Bij klinische tests, vaak cognitieve tests, wordt bij de omzetting van de ruwe testscore naar een gestandaardiseerde score rekening gehouden met specifieke kenmerken van de referentiep populatie, zoals leeftijd, gender en opleidingsniveau van de gezonde bevolking. In dat geval drukt de gestandaardiseerde score uit hoe uitzonderlijk de behaalde score is in de gezonde bevolking van dezelfde leeftijd, gender en opleidingsniveau. Dit is een goede referentiep populatie gezien het doel van klinische tests, namelijk het beoordelen van de mate van disfunctie. Een score die negatief afwijkt ten opzichte van gezonde personen met verder dezelfde kenmerken als de betrokken cliënt is een aanwijzing voor relatieve disfunctie in het geteste domein.

Bij sommige tests worden de normen gegeven voor meerdere referentiep populaties. Ten eerste zijn er meerdere referentiep populaties als er normen worden berekend op basis van één normeringssteekproef, waarbij de normen verschillen afhankelijk van bepaalde kenmerken – dus bijvoorbeeld leeftijdsafhankelijke normen, en leeftijds- en geslachtsafhankelijke normen. Ten tweede zijn er meerdere referentiep populaties als er meerdere normeringssteekproeven zijn, bijvoorbeeld één uit de algemene bevolking en één uit een klinische populatie (personen met een bepaalde aandoening). In dat geval moet de diagnosticus kiezen welke referentiep populatie(s) het meest relevant is (zijn) voor de onderzoeksvraag en deze keuze onderbouwen.

De testnormen geven voor alle mogelijke ruwe test scores de bijbehorende gestandaardiseerde score weer. Als de referentiep populatie afhangt van bijvoorbeeld de leeftijd kan dit resulteren in een aanzienlijk aantal testnormen, wanneer voor iedere leeftijd binnen het leeftijdsbereik van de testnormen worden vastgesteld. Om praktisch hanteerbare testnormen te verkrijgen, zijn er twee benaderingen. Ten eerste kunnen de normen voor leeftijdsintervallen worden verstrekt. De breedte van die intervallen moet zorgvuldig worden gekozen om vergelijkbare testnormen binnen elke leeftijdsgroep te waarborgen. Ten tweede kunnen testnormen geïntegreerd zijn in computersoftware, waardoor een gestandaardiseerde score kan worden verkregen met als referentiep populatie personen van exact dezelfde leeftijd als de geteste cliënt.

Testnormen worden berekend op basis van de testprestaties van een steekproef van personen. Het is de verantwoordelijkheid van de testontwikkelaar om ervoor te zorgen dat deze normsteekproef enerzijds een representatieve afspiegeling vormt van alle relevante referentiep populaties van de betreffende test, en anderzijds groot genoeg is om de testnormen met voldoende nauwkeurigheid te berekenen. Alle relevante referentiep populaties samen worden de normpopulatie genoemd.

Het is belangrijk om de normpopulatie te onderscheiden van de doelgroep van de test. De doelgroep van een test omvat de personen voor wie de test geschikt is om het beoogde begrip te meten. De testontwikkelaar moet de doelgroep zo nauwkeurig mogelijk definiëren, bijvoorbeeld door middel van inclusie- en exclusiecriteria. De doelgroep van een test kan breder zijn dan de normpopulatie van de test. Neem bijvoorbeeld de doelgroep van een neuropsychologische test, die bestaat uit personen met een vermoeden van een bepaald type breinaandoening, in de leeftijd van 18 tot 80, met minimaal een MBO niveau 2. De referentiep populaties omvatten op hun beurt gezonde personen in dezelfde leeftijdscategorie, met minimaal een MBO niveau 2. Bij de selectie van een test voor een cliënt is het cruciaal dat de diagnosticus nagaat of de cliënt behoort tot de spe-

cifieke doelgroep van de test. Als dit niet zo is, kan de test mogelijk minder geschikt zijn voor de betreffende cliënt. Meer informatie hierover staat in sectie 1.2.6 ('Toepassing psychodiagnostische instrumenten bij cliënten met specifieke individuele kenmerken').

Bij de interpretatie van een gestandaardiseerde score is het essentieel dat de diagnosticus weet wat de referentiep populatie is. Deze informatie moet expliciet worden vermeld in het rapport. Het kan voorkomen dat een test norm-georiënteerd is, terwijl de diagnosticus de test zou willen inzetten voor een doel waarbij een criterium-georiënteerde test meer geschikt zou zijn, bijvoorbeeld bij diagnosestelling. In dergelijke gevallen heeft een passende criterium-georiënteerde test de voorkeur. Als deze niet beschikbaar is, dan moet de diagnosticus zich goed realiseren dat een relatief weinig voorkomende score binnen de algemene (of gezonde) bevolking niet per se betekent dat de onderzochte persoon lijdt aan de (vermoedelijke) aandoening. Met andere woorden, relatieve normen (van norm-georiënteerde test scores) kunnen niet zonder meer absoluut geïnterpreteerd worden. Een diagnose moet daarom gebaseerd zijn op meer informatiebronnen dan deze test alleen.

