

Evaluatie van de ABAS-3

Met focus op de Vlaamse normering



KWALITEITSCENTRUM
DIAGNOSTIEK^{VZW}



KWALITEITSCENTRUM
DIAGNOSTIEK^{vzw}



Vlaanderen
is zorgzaam samenleven

Kwaliteitscentrum voor Diagnostiek vzw
Kortrijksesteenweg 405
9000 Gent

Website: www.kwaliteitscentrumdiagnostiek.be
E-mail: communicatie@kwaliteitscentrumdiagnostiek.be

Titel: Evaluatie van de ABAS-3 – met focus op de Vlaamse normering
Redactie: Kwaliteitscentrum voor Diagnostiek vzw
Datum: Januari 2023

Dit rapport kwam tot stand met de steun van de Vlaamse Overheid. In deze tekst komen onderzoeksresultaten van de auteur(s) naar voor en niet van de Vlaamse Overheid. Het Vlaams Gewest kan niet aansprakelijk gesteld worden voor het gebruik dat kan worden gemaakt van de meegedeelde gegevens. Niets uit deze uitgave mag worden veeveelvoudigd en/of openbaar gemaakt zonder uitdrukkelijk te verwijzen naar de bron.

Reviewprocedure

De evaluatie van de Adaptive Behavior Assessment System – versie 3 – Nederlandstalige versie (ABAS-3) kwam tot stand na raadpleging van twee onafhankelijke externe beoordelaars met expertise inzake testconstructie, psychometrie en/of het meten van adaptieve vaardigheden. Beide externe beoordelaars hebben het testmateriaal en de handleidingen van de ABAS-3 grondig onderzocht en beoordeeld aan de hand van het beoordelingsmodel voor de beschrijving en evaluatie van psychologische en educatieve testen van de European Federation of Psychologists' Associations (EFPA; Evers et al., 2013), naar het Nederlands vertaald door het Kwaliteitscentrum voor Diagnostiek vzw¹. De gemotiveerde beoordelingen werden onafhankelijk door beide externe experts bezorgd aan het Kwaliteitscentrum voor Diagnostiek vzw. Daarnaast heeft een wetenschappelijk onderzoeker van het Kwaliteitscentrum voor Diagnostiek vzw onafhankelijk een beoordeling gemaakt van de ABAS-3, eveneens aan de hand van het EFPA-beoordelingsmodel. Het Kwaliteitscentrum voor Diagnostiek vzw integreerde de drie onafhankelijke beoordelingen tot een eerste versie van dit adviesrapport dat vervolgens werd voorgelegd aan de externe beoordelaars waarna enkele zaken zijn aangepast. Deze tweede versie van het adviesrapport werd voorgelegd aan de testauteurs. Via een gemotiveerd schrijven hebben de testauteurs gereageerd op de inhoud van het adviesrapport en extra toelichting gegeven omtrent enkele opmerkingen. Op basis van de inhoudelijke toelichting door de testauteurs werd het adviesrapport herwerkt door de wetenschappelijk medewerker van het Kwaliteitscentrum. Dit resulteerde in de finale versie van het adviesrapport. Doorheen het geïntegreerd adviesrapport wordt de structuur van het EFPA-beoordelingsmodel zoveel mogelijk aangehouden.

Het adviesrapport heeft voornamelijk betrekking op het gebruik van de ABAS-3 in Vlaanderen.

¹ <http://assessment.efpa.eu/documents/> en <https://portaal.kwaliteitscentrumdiagnostiek.be/wp-content/uploads/2021/01/EFPAtestbeoordeling2020.docx>

Algemeen

Jaar van uitgave:	2020
Auteur(s):	Harrison, P. L. & Oakland, T. (2015)
Bewerker(s):	Kreemers, B., Maljaars, J., Briers, V., Bulteel, S., Storms, G., Maes, B., & Noens, I. (2020)
Uitgever:	Hogrefe uitgevers B.V. Amsterdam.
Referenties:	Harrison, P. L., & Oakland, T. (2015). <i>Adaptive Behavior Assessment System, Third Edition</i> [Manual]. Torrance, CA: Western Psychological Services. Kreemers, B., Maljaars, J., Briers, V., Bulteel, S., Storms, G., Maes, B., & Noens, I. (2020). <i>ABAS-3 Schaal voor adaptief gedrag – Instructies en verantwoording</i> . Amsterdam: Hogrefe Uitgevers B.V. Kreemers, B., Maljaars, J., Briers, V., Bulteel, S., Storms, G., Maes, B., & Noens, I. (2020). <i>ABAS-3 Schaal voor adaptief gedrag – Vlaamse normen en aanvullende analyses</i> . Amsterdam: Hogrefe Uitgevers B.V.

Meetpretentie, doelgroep en gebruiksdoel

De Schaal voor Adaptief Gedrag (2020) is de Nederlandstalige versie van de ABAS-3, Adaptive Behavior Assessment System – Third Edition (Harrison & Oakland, 2015). Het is een observatieschaal met als doel op een valide en betrouwbare wijze adaptief gedrag van kinderen (vanaf 3 maanden) en volwassenen (tot 80 jaar) in kaart te brengen. Het instrument meet vaardigheden die belangrijk zijn voor diverse gebieden in het dagelijks leven zoals communicatie en sociaal contact, vrijetijdsbesteding, het zorgen voor de eigen gezondheid en veiligheid, en het uitvoeren van activiteiten in een thuis-, school- of werkcontext.

Het instrument werd ontwikkeld om zowel gebruikt te worden in het kader van handelingsgerichte diagnostiek bij o.a. personen met een verstandelijke beperking, autismespectrumstoornissen (ASS), aandachtsdeficiëntie/hyperactiviteitsstoornis (ADHD), niet-aangeboren hersenletsel (NAH), sensorische/motorische beperkingen, of hoogbegaafdheid, als in het kader van classificerende of onderkende diagnostiek naar verstandelijke beperking. Tekorten in adaptief gedrag zijn namelijk een essentieel criterium voor het vaststellen van een verstandelijke beperking volgens de DSM-5 (APA, 2013). Dit gaat over tekorten in vaardigheden die in verschillende contexten tot uiting komen (vaardigheidstekort) en niet verklaard worden door gebrekkige motivatie en/of gebrek aan mogelijkheden om het gedrag te stellen of te oefenen (prestatietekort). Ook in het kader van handelingsgerichte diagnostiek is het van belang om aan de hand van een sterkte-zwakte profiel een correcte inschatting te kunnen maken van iemands functioneren aangezien dit aanknopingspunten biedt voor gepaste interventies. Zo kan het bepalend zijn voor het type onderwijs waartoe iemand toegang heeft of voor het soort hulp dat kan worden aangeboden (Kreemers et al., 2020).

Aangezien dit gaat over *belangrijke beslissingen op individueel niveau*² is het van cruciaal belang dat adaptieve vaardigheden op een betrouwbare en valide manier gemeten kunnen worden. De auteurs benadrukken terecht dat de resultaten van de ABAS-3 altijd in combinatie met andere (test)resultaten moeten geïnterpreteerd worden. Dergelijke belangrijke beslissingen vragen namelijk om een uitgebreid onderzoek met meerdere methoden en informanten en mogen niet louter op een momentopname gebaseerd zijn.

² Belangrijke beslissingen worden door de COTAN gedefinieerd als beslissingen die in principe op korte termijn onomkeerbaar zijn en die voor een belangrijk deel buiten de geteste persoon om worden genomen, zoals personeelsselectie, verwijzing naar speciaal onderwijs, opname/ontslag kliniek, certificering, enzovoort (Evers, et al., 2010, pg. 22).

De ABAS-3 wordt ook ingezet voor relatief *minder belangrijke beslissingen op individueel niveau*³, bijvoorbeeld voor voortgangscntrole. In het huidig adviesrapport wordt de ABAS-3 beoordeeld voor het nemen van zowel belangrijke als minder belangrijke beslissingen op individueel niveau. Voor belangrijke beslissingen zijn de kwaliteitseisen betreffende normering en betrouwbaarheid strenger dan deze gesteld aan testen voor minder belangrijke beslissingen.

'Het instrument werd ontwikkeld om zowel gebruikt te worden in het kader van classificerende of onderkende diagnostiek naar verstandelijke beperking, als in het kader van handelingsgerichte diagnostiek.'

Structuur

Er bestaan in het totaal zes versies van de ABAS-3, opgedeeld in drie leeftijdscategorieën waarvoor telkens twee informantversies bestaan:

- Jonge kinderen (0 – 4 jaar) met een versie voor ouders/verzorgers en een versie voor leerkrachten/begeleiders
- Kinderen en jongeren (5 – 17 jaar) met een versie voor ouders/verzorgers en een versie voor leerkrachten/begeleiders
- Volwassenen (18 – 80 jaar) met een zelfrapportagelijst en een informantenvragenlijst

Om tot een zo betrouwbaar mogelijke inschatting van het adaptief gedrag te komen, is het wenselijk dat meerdere informanten uit verschillende contexten de vragenlijst invullen.

De ABAS-3 onderscheidt drie domeinen van adaptief gedrag: *het conceptueel domein*, *het sociaal domein* en *het praktisch domein*. Deze worden verder onderverdeeld in negen of tien vaardigheidsgebieden. Het *conceptueel domein* bestaat uit drie vaardigheidsgebieden, namelijk 'Communicatie', 'Functionele (voor)schoolse vaardigheden' en 'Zelfsturing'. Het *sociaal domein* bestaat uit twee vaardigheidsgebieden: '(Spel en) Vrije Tijd' en 'Sociale Vaardigheden'. Het *praktische domein* bestaat, afhankelijk van de leeftijdscategorie, uit vier of vijf vaardigheidsgebieden. De gebieden zijn 'Maatschappelijke Vaardigheden', 'Huishoudelijke/Schoolse Vaardigheden', 'Gezondheid en Veiligheid', en 'Zelfverzorging'. In de versie voor 18-80-jarigen wordt ook het vaardigheidsgebied 'Werk' meegenomen. Voor jonge kinderen is er een apart vaardigheidsgebied toegevoegd om de 'Motoriek' in te schalen. Bij het invullen van de schaal wordt een onderscheid gemaakt tussen een vaardigheidstekort en een prestatietekort (zie supra).

De structuur van de ABAS-3 is compatibel met de definitie van adaptief gedrag die gehanteerd wordt door de American Association on Intellectual and Developmental Disabilities (AAIDD, 2010) en de DSM-5 (American Psychiatric Association, 2013). De totaalscore omvat alle gemeten vaardigheidsgebieden en vormt een globale beoordeling van adaptief gedrag. Bij de beoordeling wordt de nadruk gelegd op de totaalscore en de scores op de domeinen.

³ De COTAN geeft als voorbeelden voor relatief minder belangrijke beslissingen op individueel niveau: voortgangscntrole, beschrijvend gebruik van de testresultaten, therapie-indicatie en beroepskeuzebegeleiding (Evers, et al., 2010, pg. 22).

Daarnaast is er voor elke leeftijdscategorie een interventieplanner beschikbaar. Hierin worden voor elk item van de ABAS-3 suggesties gedaan over activiteiten en taken om het adaptief functioneren op verschillende domeinen en gebieden te bevorderen.

Normen

De ABAS-3 beschikt over zowel Nederlandse als Vlaamse normen voor elke versie van de vragenlijst. De normeringen werden in beide landen afzonderlijk uitgevoerd. Er zijn normen voor de totaalscores, domeinscores en vaardigheidsscores. Met het oog op het samenstellen van respectievelijk een representatieve Vlaamse en Nederlandse normeringssteekproef werd rekening gehouden met leeftijd, geslacht, opleidingsniveau (van de moeder), migratieachtergrond en regio in het land (enkel voor de volwassenversies). De normering van de ouder- en leerkrachtversies gebeurde aan de hand van continue normering en de methode van cumulatieve proporties (zie infra). Naast normscores werden eveneens leeftijdsequivalenten bepaald. Voor de zelfrapportage- en informantversie 18 – 80 jaar werd de methode van cumulatieve proporties gehanteerd.

Aangezien adaptief gedrag onlosmakelijk is verbonden met de omgeving, context, cultuur en samenleving waarin iemand opgroeit en de ABAS-3 ontwikkeld is voor personen die in Nederland of Vlaanderen leven, is het belangrijk om de nodige voorzichtigheid aan de dag te leggen bij het afnemen en interpreteren van de ABAS-3 bij een aantal doelgroepen, waaronder personen met een migratieachtergrond of met een lage sociaal-economische status (SES).

Afname en scoring

De ABAS-3 kan worden afgenomen door professionals met relevante training, kennis, en ervaring met betrekking tot de basisprincipes van psychodiagnostiek (Kreemers et. al., 2020). Onder andere op de website van de testuitgeverij⁴ wordt dit verder gespecificeerd en wordt aangegeven dat het instrument gebruikt kan worden door een psycholoog, (ortho)pedagoog, bachelor toegepaste psychologie, psychiater, of een andere opleiding die leidt tot een ruime kennis van testconstructie, testtoepassing en testinterpretatie. De auteurs wijzen erop dat op de website van het instrument⁵ vermeld wordt dat het in bepaalde settings (bv. CLB) aangewezen kan zijn dat de afname gebeurt door iemand die een nauw contact heeft met het gezin (bv. maatschappelijk werker). In een dergelijke situatie gebeurt de afname in samenwerking met, en onder supervisie van iemand die voldoende opgeleid is in psychodiagnostiek. Deze persoon blijft verantwoordelijk voor de voorbereiding, scoring en interpretatie. Verder wordt aangeraden dat het instrument wordt ingevuld in een afleidingsvrije en stille ruimte waarbij de testleider aanwezig kan zijn om het gebruik toe te lichten, vragen te beantwoorden, en toe te zien op het invullen van de vragenlijst (Kreemers et. al., 2020). Voor mensen met moeilijkheden met taal, lezen en/of begrijpen, kan het ook in de vorm van een gestructureerd interview afgenomen worden of kunnen de vragen voorgelezen worden. Een belangrijke kanttekening hierbij is dat de resultaten dan enkel kwalitatief geïnterpreteerd kunnen worden (i.e., geen gebruik maken van de normen).

Het invullen van de ABAS-3 duurt ongeveer 30 minuten per vragenlijst. De informant geeft voor elk item op een vierpuntsschaal aan in welke mate iemand bepaalde adaptieve vaardigheden laat zien. Wanneer de informant twijfelt, kan men aangeven dat de aangeduide score een inschatting is. Indien er vier of meer inschattingen zijn per vaardigheidsgebied, kan dit nog eens besproken worden met de informant om het aantal inschattingen te verminderen. Indien dit niet lukt, kan men ook een extra informant inschakelen. In de normeringssteekproef komen er maximum twee inschattingen per vaardigheidsgebied voor (Kreemers et. al., 2020).

Het scoren en interpreteren van de basisanalyses duurt ongeveer 15 minuten. Het vergt van de testleider de nodige opleiding en een gedegen inzicht in adaptieve vaardigheden om de resultaten op een gepaste manier te kunnen interpreteren. De ruwe scores worden berekend door per vaardigheidsgebied de itemscores op te

⁴ <https://www.hogrefe.com/nl/shop/abas-3-schaal-voor-adaptief-gedrag.html#1+1>

⁵ <https://ppw.kuleuven.be/ogop/abas-3>

tellen en op het scoreformulier over te nemen. De ruwe scores van de vaardigheidsgebieden worden omgezet naar een normscore met een gemiddelde van 10 en standaarddeviatie (SD) van 3. Dit staat beter bekend als de afgeleide uitslag (AU) die ook bij intelligentietesten vaak gebruikt wordt. Gemiddelde scores zijn scores die minder dan één SD boven of onder het gemiddelde liggen (AU = 8-12). Eén SD boven of onder het gemiddelde is een hoge (AU = 13-15) of een lage (AU = 5-7) score. Twee SD boven of onder het gemiddelde is een zeer hoge (AU ≥ 16) of een zeer lage (AU ≤ 4) score.

Het scoreformulier biedt een overzicht van alle scores. Zo wordt voor elk vaardigheidsgebied de ruwe score, de normscore, het leeftijdsequivalent en het leeftijdsbereik weergegeven. Per domein en voor het totaal is er een overzicht van de somscores (i.e., som van de normscores van de vaardigheidsgebieden), normscores, betrouwbaarheidsintervallen en percentielscores. Dit alles kan visueel samengevat worden op het profielformulier, waar ook bodem- en plafondeffecten worden aangeduid. Eveneens kunnen er op het scoreformulier aanvullende analyses gedaan worden zoals een vergelijking tussen de domeinen, het bekijken van de spreiding van de vaardigheidsgebieden en een sterkte- en zwakteanalyse van de vaardigheidsgebieden.

Op de website van de uitgeverij wordt een extra scoreformulier ter beschikking gesteld om de verschillen tussen informanten die elk de ABAS-3 invulden over eenzelfde kind/jongere (0-4 jaar en 5-17 jaar) na te gaan. Contextuele verschillen zijn te verwachten, maar men kan op deze manier nagaan of deze statistisch significant zijn en/of het al dan niet gebruikelijke verschillen zijn.

Afname

Afname:	Individueel
Wijze:	Papier-en-potlood
Tijdsduur:	ongeveer 30 min
Door:	Professionals met relevante training, kennis, en ervaring met betrekking tot de basisprincipes van psychodiagnostiek <ul style="list-style-type: none">- psycholoog, (ortho)pedagoog, bachelor toegepaste psychologie, psychiater- een andere opleiding die leidt tot een ruime kennis van testconstructie, testtoepassing en testinterpretatie- In bepaalde situaties kan het ook door een maatschappelijk werker (onder toezicht van iemand uit de bovenstaande groepen)

Scoring

Papier-en-potlood:	Handmatig
Tijdsduur:	ongeveer 15 min voor de basisanalyses
Door:	<ul style="list-style-type: none">- professionals met relevante training, kennis, en ervaring met betrekking tot de basisprincipes van psychodiagnostiek- psycholoog, (ortho)pedagoog, bachelor toegepaste psychologie, psychiater- een andere opleiding die leidt tot een ruime kennis van testconstructie, testtoepassing en testinterpretatie

Interpretatie

Door:	Professionals met relevante training, kennis, en ervaring met betrekking tot de basisprincipes van psychodiagnostiek <ul style="list-style-type: none">- psycholoog, (ortho)pedagoog, bachelor toegepaste psychologie, psychiater- een andere opleiding die leidt tot een ruime kennis van testconstructie, testtoepassing en testinterpretatie
-------	--

'Om tot een zo betrouwbaar mogelijke inschatting van het adaptief gedrag te komen, is het wenselijk dat meerdere informanten uit verschillende contexten de vragenlijst invullen.'

Beoordeling

Kwaliteit uitgangspunten, presentatie en beschikbare informatie

Uitgangspunten van de test

In de handleiding van de ABAS-3 wordt een beknopte beschrijving gegeven van de theoretische constructen waarop de test is gebaseerd. Het gaat hierbij om verschillende vaardigheden die belangrijk zijn in het dagelijks leven zoals sociale, communicatieve en schoolse vaardigheden, maar evenzeer over vaardigheden omtrent zelfverzorging, gezondheid, maatschappelijk functioneren, enzovoort. Deze vaardigheden worden geclusterd in drie verschillende domeinen: het sociale domein, het conceptuele domein en het praktische domein. De theoretische onderbouwing van deze constructen komt slechts gering aan bod. Een verdere uiteenzetting van relevante en recente literatuur die de structuur van de ABAS-3 onderbouwt, voornamelijk de opdeling in drie verschillende domeinen, is tevens beperkt. Aangezien alle vaardigheidsschalen sterk met elkaar correleren en ze niet als verschillende factoren te onderscheiden zijn, wordt in de handleiding geïmpliceerd dat deze opdeling vooral op basis van nuanceverschillen tot stand is gekomen. De auteurs staan voldoende stil bij de historische en recente trends op vlak van adaptief gedrag, alsook bij wat de relevantie is van het meten van adaptief gedrag. Er is voldoende informatie voor handen om de test op een correcte manier te kunnen gebruiken, maar indien de testleider onvoldoende vertrouwd is met adaptieve vaardigheden, wordt aangeraden om andere bronnen te raadplegen.

De procedure van testontwikkeling wordt uitgebreid en helder omschreven in de handleiding. Deze test kwam tot stand op basis van de Amerikaanse versie van de ABAS-3. Er bestond voorheen nog geen Nederlandstalige versie van deze test. De ontwikkeling van de ABAS-3 bestond uit verschillende fasen: vertaling en terugvertaling van het Engels naar het Nederlands, pilootstudie, dataverzameling, normeringsonderzoek, betrouwbaarheids- en validiteitsonderzoek. Er was hierbij telkens voldoende aandacht voor mogelijke biases en faire diagnostiek. De items werden vertaald en terugvertaald waarbij woorden werden aangepast rekening houdende met de linguïstiek in de Nederlandse en Vlaamse context. De vertaling werd voorzien van feedback van verschillende experts met aandacht voor cultuurgevoeligheid, toepasbaarheid in België en Nederland en de formulering van de items (complexiteit, woordkeuze, lengte). Verder wordt geen informatie verschaft omtrent het ontwikkelen van de items of analyses op itemniveau. Mogelijk komt dit wel aan bod in de handleidingen van de eerdere Amerikaanse versies van de test (ABAS en ABAS-2).

Beschikbare documentatie

De handleiding van de ABAS-3 biedt voldoende en duidelijke informatie over de ontwikkeling, normen, betrouwbaarheid en validiteit van de test. Het is toegankelijk en helder geschreven en te begrijpen zonder noodzakelijke doorgedreven psychometrische kennis.

Wat betreft de ontwikkeling van de test wordt er voldoende informatie verstrekt over het ontstaan van de Nederlandstalige bewerking. Er wordt echter weinig tot geen informatie gegeven over het tot stand komen van de oorspronkelijke ABAS-3. Over het algemeen kan gesteld worden dat de handleiding een eerder beperkte referentielijst heeft. Er zijn vooral bronnen opgenomen over (psychometrisch) onderzoek en statistiek, en minder bronnen die de inhoudelijke onderwerpen (o.m. het tot stand komen van de items) behandelen. De opgenomen, inhoudelijke bronnen zijn vaak gericht op specifieke onderwerpen en doelgroepen, en zijn weinig recent.

Wat betreft documentatie over de normen en het normonderzoek wordt duidelijk omschreven hoe de normgroepen zijn samengesteld, hoe groot ze zijn en op welke manier men beoogde om een representatieve steekproef samen te stellen.

In de handleiding wordt voldoende aandacht besteed aan de verschillende vormen van betrouwbaarheid. Zo worden zowel de interne consistentie, test-hertest betrouwbaarheid, interbeoordeelaarsbetrouwbaarheid en paralleltestbetrouwbaarheid helder besproken. Er is telkens informatie over welke maat (maten) gebruikt werd(en) en wat deze betekenen. De test-hertest betrouwbaarheid werd voor Vlaanderen (tussentijd: twee weken) onderzocht voor de vier versies van de ABAS-3 die gericht zijn op kinderen en jongeren. Er is geen informatie beschikbaar over de test-hertest betrouwbaarheid van de zelfrapportage en informant rapportage bij volwassenen. Dit werd wel onderzocht in een Nederlandse groep met een tussentijd van vier weken. De interbeoordeelaarsbetrouwbaarheid werd voor Vlaanderen enkel onderzocht voor de twee ouderversies van de ABAS-3. Tenslotte werd de paralleltestbetrouwbaarheid niet onderzocht voor de zelfrapportage en informant rapportage.

Ten slotte worden verschillende vormen van validiteit onder de loep genomen. De convergente en divergente validiteit werden niet voor elke versie van de vragenlijst onderzocht. Ook vergelijkt men slechts met een beperkt aantal andere testen. Er is slechts geringe informatie beschikbaar over de criteriumvaliditeit.

Procedurele instructies voor de testleider

Testleiders beschikken in de handleiding over uitvoerige en heldere informatie over de afname, scoring en interpretatie van de ABAS-3. Er worden richtlijnen en instructies voorzien over afnamecondities, communicatie en specifieke situaties (bv. elders invullen, meerdere informanten, anderstaligheid). Een lijst met vaak gestelde vragen, aanbevelingen omtrent faire diagnostiek en verschillende uitvoerig uitgewerkte casussen is eveneens opgenomen in de handleiding. Verder omschrijven de auteurs hoe men aan de slag moet met de verkregen ruwe scores en hoe om te gaan met ontbrekende scores. Ook wat betreft de betekenis en interpretatie van de scores wordt voldoende informatie voorzien. Voorzichtigheid is aangewezen wanneer er sprake is van bodem- en plafondeffecten, waarover evenzeer informatie terug te vinden is in de handleiding. Het profielblad geeft een duidelijke, overzichtelijke en visuele samenvatting van de scores op de ABAS-3. Er ontbreken in de handleiding echter instructies over hoe een testleider mondeling en/of schriftelijk feedback kan geven aan de betrokkenen of aan derden. Idealiter wordt er eveneens enige omkadering voorzien omtrent het bespreken van de moeilijkere items aangezien dit confronterend kan zijn voor ouders of volwassenen. De auteurs geven aan dat dit aan bod komt in de webinars⁶ over de ABAS-3.

Uit de psychometrische onderzoeken blijkt dat de normen niet zonder meer bruikbaar zijn voor kinderen, jongeren en volwassenen met een migratieachtergrond of een lage SES, of bij informanten die de Nederlandse taal niet voldoende machtig zijn. De normen moeten bij deze doelgroepen dus met grote voorzichtigheid gehanteerd worden. Bij de interpretatie en rapportage van de resultaten dient hier extra

⁶ <https://ppw.kuleuven.be/ogop/abas-3/opleiding>

aandacht aan besteed te worden. In de handleiding wordt omschreven hoe de testleider hier concreet mee kan omgaan.

Sterktes

- Er is voldoende documentatie voor de testleider beschikbaar en de kwaliteit van de uitleg van de basisprincipes is goed. De beschikbaarheid van documentatie over de normen en de kwaliteit van de procedurele instructies is uitstekend.
- De aanwezigheid van richtlijnen voor de testleider zoals aanbevelingen in het kader van faire diagnostiek, een lijst met vaak gestelde vragen en specifieke instructies die toepasbaar zijn in bepaalde situaties, is een grote meerwaarde.
- Naast de handleiding is er ook een website (<https://ppw.kuleuven.be/ogop/abas-3/abas-3>) en worden webinars georganiseerd om de testleider zoveel mogelijk voor te bereiden en te ondersteunen in het gebruik van de ABAS-3.

Aandachtspunten

- Het is belangrijk dat de testleider voldoende achtergrondkennis heeft over adaptieve vaardigheden.
- Er is geen informatie over de test-hertest betrouwbaarheid voor de Vlaamse versies gericht op volwassenen, enkel voor de Nederlandse versies van de zelf- en informant rapportage. De paralleltestbetrouwbaarheid werd niet onderzocht in een volwassen populatie.
- De normgegevens zijn niet zonder meer bruikbaar bij kinderen, jongeren en volwassenen met een migratieachtergrond of een lage SES, alsook bij informanten die de Nederlandse taal niet voldoende machtig zijn. Zoals hierboven aangegeven wordt er in de handleiding uitgebreid aandacht besteed aan de richtlijnen i.v.m. faire diagnostiek.
- Er werden geen richtlijnen of instructies opgenomen in de handleiding over mondelinge of schriftelijke rapportage van de resultaten van de ABAS-3. Dit komt wel aan bod in de webinars over de ABAS-3 (een verwijzing hiernaar in de handleiding kan aangewezen zijn).

Algemeen genomen kan de **kwaliteit van de uitgangspunten, presentatie en beschikbare informatie** als **goed** worden beoordeeld voor zowel het nemen van *belangrijke* als *minder belangrijke beslissingen op individueel niveau* voor alle versies van de ABAS-3.

'De kwaliteit van het testmateriaal kan als uitstekend beoordeeld worden.'

Kwaliteit van het testmateriaal

Algemene kwaliteit

Zoals eerder aangegeven bestaan er zes versies van de ABAS-3, opgedeeld in drie leeftijdscategorieën met telkens twee informantversies. De startkoffers bevatten telkens beide informantversies van eenzelfde leeftijdscategorie. Een ABAS-3 startkoffer bevat twee handleidingen (i.e., 'Instructies en verantwoording' en 'Vlaamse normen en aanvullende analyses'), 10 vragenlijsten van de ene informantversie met 10 bijhorende scoreformulieren en 10 vragenlijsten van de andere informantversie met 10 bijhorende scoreformulieren, en de interventieplanner van de betreffende leeftijdscategorie. Er bestaan telkens aparte koffers voor Nederland en Vlaanderen. De handleidingen, vragenlijsten met scoreformulieren en de interventieplanners zijn eveneens elk apart te verkrijgen.

De antwoordformulieren en scoreformulieren zijn zeer goed opgesteld. De instructies voor de informant worden duidelijk omschreven. Er worden richtlijnen gegeven over wanneer een bepaalde score gegeven kan worden en er worden voorbeelden voorzien. De informant heeft bij elk item de mogelijkheid om aan te duiden dat de beoordeling gebaseerd is op een inschatting. Ook hierover wordt informatie op het vragenlijstformulier voorzien. De formulering van de testitems is aangepast aan het taalgebied waar het gebruikt wordt.

Vervolgens tracht men zoveel mogelijk rekening te houden met de aanbevelingen van faire diagnostiek. Zo kan de omschrijving van sommige items aangepast worden naargelang de achtergrond van de persoon waarover de vragenlijst wordt ingevuld (bv. ouderversie 0-4 jaar – communicatie – item 12: het woord pop kan aangepast worden naar een ander evident woord in de moedertaal van het kind). Aangezien adaptieve vaardigheden erg context- en cultuurafhankelijk zijn, blijft het desondanks belangrijk dat testleiders bewust zijn van mogelijke culturele en/of contextuele invloeden op de resultaten.

De scoreformulieren voor de testleider zijn eveneens helder en gebruiksvriendelijk opgesteld voor zowel de basis- als de aanvullende analyses. Het profielblad geeft een duidelijke en visuele samenvatting van de resultaten weer. Het is mogelijk om de scores van een ouderversie te vergelijken met de scores van een leerkrachtversie van de ABAS-3 voor de leeftijdsgroepen 0-4 jaar en 5-17 jaar. Het scoreformulier dat je hiervoor kan gebruiken, is terug te vinden op de website van uitgeverij Hogrefe⁷. Testleiders kunnen dit formulier daar downloaden en invullen. Het kan mogelijk wenselijk zijn om dit in een nieuwe versie van de test toe te voegen aan het bestaande scoreformulier.

De kwaliteit van het testmateriaal kan als **uitstekend** beoordeeld worden voor zowel het nemen van *belangrijke als minder belangrijke beslissingen op individueel niveau*.

⁷ <https://www.hogrefe.com/nl/shop/abas-3-schaal-voor-adaptief-gedrag.html>

Vlaamse normen

Algemeen

In het kader van de ontwikkeling van de Schaal voor Adaptief gedrag (ABAS-3; Kreemers et. al., 2020) vond in de periode 2015 – 2016 in Vlaanderen een pilootonderzoek plaats gebaseerd op de gegevens van 222 kinderen en jongeren. Deze steekproef bestond vooral uit kinderen en jongeren uit de algemene populatie, aangevuld met een zeer heterogene klinische steekproef (i.e., verstandelijke beperking, autismespectrumstoornis, gedrags- en/of emotionele stoornis, visuele beperking, auditieve beperking, motorische beperking en/of een niet-aangeboren hersenletsel). De resultaten van het pilootonderzoek ondersteunen de internationale psychometrische eigenschappen van de ABAS-3 (Kreemers et. al., 2020).