1.2.6 Toepassing psychodiagnostische instrumenten bij cliënten met specifieke individuele kenmerken

Het gebruik van een psychodiagnostisch instrument bij een cliënt met een specifiek individueel kenmerk, zoals bijvoorbeeld een fysieke beperking, verstandelijke beperking of een niet-Nederlandse culturele achtergrond, vereist speciale aandacht in de testprocedure. Dit individuele kenmerk kan namelijk onbedoelde invloed hebben op de testresultaten en daarmee op de interpretatie hiervan. Het is daarbij niet altijd helder welke kant de vertekening op zal gaan, bijvoorbeeld of de resultaten zullen leiden tot een overschatting of onderschatting van de capaciteiten of vaardigheden van de cliënt. Zie voor meer informatie de onderwerpen in **BOX 11** en **BOX 12**.

Het is de verantwoordelijkheid van de diagnosticus om rekening te houden met relevante individuele kenmerken van een cliënt. Dit volgt uit artikel 57 'Respect voor eigenheid en diversiteit' van de Beroepscode. Dit betekent dat een diagnosticus bij elke cliënt moet nagaan in hoeverre het testresultaat onbedoeld beïnvloed kan worden door de individuele kenmerken van de cliënt. Bekende factoren die zulke onbedoelde invloeden kunnen hebben, zijn fysieke beperkingen, mentale beperkingen, culturele achtergrond, taalvaardigheid en handvoorkeur.

Het categoriseren van deze kenmerken kan complex zijn. Dit geldt met name voor een kenmerk als etniciteit en migratieachtergrond, omdat de bijbehorende categorisatie gebaseerd is op sociale normen en kan variëren over gebieden en over de tijd. Daarbij geldt dat de individuen binnen deze categorieën onderling nog veel van elkaar verschillen. Psychologen lopen het risico de culturele en/of taalverschillen te overschatten of te onderschatten op basis van een paspoort, migratieachtergrond of zelfidentificatie. Als er sprake is van één of meer van dergelijke factoren, dan moet de diagnosticus extra aandacht besteden aan de volgende aspecten.

Ten eerste, bij de keuze van een psychodiagnostisch instrument moet de diagnosticus deze specifieke individuele

kenmerken meenemen in de overwegingen. Hierbij staat altijd het doel van de testafname centraal. Als bijvoorbeeld het doel is om de intelligentie van een cliënt met een niet-westerse culturele achtergrond te bepalen, dan kan het zinnig zijn om een non-verbale intelligentietest te gebruiken om de invloed van culturele achtergrond en taalbeheersing op de testscore te minimaliseren. Merk op dat ook bij non-verbale tests culturele verschillen van invloed zijn. Als het doel is om de taalvaardigheid te bepalen, dan is het gebruik van een non-verbaal instrument juist weer niet geschikt. Bij de afname van de test kan ook gedacht worden aan de inzet van een tolk. Dit heeft echter diverse nadelen, zoals de invloed op de relatie tussen cliënt en psycholoog en het risico dat de testinstructies niet zeer precies worden vertaald en dat daarmee potentieel een afwijking van de standaardinstructie optreedt.

Ten tweede moet de diagnosticus voor cliënten met een beperking bepalen of er aanpassingen nodig zijn in de test-situatie. Hierbij is het nodig dat de diagnosticus zekerheid heeft over de gemelde beperkingen, bijvoorbeeld door zelf de beperking vast te stellen of door een officiële verklaring (zoals een dyslexieverklaring) te verkrijgen. Deze verklaring geeft ook aan welke behandelingen, materiële voorzieningen, begeleiding en compensaties/dispensaties nodig zijn.

Voor cliënten met fysieke beperkingen moet de testsituatie goed toegankelijk zijn en waar nodig moeten ergonomische aanpassingen worden gedaan. Bijvoorbeeld het gebruik van voorleessoftware voor personen met visuele beperkingen of dyslexie, of het verhogen van tijdslimieten voor personen met dyslexie of motorische problemen. Een aanpassing is geoorloofd als de test na aanpassing nog steeds het beoogde construct meet met dezelfde meetprecisie, oftewel als de validiteit en betrouwbaarheid niet verandert door de aanpassing. Idealiter vinden aanpassingen aan tests plaats op het niveau van de test zelf, en is er empirisch onderzoek dat aannemelijk maakt dat vergelijkbare metingen worden verkregen, zowel met als zonder aanpassingen. Een voorbeeld hiervan wordt gegeven in **BOX 13**.

BOX 11 Item bias

Er is sprake van *item bias* wanneer personen verschillend reageren op een bepaalde vraag zonder dat deze personen van elkaar verschillen op de eigenschap die wordt gemeten. Het gaat hierbij om een verschil in de reactie op een vraag die niet te verklaren is door de gemeten eigenschap, maar die gerelateerd is aan andere specifieke individuele kenmerken, zoals culturele achtergrond of taal (wanneer dit niet samenhangt met de te meten eigenschap). De betreffende vraag meet dan dus iets anders dan het psychodiagnostische instrument beoogt.

Een voorbeeld hiervan is een vraag in een test voor studievaardigheden waarbij gebruik gemaakt moet worden van informatie uit een kleurenplaatje. Als het kleurgebruik in dat plaatje zodanig is dat personen die kleurenblind zijn de relevante informatie er niet uit kunnen halen, dan zegt het antwoord op die vraag niets over de studievaardigheid van personen met kleurenblindheid. Er is in dit geval duidelijk sprake van item bias omdat de vraag personen met kleurenblindheid benadeelt. Een ander voorbeeld van mogelijke item bias is een verhalende som om rekenvaardigheid te meten. Zo'n item kan bij bepaalde groepen personen eerder de taalvaardigheid dan de rekenvaardigheid meten.