In een volgende stap werd gestart met grootschalige normeringsonderzoeken in Vlaanderen en Nederland. Het huidige rapport beoordeelt in hoofdzaak het Vlaams normeringsonderzoek en de daaruit voortkomende Vlaamse normen. Voor de *ouderversies 0-4 jaar en 5-17 jaar* en de *leerkrachtversies 2-4 jaar en 5-17 jaar* zijn de vaardigheidsgebieden genormeerd aan de hand van **continue normering**. De domeinen en het totaal werden genormeerd volgens de **methode van cumulatieve proporties**. Deze laatstgenoemde methode werd eveneens toegepast voor het bepalen van de normscores van de vaardigheidsgebieden, domeinen en het totaal voor zowel de *zelfrapportage 18-80 jaar* en de *informantenrapportage 18-80 jaar* (Kreemers et. al., 2020).

Dataverzameling

De Vlaamse dataverzameling vond plaats vanaf het najaar van 2016 tot de zomer van 2019. Via Kind & Gezin werden de gegevens van niet-schoolgaande kinderen verzameld. Voor schoolgaande kinderen verliep dit vooral via de Centra voor Leerlingenbegeleiding (CLB). Men trachtte voor elk kind minimaal twee informanten, idealiter uit verschillende leefomgevingen van het kind, te betrekken. De data met betrekking tot kinderen en jongeren werden verzameld via de pen-en-papier versie van de vragenlijst. De volwassen deelnemers werden vooral via Panel Inzicht (internetpanel), seniorenverenigingen en het persoonlijk netwerk van de onderzoekers gerekruteerd. De klinische groepen bestonden uit deelnemers die via het Vlaams Agentschap voor Personen met een Handicap (VAPH), Centra voor Ontwikkelingsstoornissen (COS), Centra voor Ambulante Revalidatie (CAR), Expertisecentrum Autisme, Diagnosecentrum voor Autismespectrumstoornissen (DIASS), en doctoraatsonderzoek aan de KU Leuven werden gerekruteerd.

De normering voor de ouder- en leerkrachtversies verliep via de pen-en-papier versie van de ABAS-3. Er wordt weinig informatie verstrekt over de omstandigheden waarin de normeringsdata werden verzameld. Zo wordt er niet gespecificeerd of er een testleider aanwezig was bij het invullen van de ABAS-3 en of er een aangepaste manier van afname werd toegepast in bepaalde situaties (bv. een ouder die moeilijkheden heeft met lezen). Voor de zelf- en informantenrapportage werd gebruikt gemaakt van het internetpanel Panel Inzicht. Hier konden volwassenen een digitale versie van de ABAS-3 invullen tegen een kleine vergoeding. Voorafgaand zijn er selecties gebeurd op basis van leeftijd en geslacht. Er werd daarnaast ook rekening gehouden met opleidingsniveau, migratie-achtergrond en regio. Enkel volledig ingevulde instrumenten werden meegenomen in de data-analyse. Er werden eveneens verschillende kwaliteitscontroles uitgevoerd (i.e., zorgvuldigheid van het invullen van de vragenlijst, het geven van steeds dezelfde antwoorden, realistische invulsnelheid).

De finale versies van de ABAS-3 bestaan enkel in pen-en-papier versies. De digitale versie die door de volwassenen werd ingevuld tijdens het normeringsonderzoek is in de praktijk niet beschikbaar gesteld en wordt dus niet gebruikt. Volwassenen waarvoor de ABAS-3 ingezet wordt, zullen dus een pen-en-papier versie van de ABAS-3 invullen. Er wordt echter geen informatie gegeven over of dit al dan niet een impact heeft op de resultaten.

Vlaamse normeringssteekproeven

Per versie van de ABAS-3 werd een aparte Vlaamse steekproef onderzocht. De *ouderversie 0-4 jaar* is voor 1289 kinderen tussen 0-6 jaar ingevuld en de *leerkrachtversie 2-4 jaar* is ingevuld voor 360 kinderen tussen 0-6 jaar. Wat betreft de *ouder- en leerkrachtversie 5-17 jaar* bestonden de steekproeven respectievelijk uit 1494 en 706 deelnemers tussen 3 en 19 jaar oud. In totaal vulden 629 volwassenen de *zelfrapportage 18-80 jaar* in en vulden 417 volwassenen de *informantenrapportage 18-80 jaar* in.

Er werd telkens gestratificeerd naar leeftijd en geslacht. Voor de ouder- en leerkrachtversies werd eveneens gestratificeerd naar opleidingsniveau van de moeder en nationaliteit, terwijl voor de volwassenversies ook op basis van land van herkomst en regio werd gestratificeerd. Wat betreft leeftijd hanteerde men, afhankelijk van de versie van de test, een andere leeftijdsbreedte. Dit ligt in lijn met de ontwikkeling van adaptieve vaardigheden doorheen de levensloop. Het verband tussen leeftijd en adaptief gedrag wordt in het hoofdstuk over validiteit meer in detail besproken (zie verder, Samenhang met leeftijd onder Begripsvaliditeit). De gekozen leeftijdsbreedtes zijn passend voor de verschillende versies van de ABAS-3. De breedte van deze leeftijdscategorieën werd bepaald tijdens het construeren van de normen. Het is vervolgens van belang dat de overige stratificatievariabelen representatief zijn binnen elke normgroep/leeftijdsgroep (m.a.w. gestratificeerd per leeftijdsgroep; Van Gampelaere, De Ganck, & Schoupe, 2020).

De streefcijfers voor de steekproeven zijn gebaseerd op diverse bronnen, waaronder Statistics Belgium (Statbel, 2020), Statbel Census 2011, en Vlaams onderwijs in cijfers: Schooljaar 2014-2015 (Vlaams Departement Onderwijs en Vorming, 2015). Tijdens de dataverzameling werd gepoogd deze populatiecijfers zo goed mogelijk te benaderen. Tevens werd een wegingsprocedure uitgevoerd tijdens de data-analyse voor verschillende variabelen (i.e., geslacht, opleidingsniveau moeder/opleidingsniveau, nationaliteit/land van herkomst, regio) om een zo groot mogelijke overeenkomst te bekomen tussen de steekproef en de populatie. Hierbij werd maximaal met een factor twee gewogen zoals opgelegd door COTAN (Evers, Lucassen, Meijer, & Sijsma, 2010).

Na weging bestond de normeringssteekproef voor de *ouderversie 0-4 jaar* uit 1294 deelnemers en 358 deelnemers voor de *leerkrachtversie 2-4 jaar*. De normeringssteekproeven voor de *ouder- en leerkrachtversie 5-17 jaar* hadden respectievelijk 1681 en 694 deelnames. Wat betreft de *zelf- en informantenrapportage 18-80 jaar* werden er respectievelijk 630 en 415 volwassenen in de normeringssteekproef opgenomen.

Zoals reeds aangegeven werd voor deze versies, waar mogelijk, een breder leeftijdsbereik dan beoogd toegepast, zodat er voldoende datapunten waren om de geschatte waarde accuraat te schatten. Zo bestond de normeringssteekproef voor de *ouderversie 0-4* en de *leerkrachtversie 2-4 jaar* ook uit kinderen tussen 0 en 6 jaar oud. Voor de ouder- en leerkrachtversie 5-17 jaar werden kinderen tussen 3 en 19 jaar opgenomen. De grootte van de normeringssteekproef van de *ouderversie 0-4 jaar* wordt als goed tot uitstekend beoordeeld. Voor de *leerkrachtversie 2-4 jaar* wordt de steekproefgrootte als voldoende beoordeeld. De normeringssteekproeven van de *ouder- en leerkrachtversie 5-17 jaar* zijn groot genoeg, gezien het groot aantal categorieën en het minimum wordt voor elke groep bereikt. Het is echter wel zo dat de uiterste leeftijdsgroepen vaak minder groot zijn. De testauteurs komen hieraan tegemoet door ook observaties op te nemen buiten het beoogde leeftijdsbereik. Deze normeringssteekproeven worden als goed beoordeeld. De samengenomen normeringssteekproeven voor de *zelf- en informantenrapportage 18-80 jaar* worden voor zowel belangrijke beslissingen als voor minder belangrijke beslissingen als goed beoordeeld.

Op basis van de informatie die voor handen is, worden de steekproefgroottes algemeen als voldoende tot goed beoordeeld voor minder belangrijke beslissingen en als voldoende tot goed beoordeeld voor belangrijke beslissingen. Er wordt jammer genoeg geen vergelijking gemaakt tussen de steekproefgroottes op basis van klassieke of continue normering (het plotten van standaardfouten ten opzichte van leeftijd) om na te gaan of de steekproeven eveneens voldoende groot zijn in de uiterste leeftijdscategorieën. Een aantal leeftijdscategorieën zijn voor bepaalde versies van de ABAS-3 te klein. In de finale normtabellen worden deze samengenomen waardoor de groepen alsnog (net) groot genoeg zijn.

Representativiteit

Over het algemeen kan de representativiteit van de Vlaamse steekproef als voldoende tot goed beoordeeld worden. Twee belangrijke aandachtspunten zijn echter van toepassing. Ten eerste geldt voor alle versies van de ABAS-3 dat de normgroepen niet representatief zijn wat betreft migratie-achtergrond. Ten tweede is er sprake van een oververtegenwoordiging van hoogopgeleide moeders en volwassenen, en een ondervertegenwoordiging van laag opgeleide moeders en volwassenen. Ondanks de inspanningen die bij de dataverzameling geleverd zijn om representativiteit te bekomen voor migratieachtergrond en opleidingsniveau, is het niet gelukt om deze doelgroepen voldoende in het normeringsonderzoek te kunnen betrekken.

De normeringsgroepen bestaan uit verschillende leeftijdsgroepen waarbij rekening gehouden werd met geslacht. De aantallen per leeftijdsgroep en per geslacht voor en na weging worden in tabellen 16, 18 en 20 de handleiding weergegeven. Met uitzondering van de variabele geslacht worden geen tabellen weergegeven over de verdeling van de stratificatievariabelen per leeftijdsgroep. Het is bijgevolg niet mogelijk om na te gaan of de afzonderlijke leeftijdsgroepen binnen de steekproef een representatieve afspiegeling zijn van de populatie. Voor de variabelen geslacht en sociaaleconomische status (SES) – gedefinieerd aan de hand van het opleidingsniveau van de moeder of het opleidingsniveau van de volwassene – worden de (gewogen) steekproefpercentages en de vooropgestelde streefpercentages (i.e., verdeling in de populatie) weergegeven. Voor de variabelen regio en migratie-achtergrond worden deze streefwaarden en steekproefpercentages niet weergegeven.

Wat betreft geslacht wijkt de verdeling voor alle versies van de ABAS-3 na weging licht af van de verdeling in de algemene populatie. Deze afwijkingen zijn voor elke versie van de ABAS-3 niet significant. Zowel voor de ouder- en leerkrachtversies als voor de zelf- en informant rapportage van de ABAS-3 is er een duidelijke onderrepresentatie van personen met een lage SES. Nochtans tonen literatuuronderzoek en variantieanalyses (vooral in leeftijdsgroep 5-17 jaar) aan dat kinderen van ouders met een lage SES significant lager presteren op vlak van adaptieve vaardigheden (Kreemers et al., 2020). Ondanks verschillende inspanningen bleef deze groep ondervertegenwoordigd in de normeringssteekproeven (zowel kinderen en jongeren als volwassenen). Er werd getracht om een weging toe te passen. Echter, omwille van de grote ondervertegenwoordiging van deze groep in alle steekproeven, was een weging niet meer mogelijk (overstijgen van factor 2; Evers et al., 2010). Deze subgroep is met andere woorden geen representatieve weergave van de groep kinderen, jongeren, of volwassenen met een lage SES. Daarom werd voor de ouder-leerkracht en volwassen versies van de ABAS-3 besloten om enkel een weging toe te passen voor midden en hoog opleidingsniveau omwille van de overrepresentatie van het midden opleidingsniveau. Dit is dus een erg belangrijk aandachtspunt bij het interpreteren van de resultaten van elke versie van de ABAS-3.

Aangezien adaptieve vaardigheden sterk cultuur- en contextafhankelijk zijn, is het belangrijk om hier rekening mee te houden bij het in kaart brengen van adaptief gedrag. De variantieanalyses tonen echter – in tegenstelling tot andere onderzoeken – geen significante verschillen aan op basis van culturele achtergrond. Dit is zo voor alle versies van de ABAS-3. Mogelijk is dit te wijten aan het feit dat er vooral hoger opgeleide mensen met een andere culturele achtergrond deelnamen aan het normeringsonderzoek. Om deze reden werd niet naar representativiteit gestreefd, maar geven de auteurs wel aan alert te blijven voor mogelijke culturele vertekeningen. Ze stelden eveneens richtlijnen rond faire diagnostiek in specifieke doelgroepen beschikbaar in de handleiding.

Ten slotte werd voor de *zelf- en informant rapportage 18-80 jaar* ook rekening gehouden met de woonplaats van de respondent. Hier werd een onderverdeling gemaakt op basis van regio: west (West-Vlaanderen), midden (Oost-Vlaanderen, Antwerpen, Vlaams-Brabant, Brussels Hoofdstedelijk Gewest) en oost (Limburg). Men hield hierbij rekening met de verdeling in de populatie op basis van de gegevens van Statbel 2017. In de gewogen steekproeven resulteerde dit in een kleine afwijkingen, namelijk een ondervertegenwoordiging van de regio's oost en midden en een oververtegenwoordiging van de regio west. Deze afwijkingen verschillen echter niet significant van de verdeling in de populatie.

Normeringsmethode

Continue normering

Voor het ontwikkelen van de Vlaamse normen voor de **vaardigheidsgebieden** van de *ouderversies 0-4 jaar en 5-17 jaar en de leerkrachtversies 2-4 jaar en 5-17 jaar* werd gebruik gemaakt van continue normering. Aan de basis van continue normering ligt een regressieanalyse waarbij de testscore wordt gemodelleerd als functie van de leeftijd. Dit zorgt voor meer nauwkeurige en vloeiend oplopende normen waarbij discontinuïteit tussen leeftijdscategorieën vermeden wordt (Kreemers et. al., 2020). Deze methode laat daarnaast toe dat bij het bepalen van een normscore, in combinatie met een ruwe score en leeftijd, de informatie van de gehele steekproef wordt meegenomen. Concreet werd bij de normering van de vaardigheidsgebieden een regressieanalyse toegepast op de ruwe scores, waardoor een mogelijke vertekening ten gevolge van verschillende aantallen observaties per leeftijdscategorie vermeden werd (Kreemers et. al., 2020).

Om normen te kunnen construeren is het noodzakelijk dat er eerst aan bepaalde assumpties is voldaan (i.e., normaliteit, homogeniteit van variantie, collineariteit en curvilineariteit). Wat betreft *normaliteit*, bleek dat de data voor de meeste vaardigheidsgebieden normaal verdeeld zijn. De auteurs geven aan dat er beperkte schendingen waren, maar dat deze geen opmerkelijke impact hadden op de resultaten van de regressieanalyse (Kreemers et. al., 2020). Er worden echter geen data of resultaten weergegeven in de handleiding om dit te kunnen nagaan. Op vlak van *homogeniteit van de variantie* werden significante verschillen vastgesteld. Aangezien deze verschillen niet systematisch waren en niet telkens in dezelfde richting wezen, werd besloten om geen gevolg te geven aan deze verschillen (Kreemers et. al., 2020). De resultaten van deze tests en vergelijkingen werden echter eveneens niet weergegeven in de handleiding. Vervolgens werd ook de *curvilineariteit* nagegaan. Er werd een verband gevonden tussen leeftijd en adaptief gedrag met als gevolg dat voor alle versies de kwadratische term leeftijd werd opgenomen. Ten slotte werd de *(multi)collineariteit* onderzocht. Deze analyses toonden aan dat er zoals verwacht een grote correlatie was tussen leeftijd en de kwadratische term van leeftijd. Hiervoor werd een constante afgetrokken waardoor de variance-inflation factor (VIF) zakte naar een aanvaardbare waarde ($VIF < 10$); Kreemers et. al., 2020). Samengevat kan gesteld worden dat niet aan alle assumpties voldaan is. De auteurs geven echter ook theoretisch aanvaardbare oplossingen voor deze moeilijkheden. Aangezien de data ontbreken om dit verder te onderbouwen, kan niet worden bevestigd dat dit correct gebeurde.