BOX 12 Test bias

Het verschijnsel van *bias* kan ook bij een volledig psychodiagnostisch instrument gevonden worden: als reactietijd wordt gemeten met behulp van de computer en linkshandigen moeten met een rechtshandige muis werken, dan kan dit een vertekend (slechter) testresultaat geven. Een ander bekend voorbeeld van test bias is een intelligentietest waarin veel cultuurspecifieke vragen zijn opgenomen. Respondenten met die specifieke culturele achtergrond scoren dan veelal structureel hoger dan andere respondenten, hetgeen tot de onterechte conclusie kan leiden dat deze groep intelligenter is. Zo'n instrument meet dan echter meer cultuurkennis dan intelligentie.

BOX 13

Voorbeeld van een aangepaste test voor cliënten met fysieke beperkingen

Een voorbeeld van een test met aanpassingen voor cliënten met fysieke beperkingen is de Bayley-III-SNA (Special Needs Addition, Ruiters et al., 2014) voor het meten van de cognitieve ontwikkeling van zeer jonge kinderen met motorische en/of visuele beperkingen. Het materiaal is speciaal aangepast voor deze doelgroep en onderzoek heeft aangetoond dat zo metingen met dezelfde betrouwbaarheid en validiteit worden verkregen als bij kinderen zonder deze beperkingen die de standaard Bayley-III gebruiken. Verder zijn ook de testnormen van de standaard Bayley-III direct van toepassing op de resultaten verkregen met de Bayley-III-SNA. Het is belangrijk op te merken dat er dus geen aparte testnormen zijn voor de speciale

doelgroep. Het gebruik van dergelijke testnormen zou een geheel andere interpretatie hebben dan de standaard normen, aangezien de referentiepopulatie dan zou zijn kinderen met motorische en/of visuele beperkingen van dezelfde leeftijd. Omdat kinderen met deze beperkingen vaak ook cognitief beperkt zijn, zouden de testnormen aanzienlijk verschillen (d.w.z. lager zijn). Een vergelijking met deze populatie is niet zinvol voor de toepassing, omdat de onderzoeksvraag doorgaans is of het onderzochte kind een vertraagde cognitieve ontwikkeling heeft ten opzichte van de ontwikkeling van alle kinderen van dezelfde leeftijd.

Voor de meeste tests zijn aanpassingen voor cliënten met specifieke individuele kenmerken niet gestandaardiseerd en dus ook niet systematisch onderzocht. Het is dan aan de diagnosticus om op basis van de context en de inhoud te beoordelen of aanpassingen gerechtvaardigd zijn. Indien beschikbaar kan ook bestaand onderzoek naar de invloed van bepaalde hulpmiddelen op de scores op andere, vergelijkbare tests gebruikt worden om die afweging te maken. In de meeste gevallen moeten de standaardinstructies voor de testafname behouden blijven.

In het geval een cliënt individuele kenmerken heeft die het testresultaat onbedoeld kunnen beïnvloeden moet de diagnosticus deze kenmerken beschrijven en verantwoordelijk hoe hiermee is omgegaan bij de testafname. Een aanpassing van de standaard testwijze (bijvoorbeeld,

door extra tijd te geven of iets voor te doen) moet worden onderbouwd. De factoren die van invloed kunnen zijn op de validiteit van de testresultaten (d.w.z. de beperkingen en de testaanpassingen) moeten zo goed mogelijk worden meegewogen bij de interpretatie van de testresultaten. Voor meer informatie over dit zogenaamde *testing the limits* (en voorbeelden van hoe de diagnosticus dit kan toepassen) verwijzen we naar Dek en collega's (2021).

1.3 Het psychologisch rapport

Het psychologisch rapport vormt het eindresultaat van het psychodiagnostisch onderzoek. De inhoud van het psychologisch rapport is afgestemd op de specifieke vraagstelling en bevat bevindingen, beoordelingen, beslissingen of adviezen (Beroepscode, artikel 1.17 'Rapportage'). De inhoud van het rapport kan variëren, afhankelijk van de beoogde

ontvanger. In de behandelcontext is het rapport in de eerste plaats bedoeld voor de cliënt, die tevens de opdrachtgever is (volgens artikel 89 'Rapportage in opdracht van de cliënt' van de Beroepscode). In andere situaties, zoals in het onderwijs, bij selectieprocedures, of bij rapportages over sociale zekerheidskwesties, is het rapport bedoeld voor een externe opdrachtgever (Beroepscode, artikelen 90 'Toestemmingsvereiste voor rapportage aan derden', 91 'Rapportage aan derden' en 93 'Mondelinge rapportage aan een derde'). Ook in dat geval wordt verwacht dat de cliënt in staat is het rapport te begrijpen. In principe heeft de cliënt het recht om het rapport als eerste in te zien (Beroepscode, artikel 92 'Gelegenheid tot inzage voorafgaand aan de rapportage') en te beslissen of het naar de externe opdrachtgever mag worden gestuurd (Beroepscode, artikel 95 'Blokken van de rapportage aan de externe opdrachtgever'). In bepaalde gevallen kan het doel van de rapportage of de vertrouwelijkheid ten opzichte van anderen ertoe leiden dat er beperkingen gesteld worden aan het inzage- en blokkeringsrecht (Beroepscode, artikel 96 'Inzage- en blokkeringsrecht bij rapportage over een cliëntensysteem).