Een volgende stap is het uitvoeren van regressieanalyses. De auteurs kozen ervoor om, conform de methode van Van Breukelen en Vlaeyen (2005), alle relevante voorspellers op te nemen in de regressieanalyse om vervolgens stapsgewijs de niet-significante voorspellers te elimineren uit de regressie. Uiteindelijk werd enkel leeftijd als voorspellende variabele behouden en werden geslacht, culturele achtergrond en SES niet opgenomen. Vervolgens werden ruwe scores getransformeerd naar normscores. Aangezien er niet gedifferentieerd kon worden voor de lage scores, werd besloten om voor Vlaanderen vanaf de leeftijd 0;3 jaar te normeren aangezien er vanaf dan enige differentiatie op te merken is. Eveneens merkte men een plafondeffect op de leeftijd van 5 jaar voor de *ouderversie 0-4 jaar* en de *leerkrachtversie 2-4 jaar*. Op basis van de regressieanalyses werd de breedte van de leeftijdscategorieën voor de verschillende versies bepaald. Bij zeer jonge kinderen wordt een leeftijdsrange van een maand gehanteerd, aangezien er een zeer snelle ontwikkeling in adaptief gedrag wordt opgemerkt. Vanaf 5 jaar is er een merkbare vertraging waardoor er in de groepen 5;0 tot 6;11-jarigen, 7;0 tot 11;11-jarigen en 12;0 tot 17;11-jarigen respectievelijk een leeftijdsrange van drie, zes en 12 maanden werd gehanteerd (Kreemers et. al., 2020). In de handleiding werd echter geen data of meer gedetailleerde informatie over de resultaten opgenomen.

Tot slot werden leeftijdsequivalenten bepaald voor de ruwe scores van de vaardigheidsgebieden. Hiervoor ging men voor elke leeftijdsgroep na welke ruwe score overeenkomt met een normscore van 10. Aangezien een bepaalde ruwe score vaak voor verschillende leeftijdsgroepen overeenkomt met een normscore van 10, wordt naast het leeftijdsequivalent ook het leeftijdsbereik weergegeven (Kreemers et. al., 2020). Een groot pluspunt van de ABAS-3 is dat er rekening gehouden wordt met bodem- en plafondeffecten in de laagste en hoogste leeftijdsgroepen.

Methode van cumulatieve proporties

Ouder- en leerkrachtversies 0-4 en 5-17 jaar

Voor de **domeinen en totaalscores** van de *ouder- en leerkrachtversies 0-4 jaar* en *5-17 jaar* werden de normen bepaald volgens de methode van cumulatieve proporties. Dit houdt in dat op basis van de geobserveerde cumulatieve proportie van de som van de normscores van de vaardigheidsgebieden, normscores worden berekend voor de domeinen en het totaal. Aangezien dit gebaseerd is op normscores van de vaardigheidsgebieden, zijn de leeftijdseffecten al opgenomen. Bijgevolg worden de normscores voor domeinen en het totaal per leeftijdsgroep weergegeven (Kreemers et. al., 2020).

Net zoals bij continue normering is het belangrijk dat aan een aantal assumpties (i.e., normaliteit, verschillen tussen leeftijdsgroepen en homogeniteit van de variantie) is voldaan. Deze analyses tonen echter aan dat de assumptie *normaliteit* niet voldaan is voor de uiterste leeftijdscategorieën. Andere opties werden geëxploreerd, maar niet toegepast waardoor men, zeker wat betreft normscores in de uiterste leeftijdscategorieën, heel voorzichtig moet zijn bij het interpreteren van de resultaten (Kreemers et. al., 2020).

De normscores voor de domeinen en het totaal werden vervolgens geconstrueerd op basis van de frequentietabel van de geobserveerde somscores in de normeringssteekproef, de cumulatieve proportie en de z-score die uiteindelijk werd omgezet naar een normscore (met gemiddelde 100 en standaarddeviatie 15). Het bereik van de normscores is gebaseerd op de laagste geobserveerde score. Omwille van bodemeffecten is het in de jongste groep niet mogelijk om te differentiëren naar beneden (0;3 – 0;11 jaar). Er wordt enkel aangegeven of de score onder de afkapwaarde van twee standaarddeviaties onder het gemiddelde ligt, omdat dit van belang is in het kader van het diagnosticeren van een verstandelijke beperking. Bij alle ouder- en leerkrachtversies geldt eveneens een bovengrens van twee standaarddeviaties boven het gemiddelde, aangezien een verdere differentiatie naar boven (statistisch) niet verantwoord is. Aan de uiteinden van de normtabellen zijn scores minder betrouwbaar dus zeer hoge en zeer lage scores normscores moeten altijd voorzichtig geïnterpreteerd en gerapporteerd worden (Kreemers et. al., 2020).

Zelf- en informantenrapportage 18-80 jaar

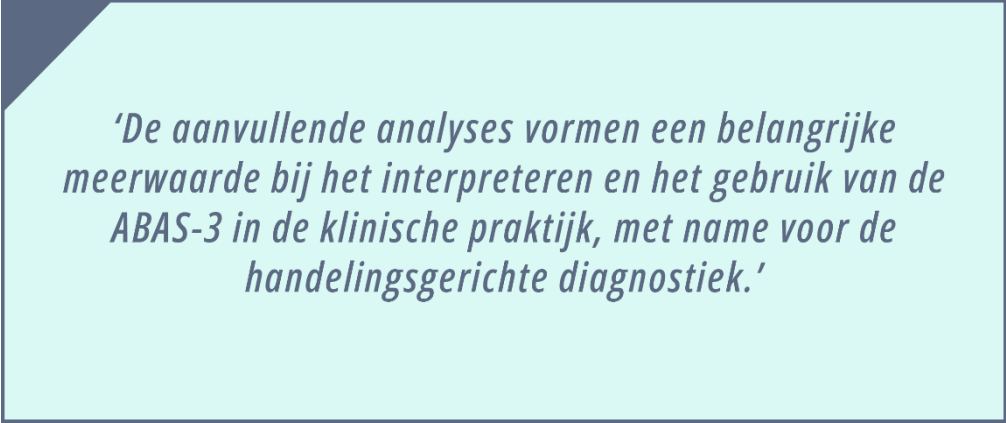
De methode van cumulatieve proporties werd eveneens toegepast om de normen te bepalen voor de **vaardigheidsgebieden, domeinen en totaalscores** van de *zelf- en informantenrapportage 18-80 jaar*.

Voorafgaand aan de normconstructie is het belangrijk om na te gaan of aan de assumpties voldaan is. Uit deze analyses bleek dat, voor zowel de zelf- als de informantenrapportage, geen enkel vaardigheidsgebied normaal verdeeld is. De afwijking van de normaalverdeling toont vooral aan dat er sterke plafondeffecten aanwezig zijn in de volwassenversies. De auteurs geven aan dat dit niet verwonderlijk is aangezien de ABAS-3 zo opgesteld is om voldoende naar beneden te kunnen differentiëren in het kader van de classificerende diagnostiek van een verstandelijke beperking. Theoretisch kan men echter verwachten dat adaptief gedrag ook bij volwassenen normaal verdeeld is. De metingen gebeuren op ordinaal niveau en de voorwaarde dat de rangorde gelijk blijft na normalisatie is voldaan. Op basis van deze informatie werd besloten om de geobserveerde verdelingen te normaliseren en vervolgens de methode van cumulatieve proportie toe te passen (Kreemers et. al., 2020).

Vervolgens werd nagegaan of de normgroep opgesplitst moest worden in verschillende leeftijdscategorieën en/of naargelang geslacht. Omdat er geen significante verbanden gevonden werden met leeftijd, werd besloten om deze versies klassiek te normeren zonder onderscheid tussen verschillende leeftijdscategorieën. De regressieanalyses met geslacht, toonden aan dat er geen aparte normen voor mannen en vrouwen nodig zijn (Kreemers et. al., 2020). De resultaten van deze analyses ontbreken in de handleiding.

De normscores voor de **vaardigheidsgebieden** werden geconstrueerd op basis van de geobserveerde frequenties, de cumulatieve proportie en de z-score die uiteindelijk werd omgezet naar een normscore (met gemiddelde 10 en standaarddeviatie 3). De maximale normscore komt in de praktijk echter niet voor omwille van de eerdergenoemde plafondeffecten. Om deze reden werd de hoogste mogelijke normscore bepaald door de geobserveerde frequenties (Kreemers et. al., 2020).

De normscores voor de **domeinen en totaalscores** werden op eenzelfde manier geconstrueerd als deze voor de ouder- en leerkrachtversies. Net als bij de vaardigheidsgebieden werd niet voldaan aan de assumptie normaliteit omwille van aanwezige plafondeffecten. Dezelfde argumenten gaan echter op om de niet normaal verdeelde domeinen en totalen toch te normaliseren. De normscores voor de domeinen en het totaal werden geconstrueerd op basis van de frequentietabel van de geobserveerde somscores in de normeringssteekproef, de cumulatieve proportie en de z-score die uiteindelijk werd omgezet naar een normscore (met gemiddelde 100 en standaarddeviatie 15; Kreemers et. al., 2020).



'De aanvullende analyses vormen een belangrijke meerwaarde bij het interpreteren en het gebruik van de ABAS-3 in de klinische praktijk, met name voor de handelingsgerichte diagnostiek.'

Aanvullende analyses

Naast resultaten voor de vaardigheidsgebieden, domeinen en totaalscores, kunnen voor elke versie van de ABAS-3 ook bijkomende analyses gedaan worden om nog beter zicht te krijgen op het adaptief profiel van een individu. Het feit dat dit mogelijk is, is een grote meerwaarde in het kader van handelingsgerichte diagnostiek en het opstellen van zorgdoelen.

Een eerste mogelijkheid is om na te gaan of er significante verschillen zijn tussen de normscores van twee domeinen. Ten tweede kan men ook nagaan of er een significant verschil is tussen twee vaardigheidsgebieden binnen een domein of binnen het totaal aan de hand van de spreiding van vaardigheidsgebieden. Als een verschil gevonden wordt binnen een domein, moet men de spreiding binnen elk domein vergelijken. Indien er geen verschil is, dan dient de spreiding binnen het totaal vergeleken te worden. Een derde optie is om een sterkte-zwakte analyse uit te voeren. Op basis van deze analyse kan men nagaan welke vaardigheidsgebieden een relatieve sterkte of zwakte zijn ten opzichte van de eigen gemiddelde score voor een domein of het totaal. De kritieke waarde geeft aan vanaf wanneer het verschil met de gemiddelde normscore als statistisch significant beschouwd kan worden. Ten slotte kan men ook verschillende informantversies vergelijken en nagaan of er een verschil is in het adaptief gedrag van een kind of jongere zoals gerapporteerd door een ouder of een leerkracht. Voor de volwassenversies kan geen vergelijking gemaakt worden tussen beide rapportageversies aangezien er geen data is over het adaptief functioneren van eenzelfde volwassen persoon. De kritieke waarde geeft aan vanaf wanneer het verschil tussen de normscores van twee vaardigheidsgebieden, twee domeinen of twee totalen als statistisch significant beschouwd kan worden. Voor alle aanvullende analyses geldt echter dat een statistisch significant verschil pas als klinisch relevant mag beschouwd worden, als het ook een uitzonderlijk verschil is. Om na te gaan of het een uitzonderlijk, en dus een klinisch relevant verschil is, moet men de *base rate* nagaan. De ABAS-3 vermeldt kritieke waarden voor *base rates* gelijk aan of lager dan 15% of 5% van de normatieve steekproef.

Deze analyses vormen een meerwaarde bij het interpreteren en het gebruik van de ABAS-3 in de klinische praktijk, met name voor de handelingsgerichte diagnostiek.

Sterktes

- De testontwikkelaars hebben veel inspanningen geleverd om de steekproeven zo groot en representatief mogelijk te maken. Waar dit onvoldoende gelukt is, zijn richtlijnen voor faire diagnostiek voorzien.
- De rapportering van het normeringsonderzoek is uitgebreid en transparant.
- De aanvullende analyses vormen een belangrijke meerwaarde bij het interpreteren en het gebruik van de ABAS-3 in de klinische praktijk, met name voor de handelingsgerichte diagnostiek.

Aandachtspunten

- De grootte van de normeringssteekproef van de leerkrachtversie 2-4 jaar wordt als te klein beoordeeld. Ook voor de zelf- en informant rapportage, met name de groep 66-80 jaar, wordt de steekproefgrootte als onvoldoende beoordeeld. Dit houdt in dat de normen onvoldoende accuraat zijn en de resultaten voorzichtig geïnterpreteerd moeten worden.
- Er werd veel aandacht besteed aan de stratificatie van de Vlaamse steekproef. Hierbij dient echter opgemerkt dat de normen niet representatief zijn voor personen met een lage SES of met een andere culturele achtergrond.
- Aan de assumptie van normaliteit werd niet voldaan voor verschillende versies van de ABAS-3. Voor de ouder- en leerkrachtversies is met name voorzichtigheid geboden bij het interpreteren van scores in de uiterste leeftijdsgroepen; alsook bij zeer lage of zeer hoge scores. In de volwassenversies is voorzichtigheid aangewezen bij hoge scores.
- Het is belangrijk om rekening te houden met bodem- en plafondeffecten. Deze worden niet meegenomen in aanvullende analyses en voorzichtigheid bij het interpreteren van deze resultaten is noodzakelijk.
- Voorzichtigheid bij het gebruik van leeftijdsequivalenten is aangewezen. Deze zijn statistisch minder goed onderbouwd en kunnen niet gebruikt worden om een vergelijking te maken tussen leeftijdsgenoten of vaardigheidsgebieden. Ze kunnen ingezet worden in het kader van handelingsgerichte diagnostiek (NIET classificerende diagnostiek) wanneer de normscores niet gebruikt kunnen worden (weinig differentiatie omwille van bodemeffecten; gebruik van een versie die niet overeenkomt met de kalenderleeftijd).

Ouderversie 0-4 jaar en leerkrachtversie 2-4 jaar

Voor het nemen van *minder belangrijke beslissingen op individueel niveau* kan de **Vlaamse normering** als **goed** worden beoordeeld voor de ouderversie en als **voldoende** voor de leerkrachtversie.

Voor het nemen van *belangrijke beslissingen op individueel niveau* kan de **Vlaamse normering** als **goed** worden beoordeeld voor de ouderversie en als **voldoende** voor de leerkrachtversie.

Ouder- en leerkrachtversie 5-17 jaar

Voor het nemen van *minder belangrijke beslissingen op individueel niveau* kan de **Vlaamse normering** als **goed** worden beoordeeld.

Voor het nemen van *belangrijke beslissingen op individueel niveau* wordt de **Vlaamse normering** als **voldoende** beoordeeld, mits er rekening gehouden wordt met de bovengenoemde aandachtspunten.

Zelf- en informant rapportage 18-80 jaar

Voor het nemen van *minder belangrijke beslissingen op individueel niveau* kan de **Vlaamse normering** als **goed** worden beoordeeld.

Voor het nemen van *belangrijke beslissingen op individueel niveau* wordt de **Vlaamse normering** als **voldoende** beoordeeld.

Betrouwbaarheid

Betrouwbaarheid verwijst naar de mate waarin scores vrij zijn van meetfoutvariantie. Betrouwbaarheidsonderzoek gaat met andere woorden na in welke mate toevallige meetfouten de test scores beïnvloeden. Er dient hierbij een onderscheid gemaakt te worden tussen verschillende types betrouwbaarheid (i.e., interne consistentie, test-hertestbetrouwbaarheid, interbeoordelaarsbetrouwbaarheid en paralleltestbetrouwbaarheid). Daarnaast kan de betrouwbaarheid en stabiliteit van een test ook worden onderbouwd door het beschrijven van standaardmeetfouten, betrouwbaarheidsintervallen en significantie van verschillen tussen scores. Bij de beoordeling dient verder rekening gehouden te worden met het doel van de test. Voor de domein- en totaalscores worden in het huidig adviesrapport de criteria voor belangrijke beslissingen gevolgd aangezien deze scores zowel in het kader van classificerende als van handelingsgerichte diagnostiek gebruikt worden. De criteria voor minder belangrijke beslissingen werden gevolgd voor de scores op de vaardigheidsgebieden omdat op basis van deze scores geen grote beslissingen worden genomen.