1.3.1 Onderdelen van het psychologisch rapport

Het psychologisch rapport bevat over het algemeen de volgende elementen:

1. Datum onderzoek, naam, gender en geboortedatum van de cliënt;
2. Aanleiding en beschrijving van de vraagstelling (opdracht);
3. Anamnese/intakegegevens;
4. De hypothesen en de hypothesetoetsingscriteria en de beschrijving van hoe alle verzamelde gegevens zullen worden geïntegreerd om de hypothesen te toetsen en daarmee de onderzoeksvragen te beantwoorden;
5. De gebruikte psychodiagnostische instrumenten (bronnen van informatie, zie hierna) en de psychometrische kwaliteit hiervan;
6. Het verloop van het onderzoek;
7. Resultaten van het onderzoek, inclusief rapportage van

onzekerheidsmarges (d.w.z. betrouwbaarheidsintervallen in plaats van puntscores), en relevante observaties en gespreksresultaten;

8. Samenvatting;
9. Conclusies en aanbevelingen;
10. De geldigheidsduur van de verschillende onderdelen van het rapport, waaronder de testresultaten;
11. De naam van de diagnosticus onder wiens verantwoordelijkheid het psychodiagnostisch onderzoek heeft plaatsgevonden. Het wordt aanbevolen dat de verantwoordelijke diagnosticus het rapport ondertekent;
12. De bewaartermijn van de testgegevens en het psychologisch rapport in het (digitale of papieren) dossier.

In het psychologisch rapport moeten de afzonderlijke bronnen van informatie kunnen worden teruggevonden (Beroepscode, artikel 98 'Rapportage beperken tot noodzakelijke gegevens'), inclusief de testresultaten en andere bronnen, zoals observaties van de cliënt en informatie verkregen door gesprekken met de cliënt en/of informanten, zoals ouders. De gebruikte psychodiagnostische instrumenten moeten in het rapport expliciet benoemd worden, in de tekst van de rapportage of in een bijlage.

In geval van een norm-georiënteerde test moet de aard van de referentiepopulatie (bijvoorbeeld: algemene Nederlandse bevolking van dezelfde leeftijd) genoemd worden en correct worden gebruikt bij de interpretatie van de scores. Voor zover een instrument keuze biedt in de referentiepopulatie (bijvoorbeeld geslachts- en leeftijdsafhankelijk of alleen leeftijdsafhankelijk; zie 1.2.5 'Criterium-georiënteerde test scores, norm-georiënteerde test scores en testnormen') moet de keuze worden gemotiveerd. Idealiter worden betrouwbaarheidsintervallen rond de gerapporteerde scores weergegeven. Zie **BOX 14** voor een uitleg over de betekenis, interpretatie en berekening van betrouwbaarheidsintervallen. Deze betrouwbaarheidsintervallen moeten gebruikt worden als basis voor de interpretatie. Zie 1.2.3 'Beschrijving van de onderzoeksvraag en de onderzoeksmethode' voor hoe dit aangepakt kan worden.

Het transparant combineren en afwegen van alle verkregen informatie over de cliënt is een zeer belangrijk onderdeel van het diagnostisch proces. Dit omvat verschillende methoden, zoals beschreven in de Handleiding voor het opstellen en gebruiken van beslisregels (Niessen et al., 2025), waarmee de diagnosticus alle relevante gegevens samenbrengt. Door op een transparante wijze te handelen, zowel in het diagnostische proces als in het rapportageproces, wordt een leermogelijkheid gecreëerd, voor de diagnosticus zelf, maar ook voor collega's wanneer deze casus wordt besproken binnen intervisie.

Het rapport moet duidelijk maken hoe de testresultaten verband houden met de geformuleerde uitspraken of adviezen. Hierbij gebruikt de diagnosticus in geval van een hypothesetoetsend onderzoek de vooraf geformuleerde criteria op basis waarvan de hypothese(s) verworpen of aanvaard worden. Als er niet gewerkt wordt met hypothesen, dan gebruikt de diagnosticus de vooraf bepaalde expliciete beschrijving van hoe alle verzamelde gegevens geïntegreerd zullen worden om de onderzoeksvragen te beantwoorden. Daarbij moeten deze resultaten steeds worden gekoppeld aan de individuele kenmerken en context van de cliënt.

De uitspraken in het rapport moeten worden onderbouwd, waarbij de diagnosticus zich beperkt tot het vermelden van die gegevens en beoordelingen die relevant zijn voor het doel van de rapportage. De conclusies hebben alleen betrekking op de aan de rapportage ten grondslag liggende doel- of vraagstelling (Beroepscode, artikel 26 'Voorkomen van onbedoeld gebruik en misbruik van rapportage').

Het is uitdrukkelijk niet de bedoeling om in het rapport alle ruwe scores, percentielen, en gestandaardiseerde scores op alle subschalen op te nemen. De informatieve waarde voor de cliënt en opdrachtgever is namelijk meestal nihil. Het is zinvol om in een algemene opmerking in het rapport te vermelden dat bij een doorverwijzing naar een volgende professional het complete cijfermatige rapport opgevraagd kan worden bij de diagnosticus.

BOX 14

Betrouwbaarheidsinterval: interpretatie en berekening

De interpretatie van een betrouwbaarheidsinterval van een testscore wordt in het artikel 'IQ Score is dringend aan modernisering toe' (Ruiter et al., 2017) als volgt uitgelegd:

‘Om een betrouwbaarheidsinterval goed te kunnen interpreteren, is het goed om iets meer te weten over de theorie rond meetfouten. Het idee is dat iedere test een bepaalde meeton nauwkeurigheid heeft. Volgens de klassieke testtheorie is iedere geobserveerde testscore (X) opgebouwd uit een betrouwbaar deel, ook wel ware score (T) genoemd, en een deel dat te wijten is aan de meetfout (E). Dat wil zeggen dat iedere individuele score op een test beïnvloed wordt door zowel het niveau van de persoon (wat we willen meten, de ware score) als door een meetfout (wat we niet willen meten). De ware score is hierbij simpelweg gedefinieerd als de gemiddelde geobserveerde score over (vele) herhaalde testafnames van dezelfde persoon, waarbij het niveau van de persoon hetzelfde blijft.

Dit laatste is in praktijk natuurlijk onmogelijk, want bij herhaalde testafnames zullen altijd herinnerings-, leer-, en vermoeidheidseffecten een rol gaan spelen. Dit betekent dat in de praktijk bij herhaalde afnamen de ware score verandert, en dus dat de meetfout niet te scheiden is van de ware score. We kunnen dan ook nooit bepalen welk deel van een geobserveerde score te danken is aan het niveau van de persoon, en welk deel te wijten is aan de meetfout. Dat betekent dat we nooit in staat zijn om exact weer te geven wat

het niveau van de persoon is. Wat we wel kunnen doen, is uitdrukking geven aan de mate waarin een geobserveerde score in het algemeen beïnvloed wordt door de meetfout. Die uitdrukking wordt gegeven door middel van een betrouwbaarheidsinterval. Bijvoorbeeld, een 95% betrouwbaarheidsinterval voor een IQ score van 78 tot 90 betekent dat we er voor 95% zeker van zijn dat de IQ score van deze persoon ligt tussen 78 en 90.

Bij iedere gestandaardiseerde testscore hoort een (bij voorkeur 90%) betrouwbaarheidsinterval gegeven te worden. Bij een goede test zorgt de testauteur ervoor dat de diagnosticus beschikt over gestandaardiseerde testcores én bijbehorende betrouwbaarheidsintervallen na testafname bij een cliënt.

Mocht alleen een gestandaardiseerde testscore gegeven zijn en het betrouwbaarheidsinterval ontbreken, dan kan de diagnosticus een betrouwbaarheidsinterval zelf berekenen, als de betrouwbaarheid van de test en de standaarddeviatie van de scores in de doelgroep van de test bekend zijn. Dit gaat als volgt:

Stap 1.

Bereken de zogenaamde Standard Error of Measurement (SEM).

$SEM = \sigma_x \sqrt{(1 - r_{xx'})}$, σ_x standaarddeviatie van de testcores in de doelgroep, $r_{xx'}$ de betrouwbaarheid van de test in de doelgroep.

Stap 2.

Bereken de grenzen van het 90% betrouwbaarheidsinterval als:

[testscore – 1.64 * SEM; testscore + 1.64 * SEM].

Voorbeeld:

Een IQ test is afgenomen bij een 22-jarige cliënt. De gestandaardiseerde IQ-score op de test is 128. De betrouwbaarheid van de test voor de doelgroep 20-30 jarigen is volgens de handleiding van de test $r_{xx'} = 0.94$.

Uitwerking:

IQ-scores zijn gestandaardiseerd, zodanig dat het gemiddelde 100 is, en de SD = 15.

Dit betekent:

$$SEM = \sigma_x \sqrt{(1 - r_{xx'})} = 15 * \sqrt{(1 - 0.94)} = 3.67$$

Het 90% betrouwbaarheidsinterval is:

$$[\text{testscore} - 1.64 * SEM; \text{testscore} + 1.64 * SEM] = [128 - 1.64 * 3.67; 128 + 1.64 * 3.67] = [128 - 6.02; 128 + 6.02] = [121.98; 134.02].$$

De diagnosticus is er dan voor 90% zeker van dat de IQ score van deze 22-jarige cliënt ligt tussen 122 en 134.

In artikel 98 'Rapportage beperken tot noodzakelijke gegevens' van de Beroepscode worden de minimale kwaliteits-eisen vermeld waaraan rapportages moeten voldoen.

Een goed rapport voldoet aan de volgende richtlijnen:

1. Het rapport vermeldt de feiten, omstandigheden en bevindingen waarop het berust;
2. Het rapport geeft blijk van een geschikte methode van onderzoek om de vraagstelling te beantwoorden;
3. In het rapport wordt op inzichtelijke en consistente wijze uiteengezet waar conclusies van het rapport op zijn gebaseerd;
4. Het rapport vermeldt de bronnen waar gebruik van is gemaakt, daaronder begrepen de gebruikte literatuur en de geconsulteerde personen;
5. De rapporteur blijft binnen de grenzen van diens deskundigheid.