De **interne consistentie** geeft weer in welke mate items binnen een vaardigheidsgebied onderling samenhangen en dus hetzelfde construct meten. Dit werd gemeten aan de hand van Cronbachs alfa (α), een statistiek met een waarde tussen 0 en 1 die wordt berekend op basis van paarsgewijze covarianties tussen items. De European Association of Psychologists' Association (EPPA) geeft de volgende richtlijnen omtrent het interpreteren van deze statistiek: uitstekend ($\alpha \geq .90$), goed ($.80 \leq \alpha < .90$), voldoende ($.70 \leq \alpha < .80$), of onvoldoende ($\alpha < .70$). Voor alle versies van de ABAS-3 werd voor elk vaardigheidsgebied, domein en totaalscore een Cronbachs alfa berekend. Vervolgens werd op basis hiervan de standaardmeetfout berekend om zo telkens het 95%-betrouwbaarheidsinterval te bepalen. Hoe kleiner de meetfout, hoe kleiner het interval en hoe betrouwbaarder de behaalde resultaten zijn. Voor de ouder- en leerkrachtversies van de ABAS-3 zijn de betrouwbaarheidscoëfficiënten berekend per leeftijdscategorie en voor de gehele normeringssteekproef (ouderversie 0-4 jaar: 5 leeftijdscategorieën; ouderversie 5-17 jaar: 12 leeftijdscategorieën; leerkrachtversie 2-4 jaar: 2 leeftijdscategorieën; leerkrachtversie 5-17 jaar: 6 leeftijdscategorieën). Voor de zelf- en informant rapportage is dit telkens berekend voor 3 leeftijdscategorieën (18-29 jaar; 30-65 jaar; 66-80 jaar) en de gehele normeringssteekproef.

In het Vlaamse betrouwbaarheidsonderzoek kwamen de steekproeven overeen met de beoogde testdeelnemers en met uitzondering van zeven groepen, waren alle steekproeven in de psychometrische studies (voldoende) groot. Algemeen kan gesteld worden dat de betrouwbaarheidscoëfficiënten voor de domein- en totaalschalen van de verschillende versies goed tot uitstekend zijn.

- *Ouderversie 0-4 jaar*: De domeinen en het totaal hebben een uitstekende betrouwbaarheidscoëfficiënt ($\alpha \geq .90$). Uitzonderingen zijn de domeinscores voor de leeftijdscategorie 0;3-0;11 jaar en het domein 'Sociaal' in de leeftijdscategorieën 3;0 - 3;11 jaar en 4;0 - 4;11 jaar. Deze resultaten worden als voldoende tot goed beschouwd ($.70 \leq \alpha < .90$). Wat betreft de vaardigheidsgebieden liggen de gemiddelde betrouwbaarheidscoëfficiënten tussen de .81 en .90, wat als goed beschouwd wordt, met uitzondering van 'Zelfverzorging' ($\alpha = .77$; voldoende).
- *Leerkrachtversie 2-4 jaar*: Met uitzondering van de betrouwbaarheidscoëfficiënt van het domein 'Praktisch' bij kinderen van 4;0 - 4;11 jaar ($\alpha = .89$, goed), wordt de interne consistentie van de domeinscores en het totaal als uitstekend beoordeeld ($\alpha \geq .90$). Voor de meeste vaardigheidsgebieden geldt een goede tot uitstekende beoordeling ($\alpha \geq .80$), met uitzondering van de schalen 'Gezondheid en Veiligheid' ($\alpha = .78$) en 'Zelfverzorging' ($\alpha = .75$) in de leeftijdscategorie 4;0 - 4;11 jaar die als voldoende beoordeeld worden.
- *Ouderversie 5-17 jaar*: De domeinen en het totaal hebben een uitstekende betrouwbaarheidscoëfficiënt ($\alpha \geq .90$) in alle leeftijdscategorieën. Voor de vaardigheidsgebieden bedraagt de Cronbachs alfa in de meeste gevallen .80 of meer. Deze kunnen als goed beschouwd worden. Een aantal vaardigheidsgebieden hebben een betrouwbaarheidscoëfficiënt die als voldoende beschouwd kan worden ($.70 \leq \alpha < .80$). Dit is het geval voor 'Communicatie' in de groep 7-jarigen; 'Gezondheid en Veiligheid' in de groep 10-jarigen en 12-jarigen en het vaardigheidsgebied 'Zelfverzorging' voor de 9-, 12-, 14-, 15-, 16- en 17-jarigen.

- *Leerkrachtversie 5-17 jaar:* Voor de domeinen en de totaalscores ligt de Cronbachs alfa voor elke leeftijdscategorie boven de .90. Deze resultaten kunnen als uitstekend beschouwd worden. Wat betreft de vaardigheidsgebieden bedraagt de betrouwbaarheidscoëfficiënt meestal .80 of meer, wat als goed tot uitstekend beoordeeld wordt. De schaal 'Gezondheid en Veiligheid' in de leeftijdscategorieën 5-6-jarigen en 7-8-jarigen en de schaal 'Functionele Schoolse Vaardigheden' in de categorie 16-17-jarigen hebben een Cronbachs alfa die tussen de .70 en .80 ligt en worden dus als voldoende beoordeeld.
- *Zelfrapportage 18-80 jaar:* De betrouwbaarheidscoëfficiënten van de domeinen en de totaalscores liggen in elke leeftijdsgroep boven .90, wat een uitstekende beoordeling krijgt. Met uitzondering van het domein 'Praktisch (zonder werk)' in de leeftijdscategorie 66-80 jaar waarvoor $\alpha = .88$ en dus een goede beoordeling krijgt. De vaardigheidsgebieden hebben meestal een betrouwbaarheidscoëfficiënt boven de .83, deze worden als goed of uitstekend (indien $\alpha \geq .90$) beoordeeld. In de leeftijdscategorie 66-80 jaar bevindt Cronbachs alfa zich tussen de .75 en .79 voor de schalen 'Maatschappelijke vaardigheden', 'Gezondheid en Veiligheid' en 'Zelfverzorging'. Deze worden als voldoende beoordeeld.
- *Informantenrapportage 18-80 jaar:* De interne consistentie van alle domeinen en totaalscores wordt als uitstekend beoordeeld ($\alpha \geq .90$). Ook wat betreft de vaardigheidsgebieden bevindt de betrouwbaarheidscoëfficiënt zich meestal boven de .90, waardoor deze ook als uitstekend beoordeeld worden, met uitzondering van 'Werk' in de leeftijdscategorie 18-29 jaar ($\alpha = .82$) die als goed beoordeeld wordt.

Samengevat zijn de betrouwbaarheidscoëfficiënten voor de domein- en totaalscores voor elke versie van de ABAS-3 te beoordelen als goed ($\geq .80$) tot uitstekend ($\geq .90$) volgens de criteria voor het nemen van belangrijke beslissingen, met uitzondering van de domeinscores voor de leeftijdscategorie 0;3-0;11 jaar; het domein 'Sociaal' in de leeftijdscategorieën 3;0 - 3;11 jaar en 4;0 - 4;11 jaar en de betrouwbaarheidscoëfficiënt van het domein 'Praktisch' bij kinderen van 4;0 - 4;11 jaar. Dit wijst erop dat in de jongste groep kinderen enige voorzichtigheid aangewezen is bij het interpreteren van de domeinscores, alsook bij het interpreteren van het domein 'Sociaal' bij kinderen tussen 3;0 en 4;11 jaar en het domein 'Praktisch' bij kinderen van 4;0 - 4;11 jaar. Wat betreft de betrouwbaarheid van de vaardigheidsgebieden, is deze meestal als voldoende tot goed te beoordelen. Op basis van deze schalen worden vooral minder belangrijke beslissingen genomen.

Verder worden ook **standaardmeetfouten (SEM) en betrouwbaarheidsintervallen** beschreven om de betrouwbaarheid van de ABAS-3 verder te onderbouwen. Een SEM is omgekeerd evenredig aan de betrouwbaarheid. Hoe kleiner de SEM, hoe betrouwbaarder de bekomen resultaten. Op basis hiervan worden 95%-betrouwbaarheidsintervallen berekend. Deze worden voor de domeinen en het totaal voor elke versie van de ABAS-3 weergegeven in de normtabellen (Kreemers et. al., 2020).

De **test-hertestbetrouwbaarheid** representeert in welke mate behaalde scores stabiel blijven over een relatief kort tijdsinterval. In de Vlaamse steekproeven werd dit onderzocht door respondenten twee keer na elkaar met een tijdsinterval van twee weken eenzelfde versie van de ABAS-3 te laten invullen. De auteurs geven aan dat men koos voor deze korte periode aangezien onderzoek aantoonde dat adaptief gedrag zeer snel kan ontwikkelen, vooral bij jonge kinderen (Kreemers, et. al., 2020). Hierbij werd echter weinig rekening gehouden met de ontwikkeling van adaptieve vaardigheden bij oudere kinderen, jongeren en volwassenen.

De gemiddelde scores en standaarddeviaties van de totalen, domeinen en vaardigheidsgebieden worden voor beide meetmomenten weergegeven. Daarnaast worden de intraklassecorrelatiecoëfficiënt (ICC), een maat van consistentie tussen twee meetmomenten, en de effectgroottes (Cohen's d) gerapporteerd. De ICC wordt als volgt geïnterpreteerd: een ICC dat groter dan of gelijk is aan .80 wordt als goed beoordeeld, een ICC tussen .60 en .80 wordt als voldoende beoordeeld en een ICC-waarde kleiner dan .60 wordt als onvoldoende beoordeeld (Kreemers et. al., 2020; Hallgren, 2012). Er wordt echter niet gerapporteerd welk type ICC werd berekend (voor een overzicht van verschillende types ICC's, zie Hallgren, 2012). Wat betreft de effectgroottes wordt vanaf .20 van een klein effect gesproken, vanaf .50 van een middelgroot effect en vanaf .80 spreken we van een groot effect (Cohen, 1988; Kreemers et. al., 2020). De effectgroottes worden echter niet besproken in de handleiding.

Voor Vlaanderen werd de test-hertestbetrouwbaarheid onderzocht voor beide ouderversies en leerkrachtversies.

- Wat betreft de *ouderversie 0-4 jaar* en de *ouderversie 5-17 jaar* werd de test-hertestbetrouwbaarheid onderzocht in steekproeven die overeenkomen met de beoogde testdeelnemers en waren de steekproeven groot genoeg (respectievelijk $N = 414$ en $N = 415$). Voor beide versies wordt voor zowel het totaal, de domeinen als de vaardigheidsgebieden telkens een ICC van groter dan .80 gevonden. Dit wijst op een goede consistentie tussen de twee meetmomenten. De effectgroottes geven aan dat er voor de domeinen, de totaalscore en bepaalde vaardigheidsgebieden een klein effect is. Dit betekent dat er mogelijk toch enig verschil is tussen beide meetmomenten. Algemeen kan gesteld worden dat beide versies een goede test-hertestbetrouwbaarheid hebben.
- De test-hertestbetrouwbaarheid van de *leerkrachtversie 2-4 jaar* en *leerkrachtversie 5-17 jaar* werd onderzocht in erg kleine steekproeven, respectievelijk $N = 28$ en $N = 34$. Om deze reden moeten deze resultaten zeer voorzichtig geïnterpreteerd worden. Zeker wat betreft de *leerkrachtversie 5-17 jaar* aangezien dit een leeftijdsbereik van 12 jaar bestrijkt, kan deze steekproef niet representatief zijn. Wat betreft de *leerkrachtversie van 2-4 jaar* wordt voor de domeinen en totaalscores een goede ICC van boven de .85 gevonden, met uitzondering van het domein 'Praktisch' waarvoor de ICC .69 bedraagt, wat als voldoende beoordeeld wordt. De ICC van de vaardigheidsgebieden ligt tussen .66 en .95, wat als voldoende tot goed beoordeeld kan worden. De effectgroottes zijn erg laag, wat betekent dat er geen verschil lijkt te zijn tussen beide meetmomenten. Beide maten duiden in de richting van een goede consistentie tussen beide meetmomenten. Voor de *leerkrachtversie 5-17 jaar* worden voor de domeinen en het totaal telkens een ICC van groter dan .80 gevonden. De ICC van de vaardigheidsgebieden bevindt zich telkens boven de .77, wat als voldoende tot goed kan beoordeeld worden. De effectgroottes zijn vrij groot wat aangeeft dat er wel een verschil is tussen beide meetmomenten. De ICC en de effectgroottes spreken elkaar tegen. De auteurs bieden hiervoor een mogelijke verklaring, namelijk dat leerkrachten systematisch een iets hogere score geven op de hertest. Hierdoor worden verschillen gevonden tussen de eerste en tweede meting waardoor de ICC-waarden hoog liggen. De resultaten van de test-hertestbetrouwbaarheid van deze versies zijn echter weinig betrouwbaar omwille van de kleine en weinig representatieve steekproef. Meer onderzoek hieromtrent is aangewezen.

Over de test-hertestbetrouwbaarheid van de zelfrapportage 18-80 jaar en de informant rapportage 18-80 jaar in Vlaanderen is er geen informatie voorhanden aangezien dit enkel werd onderzocht in een Nederlandse steekproef. Verder onderzoek is wenselijk.

In de Nederlandse steekproeven (beide ouderversies, beide versies voor volwassenen) werd een tijdsinterval van vier weken aangehouden. Opvallend is dat de ICC-waarden lager liggen dan in het Vlaams onderzoek, ook de effectgroottes zijn duidelijk verschillend. De auteurs geven hier geen verklaring voor in de handleiding. Verder onderzoek is aangewezen.

De **interbeoordeelaarsbetrouwbaarheid** geeft weer in welke mate er overeenstemming is tussen twee onafhankelijke beoordeelaars die eenzelfde informant versie invullen over eenzelfde persoon (Kreemers, et al., 2020; Lord & Novic, 1968). De interbeoordeelaars-betrouwbaarheid werd enkel onderzocht voor de *ouderversie 0-4 jaar* en de *ouderversie 5-17 jaar* in Vlaamse steekproeven. Er wordt in de handleiding geen informatie gegeven over de verschillende beoordeelaars en of de scoring van beide versies door eenzelfde persoon uitgevoerd werd. Net zoals voor de test-hertestbetrouwbaarheid worden de gemiddelde scores en standaarddeviaties voor de totaalscore, domeinen en vaardigheidsgebieden gerapporteerd. Daarnaast worden eveneens de intraklassecorrelatiecoëfficiënt (ICC; weerspiegelt de overeenkomst tussen de twee beoordeelaars) en de effectgroottes (cohen's d) gerapporteerd. Beide maten worden op dezelfde manier geïnterpreteerd als bij de test-hertestbetrouwbaarheid. Er wordt echter niet gerapporteerd welk type ICC werd berekend (voor een overzicht van verschillende types ICC's, zie Hallgren, 2012).

- *Ouderversie 0-4 jaar*: De interbeoordeelaarsbetrouwbaarheid werd onderzocht in een voldoende grote Vlaamse steekproef ($N = 127$). Alle ICC-waarden liggen ruim boven .80, wat betekent dat deze als goed beoordeeld kunnen worden. De meeste effectgroottes liggen onder de .20 (gemiddelde vaardigheidsgebieden: $d = .11$; gemiddelde domeinen: $d = .17$ en totaal: $d = .18$). Deze resultaten geven aan dat er weinig verschillen zijn tussen de twee beoordeelaars en er dus sprake is van een goede interbeoordeelaarsbetrouwbaarheid voor de ouderversie 0-4 jaar.
- *Ouderversie 5-17 jaar*: De Vlaamse steekproef waarin de interbeoordeelaarsbetrouwbaarheid werd onderzocht was voldoende groot ($N = 138$). Aangezien deze steekproef een breed leeftijdsbereik weerspiegelt, rijst de vraag echter of deze steekproef voldoende representatief is. Met uitzondering van 'Spel en Vrije tijd' (ICC = .76; voldoende), bevindt de ICC zich voor alle vaardigheidsgebieden, domeinen en de totaalscore boven .80, wat telkens als goed beoordeeld kan worden. De meeste effectgroottes liggen ook hier onder .20. Deze resultaten duiden op een goede interbeoordeelaarsbetrouwbaarheid voor de ouderversie 5-17 jaar.