1.3.2 Geautomatiseerde rapportages

Het psychologisch rapport wordt veelal door de diagnosticus zelf opgesteld. Tegenwoordig wordt steeds vaker gebruikgemaakt van (gedeeltelijk) geautomatiseerde rapportages. Bij het gebruik van geautomatiseerde rapportages is het van belang dat de verantwoordelijke diagnosticus bekend is met de interpretatieregels die in de rapportage worden toegepast. Als een normgroep gehanteerd is, dan moet de diagnosticus weten welke normgroep dat is. Als in een rapportage geautomatiseerd combinaties van scores worden berekend, zoals het vertalen van persoonlijkheidstest scores naar competentiescores, moet de diagnosticus de wijze van berekening begrijpen. Bij inzage van het geautomatiseerde rapport, moet de diagnosticus in staat zijn om uit te leggen hoe het tot stand is gekomen, zoals vereist door artikel 66 ('Inzage in en afschrift van het dossier') en artikel 92 ('Gelegenheid tot inzage voorafgaand aan de rapportage') van de Beroepscode. Dit kan problematisch zijn bij het gebruik van instrumenten die gebaseerd zijn op artificial intelligence en/of machine learning, omdat dan de berekeningswijze niet inzichtelijk is. Het is dan essentieel dat er goede evidentie is voor een goede validiteit van de resultaten, die relevant is in de context van de cliënt.

Antwoorden en ruwe scores op psychodiagnostische instrumenten worden steeds vaker digitaal opgeslagen. Vaak zijn alleen de gestandaardiseerde schaalscores dan in de rapportage opgenomen. Het is noodzakelijk dat de diagnosticus toegang heeft tot de antwoorden en de ruwe (item)scores, en ook cliënten dienen inzage te kunnen krijgen in deze gegevens.

1.3.3 Verantwoording

De verantwoording voor het psychologisch rapport is terug te vinden in het (digitale en/of papieren) dossier, zoals bepaald in artikel 33 ('Aflleggen van verantwoording') van de Beroepscode. De diagnosticus moet aantekening maken van de bronnen (zoals literatuurreferenties) voor diens uitspraken, tenzij deze al in de handleiding van het psychodiagnostisch instrument zijn vermeld. Als de interpretatie duidelijk afwijkt van wat in de handleiding staat, moet de diagnosticus een notitie maken van de redenering die tot de uitspraak heeft geleid.

1.3.4 Rechten van de cliënt

Het psychologisch rapport wordt doorgaans schriftelijk uitgebracht. De cliënt moet vóór aanvang van het onderzoek duidelijk en bij voorkeur schriftelijk worden geïnformeerd over diens rechten met betrekking tot het rapport. Deze rechten omvatten onder andere:

- Nabespreking van het onderzoek;
- Inzage in het rapport voorafgaand aan het uitbrengen ervan;
- Mogelijkheid tot blokkering van het rapport;
- Recht op correctie, aanvulling of verwijdering van gegevens;
- Kopie van het rapport na afgifte;
- Begeleide inzage in de ruwe testgegevens (zie ook 1.2.2.1 'Ruwe testgegevens'). De cliënt heeft recht op inzage in de ruwe testgegevens op itemniveau, maar geen recht op een kopie van deze ruwe testgegevens op itemniveau. De cliënt heeft wel recht op een kopie van ruwe schaalscores en genormeerde scores.

De cliënt heeft het recht om de rapportage aan de externe opdrachtgever te blokkeren, tenzij een wettelijke regeling dit recht uitsluit of beperkt. Als er geen wettelijke regeling van toepassing is, kan de diagnosticus, op basis van een zwaarwegend belang, de afweging maken het blokkeringsrecht niet aan de cliënt toe te kennen. De diagnosticus moet de cliënt hier voorafgaand aan het onderzoek schriftelijk over informeren (Beroepscode, artikel 95 'Blokken van de rapportage aan de externe opdrachtgever'). In gevallen waar de wettelijke regeling niet eenduidig is, kan de diagnosticus advies inwinnen tijdens het spreekuur Beroepsethiek van het NIP.

De externe opdrachtgever moet voorafgaand aan de aanvaarding van de onderzoeksopdracht op de hoogte worden gesteld van het blokkeringsrecht. Als de rapportage wordt geblokkeerd, moet de diagnosticus de externe opdrachtgever hiervan in kennis stellen zonder verdere toelichting.

Bij digitale psychodiagnostische instrumenten met automatisch gegenereerde rapportages geldt het recht op inzage en blokkering door de cliënt onverminderd. Dit betekent dat automatisch gegenereerde rapportages eerst aan de cliënt moeten worden verstrekt en pas na toestemming van de cliënt naar de opdrachtgever kunnen worden verzonden, zie **BOX 15**.

Het verzenden van psychologische rapportages moet op een veilige manier plaatsvinden. Het verdient de voorkeur om gebruik te maken van een met login en wachtwoord beveiligde webportal waar de cliënt de rapportage kan downloaden. De diagnosticus dient immers de vertrouwelijkheid van deze gegevens te waarborgen (Beroepscode, artikel 72 'Zorgvuldigheid in de communicatie').