Wat betreft de ouderversie 0-4 jaar en de ouderversie 5-17 jaar kan dus gesteld worden dat er weinig tot geen verschillen zijn als het instrument door een andere beoordelaar binnen eenzelfde context en over eenzelfde persoon wordt ingevuld. De invloed van de beoordelaar is met andere woorden minimaal. Over de andere versies van de ABAS-3 kan geen uitspraak omtrent de interbeoordeelaarsbetrouwbaarheid gedaan worden aangezien dit niet onderzocht werd. Verder onderzoek is wenselijk.

Paralleltestbetrouwbaarheid verwijst naar de overeenstemming in testcores van twee verschillende informantversies die over eenzelfde kind zijn ingevuld (Kreemers et. al., 2020; Lord & Novic, 1968). Dit werd onderzocht door een ouderversie (0-4 jaar of 5-17 jaar) en een leerkrachtversie (2-4 jaar of 5-17 jaar) van de ABAS-3 te laten invullen over eenzelfde kind. De paralleltestbetrouwbaarheid werd nog niet onderzocht voor volwassenen (zelf- en informant rapportage 18-80 jaar). Er wordt naar gestreefd om de ABAS-3 te laten invullen door verschillende informanten. Aangezien adaptief functioneren cultuur- en contextafhankelijk is, wordt verwacht dat de resultaten tussen ouder- en leerkrachtversies verschillen. Afhankelijk van de context kunnen er immers andere verwachtingen zijn op vlak van adaptief gedrag (Kreemers et. al., 2020). De gemiddeldes en standaarddeviaties voor de vaardigheidsgebieden, domeinen en het totaal worden weergegeven. De intraklassecorrelatiecoëfficiënt (ICC; mate van overeenstemming tussen beoordeelaars van beide informantversies) en effectgrootte (cohen's d) worden eveneens gerapporteerd (Kreemers et. al., 2020). Zoals eerder aangegeven wordt nergens vermeld welk type ICC werd berekend. Deze maten (i.e., ICC en cohen's d) worden op dezelfde manier geïnterpreteerd als bij de test-hertestbetrouwbaarheid en interbeoordeelaarsbetrouwbaarheid. In de handleiding wordt onvoldoende informatie gegeven over of voldaan is aan de aannames van parallelisme.

- Voor de *ouderversie 0-4 jaar* en *leerkrachtversie 2-4 jaar* werd de paralleltestbetrouwbaarheid onderzocht in een steekproef die overeenkomt met de beoogde testdeelnemers en die voldoende groot is ($N = 152$). De effectgroottes liggen voor het totaal, alle domeinen en de meeste vaardigheidsgebieden onder .20, wat wijst op weinig verschil in testcores tussen beide versies. Voor de vaardigheidsgebieden 'Zelfverzorging' ($d = .80$), 'Huishoudelijke/schoolse activiteiten' ($d = .26$) en 'Motoriek' ($d = .30$), wordt een grotere effectgrootte gevonden, wat betekent dat er op deze vaardigheidsgebieden een groter verschil is tussen wat ouders en leerkrachten rapporteren. Met uitzondering van het vaardigheidsgebied 'Functionele Voorschoolse Vaardigheden' (ICC = .69), zijn alle intraklassecorrelatiecoëfficiënten kleiner dan .60 waardoor deze allen als onvoldoende beoordeeld worden. Dit betekent dat er weinig overeenstemming is tussen de testcores op beide versies van de ABAS-3 ingevuld door verschillende beoordeelaars over eenzelfde kind.
- De paralleltestbetrouwbaarheid van de *ouder- en leerkrachtversie 5-17 jaar* werd eveneens onderzocht in een representatieve en grote steekproef ($N = 463$). Alle effectgroottes zijn kleiner dan .20. Met uitzondering van het vaardigheidsgebied 'Functionele schoolse vaardigheden' (ICC = .76) en het domein 'Conceptueel' (ICC = .63), zijn alle intraklassecorrelatiecoëfficiënten kleiner dan .60, wat als onvoldoende beoordeeld wordt. Dit betekent dat er weinig overeenstemming is tussen de testcores op beide versies van de ABAS-3 ingevuld door verschillende beoordeelaars over eenzelfde kind.

We kunnen concluderen dat er een lage paralleltestbetrouwbaarheid is voor deze vier versies van de ABAS-3. De auteurs geven aan dat dit niet verrassend is aangezien adaptief gedrag cultuur- en contextafhankelijk is.

Sterktes

- De steekproeven opgenomen in het Vlaams betrouwbaarheidsonderzoek zijn, met een aantal uitzonderingen (zie hieronder), voldoende groot.
- De interne consistentie voor de domein- en totaalscores van alle versies zijn als goed tot uitstekend te beoordelen. Dit betekent dat de items binnen een vaardigheidsgebied onderling goed samenhangen en dus hetzelfde construct meten.

Aandachtspunten

- De **interne consistentie** bij kinderen jonger dan een jaar voor de domein- en totaalscore varieert tussen voldoende tot uitstekend; voor de vaardigheidsscores ligt dit iets lager. Deze scores moeten met voldoende voorzichtigheid geïnterpreteerd worden.
- De **test-hertestbetrouwbaarheid** werd voor Vlaanderen niet bij volwassenen onderzocht. De steekproeven waarin dit onderzocht werd voor de leerkrachtversies 2-4 jaar en 5-17 jaar zijn erg klein en weinig representatief waardoor deze resultaten met voorzichtigheid geïnterpreteerd moeten worden.
- Er is geen informatie over de **interbeoordeelaarsbetrouwbaarheid** van beide leerkrachtversies of de informantenrapportage 18-80 jaar.
- De **paralleltestbetrouwbaarheid** is niet onderzocht voor volwassenen. Er is een lage paralleltestbetrouwbaarheid voor de ouderversie 0-4 jaar en leerkrachtversie 2-4 jaar, en de ouder- en leerkrachtversie 5-17 jaar. Dit wijst vermoedelijk op de cultuur- en contextafhankelijkheid van de test en toont aan dat context steeds moet worden meegenomen in de interpretatie van de resultaten.

De **betrouwbaarheid** van de ouderversie 0-4 jaar en de leerkrachtversie 2-4 jaar wordt als **goed** beoordeeld voor zowel het nemen van *belangrijke* als *minder belangrijke beslissingen op individueel niveau*. Bovenstaande aandachtspunten zijn van toepassing.

De **betrouwbaarheid** van de ouder- en leerkrachtversie 5-17 jaar wordt als **uitstekend** beoordeeld voor het nemen van *minder belangrijke beslissingen*, en als **goed** beoordeeld voor het nemen van *belangrijke beslissingen op individueel niveau*. Bovenstaande aandachtspunten zijn van toepassing.

Wat betreft de **betrouwbaarheid** van de zelf- en informantenrapportage 18-80 jaar wordt dit als **voldoende** beoordeeld voor het nemen van *minder belangrijke beslissingen*, en als **onvoldoende** beoordeeld voor het nemen van *belangrijke beslissingen op individueel niveau*. Bovenstaande aandachtspunten zijn van toepassing.

Validiteit

De validiteit van een test verwijst naar de mate waarin de test aan zijn doel beantwoordt. Er bestaan verschillende types of vormen van validiteit, die elk een bijdrage kunnen leveren tot het ondersteunen van de algemene validiteit van de test. In deze beoordeling volgen we de klassieke driedeling; inhoudsvaliditeit, begripsvaliditeit en criteriumvaliditeit.

Inhoudsvaliditeit

Inhoudsvaliditeit betreft de mate waarin de inhoud van de test een goede weergave is van het te meten construct, in dit geval adaptieve vaardigheden. De inhoudsvaliditeit van de ABAS-3 wordt als voldoende beoordeeld. Ondersteuning hiervoor wordt voornamelijk geboden aan de hand van verwijzingen naar de literatuur. Op basis van de omschrijving en het theoretisch kader lijkt deze vragenlijst inhoudelijk adaptieve vaardigheden te capteren. Er is echter weinig informatie over hoe de items en de domeinen geselecteerd werden en hoe de oorspronkelijke vragenlijst tot stand is gekomen. Er is meer opzoekwerk nodig om deze informatie te verkrijgen (cfr. Uitgangspunten van de test).

Begripsvaliditeit

Om na te gaan of de ABAS-3 daadwerkelijk meet wat hij bedoelt te meten, wordt de begripsvaliditeit onderzocht. Door het beschrijven van een confirmatorische factoranalyse, de interne structuur met analyse van intercorrelaties, verschillen tussen relevante groepen, en relaties met andere testen (convergente en divergente validiteit) biedt de handleiding hiervoor evidentie.

Er werd een **confirmatorische factoranalyse (CFA)** uitgevoerd voor elke versie van de ABAS-3. Deze analyses gaan na of het vooropgestelde model bevestigd wordt door de data. Op basis van de Vlaamse normeringssteekproeven is er besloten om de items per vaardigheidsgebied in drie parcels te clusteren.

Twee modellen werden telkens getoetst, namelijk het één-factor-model (alle vaardigheidsgebieden laden op het totaal) en het drie-factoren-model (de vaardigheidsgebieden laden op de domeinen 'Conceptueel', 'Sociaal' en 'Praktisch'). Voor de *ouderversie 0-4 jaar* en de *leerkrachtversie 2-4 jaar* werd ook een hogere-orde-model getoetst (de vaardigheidsgebieden laden op de drie domeinen die op hun beurt laden op het totaal, cfr. de structuur van de ABAS-3). In de handleiding wordt omschreven waarom het hogere-orde-model enkel voor deze twee versies werd getoetst. Naast de χ^2 -toets werden telkens vier fitindices berekend: Comparative Fit Index (CFI), Tucker-Lewis Fit Index (TLI), Root Mean Square Error of Approximation (RMSEA) en de Standardized Root Mean Square Residual (SRMR; Byrne, 2012, Kreemers, et. al., 2020). Hieronder volgt een bespreking van de factoranalyses voor elke versie van de ABAS-3.

- *Ouderversie 0-4 jaar, leerkrachtversie 2-4 jaar* en de *ouderversie 5-17 jaar*: Er kan voor de getoetste modellen gesteld worden dat een voldoende tot goede fit gevonden wordt voor drie van de vier fitindices. De RMSEA-waarde is te hoog om van een goede fit te kunnen spreken.
- *Leerkrachtversie 5-17 jaar*: Verschillende fitindices wijzen op een onvoldoende fit voor beide modellen. Om hieraan tegemoet te komen, werd een correlatie toegelaten tussen de vaardigheidsgebieden 'Maatschappelijke vaardigheden' en 'Functionele schoolse vaardigheden' waardoor drie van de vier fitindices uiteindelijk wijzen op een goede fit (CFI, TLI en SRMR). Deze correlatie is theoretisch te verantwoorden en wordt besproken in de handleiding.
- Ook voor de *zelf- en informant rapportage 18-80 jaar* wordt initieel een onvoldoende fit gevonden voor beide getoetste modellen.
 - o Voor de *zelfrapportage 18-80 jaar* werd een correlatie toegelaten tussen de vaardigheidsgebieden 'Communicatie' en 'Maatschappelijke vaardigheden'. Aangezien deze beiden tot het domein 'Conceptueel' behoren en inhoudelijk overlappen, is deze correlatie theoretisch te verantwoorden. Dit resulteert in een aanzienlijk verbeterde fit, met uitzondering dat de waarde van RSMEA te hoog blijft.

- Voor de *informantenrapportage 18-80 jaar* werd een correlatie toegelaten tussen de vaardigheidsgebieden 'Maatschappelijke vaardigheden' en 'Functionele schoolse vaardigheden' waardoor de fit opnieuw opmerkelijk verbeterde, met uitzondering van de RMSEA die een te hoge waarde heeft om van een goede fit te kunnen spreken.
- Het één-factor-model en het drie-factoren-model werden eveneens getoetst in een *klinische groep* voor de *ouderversie 5-17 jaar*. Er worden opnieuw gelijkaardige resultaten gevonden. Er is namelijk een voldoende tot goede fit gevonden voor drie van de vier indices (i.e., CFI, TLI, SRMR).
- Ten slotte werden beide modellen getoetst in een *klinische groep* voor de *ouderversie 0-4 jaar*. Deze resultaten zijn echter niet betrouwbaar aangezien dit in een te kleine steekproef werd onderzocht.

Over het algemeen kan gesteld worden dat de modellen, mits het toelaten van bepaalde correlaties waarvoor een theoretische rationale aanwezig is, de data voldoende goed lijken te fitten voor de indices CFI, TLI, SRMR. De waarde van de RMSEA is voor alle informantversies te hoog om te kunnen spreken van een goede fit.

Een aantal bedenkingen worden echter best in acht genomen:

- Ten eerste valt het op dat de fitindices heel weinig verschillen tussen het één-factor-model en het drie-factoren-model. Het hogere-orde-model dat getoetst wordt in de versies voor de jongste doelgroep, verschilt helemaal niet van het één-factor-model en het drie-factoren-model voor de ouderversie 0-4 jaar. Wat betreft de leerkrachtversie 2-4 jaar is er amper verschil tussen het één-factor-model en het hogere-orde-model, en zijn er slechts kleine verschillen met het drie-factoren-model. In de resultaatbespreking wordt hier weinig op ingegaan, ongeacht dat deze resultaten de structuur van de test (cfr. opdeling in drie domeinen) weinig verder onderbouwen.
- Ten tweede kan de vraag gesteld worden waarom het hogere-orde-model niet voor elke versie werd onderzocht (in plaats van het drie-factoren-model) aangezien dit de structuur van elke versie van de test is.
- Ten derde wordt in de handleiding niet vermeld op welke manier bepaald werd welke correlaties toegelaten worden. Er wordt een theoretische rationale gekoppeld aan waarom een bepaalde correlatie te verantwoorden is. Echter, de auteurs vermelden niet hoe men de keuze maakte om juist die specifieke vaardigheidsgebieden met elkaar te correleren. Er zou immers voor de meeste correlaties wel een post-hoc verklaring gegeven kunnen worden.
- Er wordt geen informatie voorzien over of de factorstructuur invariant is over groepen heen en/of de test vrij is van itembias. Dit wordt gezien als een beperking aangezien er mogelijk een andere opbouw zou kunnen zijn rekening houdend met sociaal-economische status, culturele achtergrond of geslacht.

De **interne structuur** van de ABAS-3 wordt besproken aan de hand van een beschrijving van intercorrelaties tussen vaardigheidsgebieden, domeinen en het totaal. Per versie van de test werden alle correlaties berekend en gerapporteerd voor de gehele leeftijdsrange van een bepaalde informantversie. Er wordt geen informatie verstrekt over de verbanden tussen items en vaardigheidsgebieden. Volgende richtlijnen worden gebruikt voor het interpreteren van de correlatiecoëfficiënten (r): $r \geq .50$ weerspiegelt een sterke samenhang, wanneer r tussen $.30$ en $.50$ ligt, spreekt men van een matige samenhang en wanneer $r < .30$ is er een zwakke samenhang (Cohen, 1988; Kreemers et. al., 2020).

Op basis van de structuur van de test, wordt verwacht dat de vaardigheidsgebieden van eenzelfde domein onderling sterker correleren dan vaardigheidsgebieden behorend tot een verschillend domein. De resultaten impliceren echter dat er geen duidelijk onderscheid is tussen de mate waarin vaardigheidsgebieden samenhangen die al dan niet tot eenzelfde domein behoren. Met andere woorden, vaardigheidsgebieden die tot eenzelfde domein behoren correleren niet sterker met elkaar dan vaardigheidsgebieden die niet tot eenzelfde domein behoren. Eveneens zijn er omgekeerd ook geen zwakkere correlaties tussen vaardigheidsgebieden die tot een verschillend domein behoren. Aangezien de correlaties tussen alle schalen heel hoog zijn en er geen distinctie op te merken is, doet dit de vraag rijzen of er wel drie verschillende domeinen te onderscheiden zijn. Zowel de confirmatorische factoranalyses als deze intercorrelaties onderbouwen namelijk de interne structuur van vaardigheidsgebieden die behoren tot een bepaald domein niet. Om deze reden wordt sterk afgeraden om conclusies te trekken op basis van domeinscores. De auteurs

geven aan dat de hoge correlaties bevestigen dat in de algemene populatie adaptieve vaardigheden zich doorgaans gelijkmatig ontwikkelen (Kreemers et. al., 2020). In een relatief kleine groep volwassenen met een autismespectrumstoornis ($N = 81$) lijken er echter iets meer verschillen tussen de domeinen zichtbaar te worden, al zijn deze eerder beperkt.

De verwachtingen die werden geformuleerd met betrekking tot de hierboven beschreven intercorrelaties zijn gebaseerd op vorige versies van de ABAS-3. De data lijken deze echter niet te onderbouwen. Een exploratieve analyse van de interne structuur van de test, zoals een exploratieve factoranalyse of principale componentenanalyse, zou toelaten om na te gaan in welke mate de ABAS-3 in staat is een onderscheid te maken tussen verschillende domeinen of vaardigheidsgebieden van adaptieve vaardigheden. Het is jammer dat dergelijke exploratieve analyses niet werden uitgevoerd om de interne structuur van de ABAS-3 mee te bepalen. De auteurs geven hiervoor de volgende verklaring: *"Bij het construeren van de ABAS-3 was de opdracht om een instrument te ontwikkelen, waarmee deficiënties in het adaptief functioneren, in het kader van classificerende diagnostiek van een verstandelijke beperking (cf. DSM-5), getoetst kunnen worden. Hierbij werd beslist om niet af te wijken van de internationale literatuur over adaptief gedrag en verstandelijke beperking en dus om de drie domeinen Conceptueel, Sociaal en Praktisch te onderscheiden. De aanbeveling om geen belangrijke beslissingen te nemen op basis van domeinscores is dus verdedigbaar vanuit statistisch oogpunt, maar staat haaks op de criteria om de diagnose van verstandelijke beperking te stellen, zoals geformuleerd in de AAIDD door Schalock en Robert (2021) en in de DSM-5. Om die reden is er gekozen om een confirmatorische factoranalyse uit te voeren, waarbij de structuur getoetst werd die vereist is om de criteria voor verstandelijke beperking te toetsen, zoals geformuleerd in AAIDD en DSM-5, en dus geen exploratieve factoranalyse uit te voeren."*

Wanneer verwachte **verschillen tussen relevante groepen** worden teruggevonden, biedt dit verdere evidentie voor de begripsvaliditeit van een test. In de handleiding wordt eerst de **samenhang tussen leeftijd en adaptief gedrag** besproken. Er wordt verondersteld dat adaptief gedrag zich ontwikkelt en aangeleerd wordt naarmate een kind ouder wordt. Sommige vaardigheden verwerft men op jonge leeftijd, andere vaardigheden worden later verworven. Dit impliceert een positieve samenhang tussen adaptieve vaardigheden en leeftijd. Aangezien sommige vaardigheden een plafond bereiken in de volwassenheid, wordt verwacht dat de positieve samenhang tussen adaptief gedrag en leeftijd op groepsniveau steeds zwakker wordt (Tassé et. al., 2012). Hieronder volgt een korte bespreking van de gevonden resultaten. Deze liggen grotendeels in lijn met de verwachtingen. Er wordt echter geen informatie verschaft over de statistische significantie van de gevonden correlaties.

- Jonge kinderen ontwikkelen doorgaans verschillende adaptieve vaardigheden op relatief korte termijn (Kreemers et. al., 2020). Dit ligt in lijn met de sterke positieve correlaties die gevonden worden tussen leeftijd en de vaardigheidsgebieden, domeinen en het totaal van de *ouderversie 0-4 jaar* ($.80 \leq r \leq .93$). Voor de *leerkrachtversie 2-4 jaar* worden matige tot sterke correlaties gevonden ($.36 \leq r \leq .75$). De auteurs verklaren dit verschil door het kleinere leeftijdsbereik in de leerkrachtversie.
- Ook bij oudere kinderen en jongeren wordt een positieve samenhang verwacht tussen leeftijd en adaptief gedrag, al verwacht men dat deze minder sterk zal zijn aangezien bepaalde vaardigheden reeds een plafond bereiken in deze leeftijdsrange (Kreemers et. al., 2020). Voor de *ouderversie 5-17 jaar* wordt voor de domeinen en het totaal een sterk verband gevonden ($.52 \leq r \leq .82$). De vaardigheidsgebieden zijn matig tot sterk gecorreleerd met de leeftijd ($.46 \leq r \leq .85$). Wat betreft de relatie tussen leeftijd en de domeinen en het totaal van de *leerkrachtversie 5-17 jaar* wordt doorgaans een matig verband gevonden ($.43 \leq r \leq .54$), met uitzondering van 'Sociaal' waarvoor een zwak verband wordt vastgesteld ($r = .09$).
- Aangezien de ontwikkeling van adaptieve vaardigheden in de volwassenheid doorgaans een plafond bereikt heeft, wordt er geen samenhang verwacht tussen leeftijd en adaptief gedrag. De resultaten van de analyses voor de *zelf- en informant rapportage 18-80 jaar* bevestigen deze verwachting. Voor de meeste vaardigheidsgebieden, domeinen en het totaal wordt een heel zwak verband gevonden ($r \leq |.14|$). Uitzonderingen zijn het domein 'Sociaal' voor de *zelfrapportage 18-80 jaar* ($r = -.27$) en de schaal "Vrije tijd" in beide versies (respectievelijk $r = -.35$ en $r = -.22$). De auteurs geven aan dat dit negatief verband aantoont dat deze vaardigheden enigszins verzwakken naarmate mensen ouder worden (Kreemers, et. al., 2020).

Vervolgens worden **verschillen tussen klinische groepen** (i.e., verstandelijke beperking, autismespectrumstoornis en een gemengde klinische groep) en controlegroepen besproken. De controlegroepen werden geselecteerd uit de normeringssteekproeven en gematcht op basis van leeftijd en geslacht.

Voor de *ouderversie 0-4 jaar* ($N = 24$) en de *ouderversie 5-17 jaar* ($N = 156$) onderzocht men of er verschillen zijn tussen kinderen met en zonder een verstandelijke beperking. In lijn met de verwachtingen, vindt men in beide leeftijdsgroepen dat kinderen met een verstandelijke beperking zwakkere adaptieve vaardigheden hebben dan kinderen zonder een verstandelijke beperking. Dit verschil is statistisch significant, maar vooral wat betreft de ouderversie 0-4 jaar betreft, zijn de groepen erg klein waardoor de resultaten zeer voorzichtig geïnterpreteerd moeten worden. De auteurs bevestigen dat het gaat over Vlaamse kinderen, dit kan het best ook in de handleiding geëxpliciteerd worden.

Daarnaast ging men na of er verschillen zijn wat betreft adaptief gedrag tussen mensen met en zonder een autismespectrumstoornis (ASS) in Vlaamse steekproeven. Voor de *ouderversie 0-4 jaar* ($N = 76$) en de *ouderversie 5-17 jaar* ($N = 120$) wordt voor elk vaardigheidsgebied, domein en totaalscore een significant verschil in adaptief gedrag gevonden, waarbij de scores over de hele lijn lager zijn in de ASS-groepen vergeleken met de controlegroepen. Op basis van de effectgroottes blijken de verschillen het grootst te zijn voor het 'Sociaal' domein en het 'Conceptueel' domein (vooral 'Communicatie' en 'Zelfsturing'). Op basis van de *zelfrapportage 18-80 jaar* ($N = 120$) vindt men, in lijn met de verwachtingen, dat volwassen met ASS lagere scores rapporteren voor het 'Sociaal' domein (vaardigheidsgebieden: 'Vrije Tijd' en 'Sociale Vaardigheden') en de vaardigheidsgebieden 'Zelfsturing' en 'Huishoudelijke vaardigheden'. In de subgroep van volwassenen die een job hebben ($N = 40$) werd eveneens het vaardigheidsgebied 'Werk' opgenomen. Uit deze data blijkt, tegen de verwachtingen in, dat mensen met ASS hogere scores halen dan volwassenen zonder ASS voor het domein 'Praktisch'.

Tenslotte werden de verschillen tussen een controlegroep en een gemengde klinische groep uit de Vlaamse populatie onderzocht. De gemengde klinische groep bestond uit kinderen met (een vermoeden van) verstandelijke beperking, autismespectrumstoornis, ADHD, leerstoornissen, een visuele beperking, een auditieve beperking en/of een motorische beperking. De vergelijking op basis van de *ouderversie 0-4 jaar* ($N = 202$) en de *ouderversie 5-17 jaar* ($N = 778$) tonen opnieuw aan dat kinderen uit de gemengde klinische groep lagere scores behalen op alle schalen van de ABAS-3. De verschillen in adaptief gedrag zijn groter bij oudere dan bij jongere kinderen.

Bovengenoemde onderzoeken bieden verdere ondersteuning voor de begripsvaliditeit van de test, aangezien de resultaten in lijn liggen met de theoretische verwachtingen. Het is echter jammer dat er geen uitgebreider onderzoek gevoerd is naar de verschillen in adaptief gedrag tussen personen met en zonder een verstandelijke beperking. In de huidige studies werd dit namelijk niet onderzocht bij volwassenen en slechts bij een kleine groep jonge kinderen. Het zou een meerwaarde zijn om hier verder onderzoek naar uit te voeren, aangezien het in kaart brengen van de adaptieve vaardigheden een belangrijk onderdeel vormt in het stellen van een diagnose verstandelijke beperking.

Er werd eveneens onderzoek gedaan naar de **relatie tussen de ABAS-3 en andere gerelateerde instrumenten**. Wanneer verwachte verbanden worden bevestigd, biedt dit evidentie voor de **convergente validiteit**, een specifiek onderdeel van begripsvaliditeit.

Om de convergente validiteit van de ABAS-3 na te gaan, werden de resultaten van de *ouderversie 0-4 jaar* vergeleken met deze van de *Vineland Screener 0-6 jaar* (Vineland-S; Scholte, van Duijn, Dijkxhoorn, Noens, & van Berckelaer-Onnes, 2014). Dit is een screeningsvragenlijst die het adaptief gedrag in kaart brengt voor de domeinen 'Communicatie', 'Dagelijkse Vaardigheden', 'Socialisatie', en 'Motoriek'. Opvallend is dat alle schalen van de ABAS-3 heel sterk correleren met alle schalen van de Vineland Screener (bv. Vineland Screener Motoriek met ABAS-3 Sociaal: $r = .93$). Er is met andere woorden geen onderscheid tussen de verschillende schalen, waardoor de vraag opnieuw rijst of er wel een onderscheid gemaakt kan worden in verschillende domeinen. De sterke correlaties tussen de totaalscores van beide testen onderbouwen de convergente validiteit van de ABAS-3. Daarnaast werden de resultaten op de *leerkrachtversie 5-17 jaar* vergeleken met de

leerkracht versie 12-18 jaar van de *RSI Beoordelingsschaal voor functionele beperkingen* in een Nederlandse steekproef (RSI; Goldstein & Naglieri, 2018). Dit instrument brengt de functionele beperkingen in kaart en is bedoeld om de impact van psychische, fysieke en cognitieve stoornissen op het dagelijkse leven in kaart te brengen. De RSI omvat drie schalen namelijk 'Sociaal', 'Mobiliteit' en 'School'. Aangezien de RSI functionele beperkingen en de ABAS-3 adaptieve vaardigheden in kaart brengen, worden negatieve correlaties verwacht. De resultaten tonen aan dat er matige tot sterke negatieve correlaties zijn tussen de verschillende domeinen en totaalscores van beide instrumenten. Er is geen informatie over de samenhang tussen deze instrumenten in Vlaamse steekproeven.

Deze studies geven een indicatie van een voldoende tot goede convergente validiteit voor de *ouderversie 0-4 jaar* (Vlaamse steekproef) en *leerkrachtversie 5-17 jaar* (Nederlandse steekproef). Er is geen informatie over de versies gericht op volwassenen. Verder onderzoek met andere instrumenten in Vlaamse steekproeven om de congruente validiteit van de andere versies van de ABAS-3 na te gaan is wenselijk. We erkennen echter dat er weinig tot geen instrumenten zijn die genormeerd zijn voor Vlaanderen/Nederland die adaptief gedrag van volwassenen in kaart brengt.

Tot slot werd de **divergente validiteit** onderzocht om na te gaan in welke mate de meetpretentie van de ABAS-3 verschilt van de meetpretentie van andere instrumenten (Kreemers et. al., 2020). Hiervoor werden de resultaten van de *ouderversies 0-4 jaar* en *5-17 jaar* van de ABAS-3 vergeleken met twee versies van de *Child Behaviour Checklist (CBCL) 1-5 jaar* en *6-18 jaar* (Achenbach & Rescorla, 2011). Deze vragenlijsten meten onder andere internaliserend en externaliserend probleemgedrag. Er worden zwakke tot matige negatieve verbanden gevonden. Deze resultaten tonen aan dat de divergente validiteit als voldoende kan beoordeeld worden. Verschillende versies van de ABAS-3 werden echter niet onderzocht, waardoor geen beoordeling kan gedaan worden van de leerkracht- en volwassenenversies. Daarnaast werden de ouderversies enkel vergeleken met de CBCL en werd geen vergelijkend onderzoek met andere instrumenten gedaan.

Criterionvaliditeit

Criterionvaliditeit wordt onderzocht door de samenhang na te gaan tussen een test en een 'real-world' criterium meting. Hieronder plaatsen we eveneens het onderzoek waarin wordt nagegaan of de ABAS-3 goed kan discrimineren tussen personen met en zonder (een vermoeden van) een verstandelijke beperking. Dit werd gedaan op basis van ROC-analyses (*Relative Operating Characteristic*) in een Vlaamse steekproef. Hierbij worden ROC-curves opgesteld waarbij de oppervlakte onder de curve (*area under the curve*, AUC) aangeeft hoe goed de test kan discrimineren. Een AUC-waarde tussen .90 en 1.00 wijst op een uitstekend discriminerend vermogen, waarden van .80 tot .90 geven een goed discriminerend vermogen weer, bij waarden van .70 tot .80 wordt gesproken van een redelijk discriminerend vermogen, bij waarden tussen .60 en .70 is het discriminerend vermogen matig en waarden onder .60 zijn onvoldoende (Kreemers et. al., 2020).

Voor de *ouderversie 0-4 jaar* werd een AUC-waarde van .95 gevonden, wat erop wijst dat deze versie van de ABAS-3 een uitstekend discriminerend vermogen heeft. Deze resultaten worden echter best met enige voorzichtigheid geïnterpreteerd aangezien de klinische groep slechts uit 22 kinderen bestond (16 met een diagnose verstandelijke beperking en 6 kinderen met een vermoeden van verstandelijke beperking). De vergelijkingsgroep bestond uit 860 kinderen. Alle kinderen waren ouder dan 1 jaar.

Voor de *ouderversie 5-17 jaar* werd een AUC-waarde van .98 gevonden, wat er eveneens op wijst dat deze versie van de ABAS-3 een uitstekend discriminerend vermogen heeft. De klinische groep bestond uit 93 kinderen die allen de diagnose verstandelijke beperking hadden gekregen. De vergelijkingsgroep bestond uit 1412 kinderen en jongeren.

Algemeen kan gesteld worden dat het onderzoek naar de criteriumvaliditeit van de ABAS-3 om verschillende redenen nog onvoldoende is. Ten eerste zijn de klinische groepen waarmee vergeleken wordt klein, waardoor voorzichtigheid zeker aangewezen is. Ten tweede wordt er enkel een vergelijking gemaakt met kinderen met (een vermoeden van) een diagnose verstandelijke beperking en werden geen andere vergelijkingen met 'real-world' metingen gedaan. Tot slot werd de criteriumvaliditeit niet onderzocht voor de leerkrachtversies en de

volwassenversies. Meer onderzoek is noodzakelijk om ook de voorspellende criteriumvaliditeit van de ABAS-3 te onderbouwen.