BOX 15 Blokkeringsrecht

In het kader van (online) automatisch gegenereerde testrapportages is de uitspraak d.d. 16 december 2015 van het College van Toezicht (CvT) van het NIP in casus 15/31 informatief.

In deze casus werd een klacht behandeld over het verstrekken van de geautomatiseerde rapportage van een online gemaakte test aan de externe opdrachtgever. Het adviesbureau, dat onder de verantwoordelijkheid van de psycholoog was ingeschakeld om de testafname bij cliënt te faciliteren, had de testresultaten van cliënt direct doorgestuurd naar de opdrachtgever, zonder cliënt eerst inzage te geven en zonder cliënt de mogelijkheid te bieden om het blokkeringsrecht uit te oefenen.

Het CvT oordeelt dat de techniek bij het afnemen van de psychodiagnostische instrumenten zodanig had moeten worden ingericht dat aan de artikelen 91 (Beroepscode 2015, artikel 92 Beroepscode 2024) 'Gelegenheid tot inzage voorafgaand aan de rapportage' en 95 'Blokken van de rapportage aan de externe opdrachtgever' van de Beroepscode (2024) wordt voldaan. Het CvT is in deze casus tot het oordeel gekomen dat deze artikelen 92 en 95 van de Beroepscode (2024) zijn overtreden.

1.3.5 Verstrekking van gegevens verkregen uit tests aan derden

Verstrekking van gegevens uit tests aan derden vereist zorgvuldige aandacht voor vertrouwelijkheid. Als het rapporteren aan derden deel uitmaakt van de opdracht, dan geldt voor de gegevens die zijn opgenomen in het psychologisch rapport, geen geheimhoudingsplicht voor de ontvanger van het rapport (Beroepscode, artikel 73 'Geheimhouding bij rapportage en gegevensverstrekking'). Zonder expliciete toestemming van de cliënt kunnen derden of externe opdrachtgevers geen toegang krijgen tot het volledige dossier. Inzage is beperkt tot de bewerkte uitslagen van de tests. Ruwe testgegevens op item- of testniveau, ingevulde tests zelf, testprotocollen en gegevens uit observatieformulieren zijn alleen toegankelijk voor de cliënt, aangezien deze tot het dossier behoren (zie ook 1.2.2 'Dossiervorming').

Voor wetenschappelijk onderzoek kunnen gegevens aan derden worden verstrekt, maar alleen als de identiteit van de cliënt niet herleidbaar is (Beroepscode, artikel 85 'Gegevens voor publicaties, onderwijs, kwaliteitszorg, supervisie en intervisie').

Met de toegenomen digitalisering van tests is het van belang extra veiligheidsmaatregelen te treffen en afspraken te maken om de privacy van de cliënt te waarborgen, met name als testcores in een database worden opgeslagen. Dit kan onder meer worden geregeld via verwerkersovereenkomsten met leveranciers van digitale psychodiagnostische instrumenten. De 'Guidelines for Technology-based assessment' (International Test Commission and the Association of Test Publishers, 2022) bieden richtlijnen voor deze kwestie.

1.4 Dossiervoering

Gegevens worden niet langer bewaard dan noodzakelijk voor het doel van het psychodiagnostisch onderzoek of volgens wettelijke regels. Alle dossiers van behandelend psychologen die betrekking hebben op geneeskundige behandeling van patiënten vallen onder de wettelijke bewaartermijn van de Wet op de Geneeskundige Behandelingsovereenkomst (WGBO). Die bewaartermijn is in 2020 gesteld op 20 jaar voor een dossier, met mogelijke uitzonderingen als dit van belang is voor de continuïteit van zorg. Merk op dat lang niet alle dossiers volgens de Beroepscode vallen onder de WGBO. Voor deze dossiers dient de diagnosticus een bewaartermijn te kiezen op basis van wat gebruikelijk is in het betreffende werkveld. Dossiers op basis van de Beroepscode dienen minimaal te worden bewaard gedurende de geldigheidsduur van het psychologisch rapport, met een minimum van één jaar (Beroepscode, artikel 34 'Bewaartermijn van het dossier'). Als er een lopende klachtbehandeling is, mag de diagnosticus het dossier niet vernietigen na afloop van de bewaartermijn.

Geanonimiseerde gegevens ten behoeve van onderzoek, zoals testgegevens, kennen binnen de Beroepscode geen specifieke bewaartermijn. Ze mogen zonder termijn worden bewaard omdat ze niet (meer) tot de cliënt herleidbaar zijn (Beroepscode, artikel 86 'Gegevens voor publicatie, onderwijs, kwaliteitszorg, supervisie en intervisie'). De diagnosticus dient er voor te zorgen dat deze gegevens toegankelijk blijven voor statistische of psychometrische analyses, bijvoorbeeld door ze in een geanonimiseerd bestand op te slaan. Het is belangrijk dat deze gegevens gescheiden zijn van andere dossiergegevens.