Sterktes

- De inhouds- en begripsvaliditeit worden in het algemeen als voldoende beoordeeld.

Aandachtspunten

- Het onderzoek naar de interne structuur van de ABAS-3 en de confirmatorische factoranalyse impliceren dat de ABAS-3 vooral één grote factor “adaptieve vaardigheden” meet en niet zonder meer in staat is om de verschillende domeinen (of vaardigheidsgebieden) te onderscheiden.
- De verschillen tussen kinderen met en zonder een verstandelijke beperking zijn nog beperkt onderzocht, vooral onderzoek in volwassenen ontbreekt.
- De convergente en divergente validiteit zijn niet voor elke versie van de ABAS-3 onderzocht. Aanvullend en uitgebreider onderzoek is aangewezen.
- De (voorspellende) criteriumvaliditeit dient verder onderzocht te worden.

De **inhoudsvaliditeit** wordt voor alle versies als **voldoende** beoordeeld voor zowel het nemen van *belangrijke* als *minder belangrijke beslissingen op individueel niveau*.

Omwille van bovenstaande aandachtspunten wordt de **begripsvaliditeit** van de ouder- en leerkrachtversies als **voldoende** beoordeeld voor het nemen van zowel *belangrijke* als *minder belangrijke beslissingen op individueel niveau*. Wat betreft de **begripsvaliditeit** voor de volwassenen versies wordt dit als **voldoende** beoordeeld voor het nemen van *minder belangrijke*, en als **onvoldoende** beoordeeld voor het nemen van *belangrijke beslissingen op individueel niveau*.

Er is te weinig onderzoek gedaan naar de **criteriumvaliditeit** waardoor deze voor alle versies van de ABAS-3 als **onvoldoende** wordt beoordeeld voor zowel het nemen van *belangrijke* als *minder belangrijke beslissingen op individueel niveau*.

‘De context (...) moet steeds worden meegenomen in de interpretatie van de resultaten.’

Samenvatting beoordeling

Het voorliggende adviesrapport handelt over de evaluatie van de Adaptive Behavior Assessment System – versie 3 – Nederlandstalige versie (ABAS-3). Het rapport heeft voornamelijk betrekking op het gebruik en de normering van de ABAS-3 in Vlaanderen. Tot op heden is er nog geen beoordeling voor het gebruik en de normering in Nederland.

Er kan worden besloten dat de ABAS-3 een gebruiksvriendelijk instrument is om adaptieve vaardigheden in kaart te brengen bij kinderen, jongeren en volwassenen in verschillende contexten (thuis, school, werk). De handleiding is goed uitgewerkt en transparant, en biedt voldoende handvatten voor de testleider om mee aan de slag te gaan. Het is echter belangrijk om met de volgende punten rekening te houden:

- Voor de ouder- en leerkrachtversies is voorzichtigheid geboden bij het interpreteren van scores in de uiterste leeftijdsgroepen; alsook bij zeer lage of zeer hoge scores. In de volwassenversies is voorzichtigheid aangewezen bij hoge scores.
- Voorzichtigheid is zeker ook geboden bij het gebruik van het instrument in doelgroepen met een lage SES en/of migratie-achtergrond.
- Hou rekening met bodem- en plafondeffecten en wees voorzichtig met het gebruik van leeftijdsequivalenten.

De context (thuissituatie versus school, werk of dagbesteding) moet steeds worden meegenomen in de interpretatie van de resultaten, deze kunnen namelijk niet (zomaar) gegeneraliseerd worden naar andere contexten.

Onderstaande kwaliteitslabels werden toegekend aan de ABAS-3 voor gebruik in Vlaanderen door het Kwaliteitscentrum voor Diagnostiek vzw na een grondige beoordeling van de test door twee onafhankelijke experts en een wetenschappelijk onderzoeker van het Kwaliteitscentrum zelf. Voor de beoordeling werd gebruik gemaakt van het beoordelingsmodel voor de beschrijving en evaluatie van psychologische en educatieve testen van de European Federation of Psychologists' Associations (EFPA; Evers et al., 2013), naar het Nederlands vertaald door het Kwaliteitscentrum voor Diagnostiek vzw.

ABAS-3 – Ouderversie 0-4 jaar en leerkrachtversie 2-4 jaar

Beoordeling voor gebruik in Vlaanderen	<i>Relatief minder belangrijke beslissingen*</i>	<i>Belangrijke beslissingen**</i>
Kwaliteit uitgangspunten, presentatie en beschikbare informatie	Goed	Goed
Kwaliteit van het testmateriaal	Uitstekend	Uitstekend
Vlaamse normen	Goed	Voldoende
Betrouwbaarheid	Goed	Goed
Inhoudsvaliditeit	Voldoende	Voldoende
Begripsvaliditeit	Voldoende	Voldoende
Criteriumvaliditeit	Onvoldoende	Onvoldoende
Kwaliteit van computergegenereerde rapporten	NVT	NVT

ABAS-3 – Ouder- en leerkrachtversie 5-17 jaar

Beoordeling voor gebruik in Vlaanderen

	<i>Relatief minder belangrijke beslissingen*</i>	<i>Belangrijke beslissingen**</i>
Kwaliteit uitgangspunten, presentatie en beschikbare informatie	Goed	Goed
Kwaliteit van het testmateriaal	Uitstekend	Uitstekend
Vlaamse normen	Goed	Voldoende
Betrouwbaarheid	Uitstekend	Goed
Inhoudsvaliditeit	Voldoende	Voldoende
Begripsvaliditeit	Voldoende	Voldoende
Criteriumvaliditeit	Onvoldoende	Onvoldoende
Kwaliteit van computergegenereerde rapporten	NVT	NVT

ABAS-3 – Zelf- en informantenrapportage 18-80 jaar

Beoordeling voor gebruik in Vlaanderen

	<i>Relatief minder belangrijke beslissingen*</i>	<i>Belangrijke beslissingen**</i>
Kwaliteit uitgangspunten, presentatie en beschikbare informatie	Goed	Goed
Kwaliteit van het testmateriaal	Uitstekend	Uitstekend
Vlaamse normen	Goed	Voldoende
Betrouwbaarheid	Voldoende	Onvoldoende
Inhoudsvaliditeit	Voldoende	Voldoende
Begripsvaliditeit	Voldoende	Onvoldoende
Criteriumvaliditeit	Onvoldoende	Onvoldoende
Kwaliteit van computergegenereerde rapporten	NVT	NVT

* Voorbeelden van *relatief minder belangrijke beslissingen* zijn: voortgangscontrole, beschrijvend gebruik van de testresultaten, therapie-indicatie en beroepskeuzebegeleiding (Evers et al., 2010, pg. 22). Hierbij werd rekening gehouden met zowel totaal- en domeinscores als de scores op de vaardigheidsgebieden.

** Met *belangrijke beslissingen* wordt bedoeld: beslissingen die op basis van de testcores worden genomen, die in principe, of op korte termijn, onomkeerbaar zijn, en die voor een belangrijk deel buiten de geteste persoon om worden genomen. Voorbeelden van belangrijke beslissingen zijn: personeelselectie, verwijzing naar speciaal onderwijs, opname/ontslag kliniek, certificering (Evers et al., 2010, pg. 22). Hierbij werd rekening gehouden met de totaal- en domeinscores. De scores op vaardigheidsgebieden zijn niet geschikt om belangrijke beslissingen op te baseren.

Errata

Handleiding 'Instructies en verantwoording'

- pg. 24 (2.3.1): In deze alinea wordt gesproken van 'een grote verscheidenheid aan kinderen'. Dit laat echter uitschijnen dat de ABAS-3 niet bij volwassenen zou kunnen worden afgenomen. Om deze reden stellen we voor om de zin aan te passen naar 'een grote verscheidenheid aan kinderen en volwassenen'.
- pg. 70: De losstaande zin klopt grammaticaal niet: 'staan nog een aantal voorbeeld waarvoor...'. Dit zou aangepast moeten worden naar 'staan nog een aantal voorbeelden waarvoor...'.
- pg. 87 (4.4.2): De informatie in de tekst en de bijhorende tabel 17 is verschillend van elkaar. In de tekst wordt aangegeven dat de leerkrachtversie 2-4 jaar voor 360 kinderen is ingevuld. In de tabel wordt aangegeven dat dit gaat over 338 kinderen (voor weging). Hierdoor is het niet duidelijk hoeveel kinderen voor en na weging deel uitmaakten van de steekproef.
- pg. 89 (4.4.2): Er staat een lidwoord te veel in de volgende zin: '*In totaal vulden 629 volwassenen tussen de 18 en 80 jaar een de zelfrapportageversie van de ABAS over zichzelf in*'. Het onderstreepte woord zou geschrapt moeten worden.
- pg. 95 (5.1): Onderaan de pagina wordt verwezen naar EFPA. Deze afkorting wordt echter verkeerd voluit geschreven. Er staat nu 'Psychologist...' dit zou veranderd moeten worden naar 'Psychologists...'.
- pg. 99-105: Op dit moment staan de tabellen met de Vlaamse en Nederlandse data door elkaar. We zouden voorstellen om de Vlaamse tabellen te groeperen alsook de Nederlandse tabellen te groeperen.
- pg. 100 (tabel 26): De titel van de laatste kolom is 'Totaal' dit zou aangepast moeten worden naar '11 jaar'.
- pg. 113 (tabel 43): In het bijschrift van tabel 43 staat 'totaall', dit zou 'totaal' moeten zijn.
- pg. 114-115: De tabellen 40-45 hebben verkeerde titels. Er staat interbeoordeelaarsbetrouwbaarheid terwijl het gaat over paralleltestbetrouwbaarheid. Daarnaast worden de twee afnames onderscheiden met de titels 'meetmoment 1 en meetmoment 2'. We zouden de suggestie willen doen om dit te vervangen door 'ouderversie en leerkrachtversie'.
- pg. 121 (5.5): Elke alinea heeft een cursieve tussentitel, echter de titel van de eerste alinea op deze pagina, *ouderversie 5-17 jaar*, staat vet in plaats cursief.
- pg. 143-149: In de tabellen 70 t.e.m. 76 staat als derde domein 'Conceptueel' waar 'Praktisch' zou moeten staan.
- pg. 141: In de volgende zin wordt naar de verkeerde test verwezen: "*De correlatiecoëfficiënten tussen de schalen van de RSI Leerkrachtversie 12-18 jaar en de domeinen en het totaal van de ABAS-3 Ouderversie 0-4 jaar worden weergegeven in tabel 67.*" Dit zou aangepast moeten worden naar 'leerkrachtversie 5-17 jaar'.
- pg. 144: In tabel 71 staat de regel van zelfverzorging schijnbaar zonder reden volledig in het vet.
- pg. 144 en pg. 147: In de tekst op pg. 144 staat er dat voor de zelfrapportageversie telkens 65 personen opgenomen zijn in de steekproeven met personen met en zonder ASS. De voetnoot bij de tabel geeft echter aan dat er in elke steekproef 60 personen zijn opgenomen. Deze informatie spreekt elkaar tegen. Het is wenselijk om dit te verduidelijken.
- pg. 158: Er wordt vermeld dat 'hierna' aanbevelingen worden gedaan over de interpretatie van de resultaten bij doelgroepen met een TMA-achtergrond of lage SES, maar deze aanbevelingen volgen niet.
- pg. 173 (1.1): Het woord curvilineariteit staat verkeerd geschreven (curvilineariteit). Het tussentiteltje 'Curvilineariteit' wordt vet weergegeven terwijl de andere tussentitels schuingedrukt staan.

Referenties

- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA School-Age Forms & Profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.
- American Association on Intellectual and Developmental Disabilities (2010). *Intellectual disability: Definition, classification, and systems of supports* (11th ed.). Washington, DC: AAIDD.
- American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: APA.
- Byrne, B. M. (2012). *Structural equation modeling with Mplus*. New York: Routledge.
- Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). New York: Academic Press.
- Davis, F. B. (1959). Interpretation of differences among averages and individual test scores. *Journal of Educational Psychology*, 50, 162–170.
- Evers, A., Hagemester, C., Hostmaelingen, A., Lindley, P., Muniz, J., & Sjöberg, A. (2013). EFPA Review Model for the description and evaluation of psychological and educational tests. Geraadpleegd via: <http://www.efpa.eu/download/650d0d4ecd407a51139ca44ee704fda4>
- Evers, A., Lucassen, W., Meijer, R., & Sijtsma, K. (2010). *COTAN Beoordelingssysteem voor de kwaliteit van tests*. Geraadpleegd via: <https://www.psynip.nl/wp-content/uploads/2016/07/COTAN-Beoordelingssysteem-2010.pdf>
- Goldstein, S. & Naglieri, J. A. (2009). *Assessing impairment: From theory to practice*. New York: Springer.
- Hallgren, K. A. (2012). Computing Inter-Rater Reliability for Observational Data: An Overview and Tutorial. *Tutorials in Quantitative Methods for Psychology*, 8(1), 23–34.
- Harrison, P. Oakland, T. (2015). *Adaptive Behavior Assessment System, Third Edition [Manual]*. Torrance, CA: Western Psychological Services.
- Hervé, A. (2007). The Bonferonni and Šidák corrections for multiple comparisons. In N. Salkind (Ed.), *Encyclopedia of measurement and statistic, Volume 1* (pp. 103-106). Thousand Oaks: Sage Publications.
- Kreemers, B., Maljaars, J., Briers, V., Bulteel, S., Storms, G., Maes, B., & Noens, I. (2020). *ABAS-3 Schaal voor adaptief gedrag – Instructies en verantwoording*. Amsterdam: Hogrefe Uitgevers B.V.
- Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading: Addison-Wesley.
- Norfolk, P. A., Farmer, R. L., Floyd, R. G., Woods, I. L., Hawkins, H. K., & Irby, S. M. (2015). Norm block sample sizes: A review of 17 individually administered intelligence tests. *Journal of Psychoeducational Assessment*, 33(6), 544-554.
- Scholte, E., van Duijn, G., Dijkxhoorn, Y., Noens, I., & van Berckelaer-Onnes, I. (2014). *Vineland-S: Vineland Screener 0-6 jaar [Handleiding]*. Amsterdam, Nederland: Hogrefe.
- Statbel – Census 2011 (2020, augustus). https://www.census2011.be/download/downloads_nl.html
- Statbel (2020, augustus). <https://statbel.fgov.be/nl/themas/bevolking/structuur-van-de-bevolking#panel-14>
- Tassé, M. J., Schalock, R. L., Balboni, G., Bersani, H., Duffy, S. A., De Valenzuela, J. S., ... Zhang, D. (2012). The construct of adaptive behavior: It's conceptualization, measurement, and use in the field of intellectual disability. *American Journal on Intellectual and Developmental Disabilities*, 117, 291-303.
- Van Breukelen, G. J. P., & Vlaeyen, J. W. S. (2005). Norming clinical questionnaires with multiple regression: The Pain Cognition List. *Psychological Assessment*, 3, 336-344.
- Van Gampelaere, C., De Ganck, J., & Schoupe, N. (2020). *Richtlijn met betrekking tot stratificatiecriteria voor normeringsonderzoek van psychodiagnostische testen in Vlaanderen*. Kwaliteitscentrum voor Diagnostiek. Geraadpleegd op 8 augustus 2022 via <https://portaal.kwaliteitscentrumdiagnostiek.be/>
- Vlaams Departement Onderwijs en Vorming, V. (2015). *Vlaams onderwijs in cijfers 2014-2015*. <https://www.vlaanderen.be/publicaties/vlaams-onderwijs-in-cijfers-2014-2015> Het Vlaams Ministerie van Onderwijs en Vorming (2017a). Leerlingkenmerken.
- Wechsler, D. (2018). *WISC-V-NL. Wechsler Intelligence Scale for Children – Fifth Edition – Nederlandstalige bewerking. Technische handleiding*. Amsterdam, Nederland: Pearson.