Referenties

- Autoriteit Persoonsgegevens. (2008, 15 juli). Inzage testgegevens. Geraadpleegd op 22 december 2023, van <https://autoriteitpersoonsgegevens.nl/sites/default/files/downloads/overig/nip.pdf>
- Collie, A., Maruff, P., Darby, D.G., & McStephen, M. (2003). The effects of practice on the cognitive test performance of neurologically normal individuals assessed at brief test-retest intervals. *Journal of the international neuropsychological society*, 9(3), 419-428.
- De Bruyn, E.E.J., Ruijsenaars, A.J.J.M., Van Pameijer, N.K., & Van Aarle, E.J.M. (2003). *De diagnostische cyclus: Een praktijkleer*. Acco.
- Dek, J. & Hendriksen, J., & Hurks, P. (2021). Zo haal je meer uit een testafname: Stappenplan om extra diagnostische informatie te verzamelen. *Kind & Adolescent Praktijk*, 20, 16-22. <https://doi.org/10.1007/s12454-021-0662-8>
- EFPA. (2012). *EFPA Standards for Test Use: Work, Education, and Health & Social Care, Levels 1, 2 and 3*.
- Evers, A., Lucassen, W., Meijer, R., & Sijtsma, K. (2009). *COTAN Beoordelingssysteem voor de kwaliteit van tests* (geheel herziene versie, mei 2009; gewijzigde herdruk mei 2010). NIP.
- Frima, R.M., & Visser, R. (2008). Ruwe testgegevens: spanningsveld tussen openheid en bescherming. Een poging tot verzoening. *De Psycholoog*, 43(4), 236-238.
- International Test Commission and the Association of Test Publishers. (2022). *Guidelines for Technology-Based Assessment*. Geraadpleegd op 23 juli 2024, van <https://www.intestcom.org/upload/media-library/tba-guidelines-final-2-23-2023-v4-167785144642TgY.pdf>
- International Test Commission. (2017). *The ITC Guidelines for Translating and Adapting Tests* (Second edition).
- Niessen, A.S.M., Hurks, P.P.M., Neumann, M., & Meijer, R.R. (2025). *Handleiding voor het opstellen en gebruiken van beslisregels*. NIP.
- NIP. (2024). *Beroepscode voor psychologen 2024*.
- Ruiter, S.A.J., Hurks, P.P.M., & Timmerman, M.E. (2017). IQ-score is dringend aan modernisering toe: Naar een nieuwe interpretatie en classificatie van geschatte intelligentie. *Kind & Adolescent Praktijk*, 16, 16-23.
- Ruiter, S.A.J., Visser, L., van der Meulen, B.F., & Timmerman, M.E. (2014). *Bayley-III. Special Needs Addition (SNA)*. Pearson Assessment and Information B.V.
- Van de Vijver, F.J.R., & Leung, K. (2021). Methods and Design. In V.H. Fetvadjev, J., He, & J.R.J. Fontaine (Eds). *Methods and Data Analysis for Cross-Cultural Research* (pp. 29-63). Cambridge University Press.

Totstandkoming van de AST-NIP 2024

Redactie

De herziening van de AST-NIP is voorbereid door prof. dr. Marieke E. Timmerman en prof. dr. Petra P.M. Hurks (leden van de Commissie Testaangelegenheden Nederland (COTAN)), in samenwerking dr. Remko H. van den Berg (COTAN lid), dr. Wendy E. de Leng (senior eindredacteur COTAN) en mr. Rosalinde Visser (senior stafmedewerker Beroepsethiek van het NIP). Bij de herziening is dankbaar gebruik gemaakt van suggesties van dr. Selma Ruiters (gedragswetenschapper bij de Kinderacademie Groningen), prof. dr. Maretha M.V. de Jonge (hoogleraar orthopedagogiek, Universiteit Leiden), dr. Jolinde Spoelstra-Niezink (klinisch neuropsycholoog), drs. Yaron Kaldenbach (kinder- en jeugdpsycholoog, gz-psycholoog en cognitief gedragstherapeut, Centrum Vela en Apollo Praktijk), dr. Susan Niessen (universitair docent Psychometrie & Statistiek, Rijksuniversiteit Groningen, externe beoordelaar COTAN), de kamer Gz-psychologen, sectie Forensische psychologie en Centrale Commissie Culturele Diversiteit en Psychologie.

Besluitvorming

De COTAN heeft de AST-NIP besproken en geaccordeerd in haar vergadering van 26 september 2024. De Centrale Commissie Beroepsethische Zaken (BEZ) van het NIP heeft de AST-NIP besproken in haar vergadering van 9 oktober 2024 en heeft daarbij in het advies aan het Algemeen Bestuur van het NIP met name beoordeeld of de AST-NIP in overeenstemming is met de Beroepscode.

Het Algemeen Bestuur van het NIP heeft de AST-NIP vastgesteld op 16 september 2024. De AST-NIP 2024 is in werking getreden op 13 maart 2025.

Het Nederlands Instituut van Psychologen

Het Nederlands Instituut van Psychologen (NIP) is met ruim 17.000 leden al meer dan 85 jaar de grootste vereniging van psychologen in Nederland.

Het NIP behartigt de belangen van academisch opgeleide psychologen, bevordert de inzet van psychologische kennis, versterkt de positie van psychologen van alle beroepsniveaus en in alle werkvelden en ondersteunt haar leden in hun beroepsuitoefening.

Het NIP is de beroepsvereniging van alle wetenschappelijk opgeleide psychologen in Nederland, vanaf hun studie, tijdens hun hele loopbaan en in ieder werkveld. Het NIP verenigt en ondersteunt psychologen, bepaalt de professionele norm van het vak en positioneert de psychologen en de psychologie.

Lees meer over het NIP op de website: nip.nl