

Een non-parametrische frisse blik op de MHC-SF

Schaalbaarheid van de Mental Health Continuum-Short Form (MHC-SF) onderzocht aan de hand van exploratieve Mokkenschaaanalyse

Jef Poelman, S1266829

Orthopedagogiek, Faculteit Gedrags- en Maatschappijwetenschappen

Masterthese Orthopedagogiek PAMA5166

Rijksuniversiteit Groningen

Begeleider: dr. M.S.C. Paap

Tweede beoordelaar: dr. W.J. Post

Datum: 30-05-2025

Aantal woorden: 6250

Abstract

In recent decades, much research has been conducted on the psychometric qualities and scalability of the Mental Health Continuum - Short Form (MHC-SF). However, no studies have yet been published that investigate the MHC-SF based on non-parametric Mokken Scale Analysis (MSA). Because MSA requires fewer assumptions, compared to frequently used factor analysis, MSA is an analysis that lends itself well to situations where items are scored on an ordinal scale making use of Likert scales. Using data from 650 participants from Dutch households collected by the LISS panel, the scalability of the MHC-SF was examined using exploratory MSA. The findings demonstrate a moderate but acceptable scalability for the unidimensional model, whereby it is plausible that all assumptions are met. The findings of exploratory MSA on different age groups revealed subtle differences, but moderate scalability remained evident. For application in practice, it is recommended that the purpose of administering the questionnaire guides the choice of a specific model, as different models lead to acceptable scalability.

Samenvatting

De afgelopen decennia is veel onderzoek gedaan naar de psychometrische kwaliteiten en schaalbaarheid van de Mental Health Continuum - Short Form (MHC-SF). Er zijn vooralsnog geen onderzoeken gepubliceerd die onderzoek doen naar MHC-SF op basis van de non-parametrische Mokschaalanalyse (MSA). Omdat het aan minder assumpties hoeft te voldoen dan de vaak gebruikte factoranalyse, is MSA een analyse die zich prima leent voor situaties waarin items ordinaal worden gescoord op basis van Likert schalen. Aan de hand van data van 650 participanten, afkomstig uit Nederlandse huishoudens opgehaald door het LISS-panel, is aan de hand van exploratieve MSA de schaalbaarheid van de MHC-SF onderzocht. De uitkomst toont een matige maar acceptabele schaalbaarheid voor een unidimensionele schaal, waarbij bovendien aannemelijk is gemaakt dat aan alle assumpties wordt voldaan. De uitkomsten van exploratieve MSA op verschillende leeftijdsgroepen laten nuanceverschillen zien, maar er bleef sprake van een matige schaalbaarheid. Voor het praktisch gebruik wordt aanbevolen het doel van de afname van de vragenlijst leidend te laten zijn bij de keuze voor een specifiek model, aangezien verschillende modellen tot een acceptabele schaalbaarheid leiden.

Inhoudsopgave

Inleiding	4
Methode	8
Design	8
Ethische kwesties	8
Populatie, steekproef en procedure	9
Variabelen en instrumenten	10
Analyseplan	10
Resultaten	12
Beschrijvende statistiek	12
Hoofdbevindingen: Dimensionaliteit	13
Assumptie checks	15
Monotoniciteit	15
Lokale onafhankelijkheid	16
Betrouwbaarheid	16
Conclusie en discussie	16
Literatuurlijst	19

Inleiding

De staat van de mentale gezondheid van jongeren binnen de Nederlandse maatschappij staat ter discussie. In opdracht van de Tweede Kamer is onderzoek verricht naar de mentale gezondheid van jongeren en studenten. Hieruit blijkt dat de mentale gezondheid van tieners en vooral jongvolwassenen al enige tijd afneemt en vooral veel slechter werd tijdens de coronaperiode. Vooral jongvolwassen vrouwen en ook meisjes in het voortgezet onderwijs vallen op door de sterke verslechtering in mentale gezondheid (Arslan et al., 2024). Levi ten Dam, in 2023 benoemd tot bijzonder hoogleraar aan de universiteit van Amsterdam (Universiteit van Amsterdam, 2023), spreekt dit tegen door te stellen dat de mentale gezondheid van jongeren niet achteruit gaat (Universiteit van Amsterdam, 2024). Hij voert de ‘prevalentie- inflatie hypothese’ aan, waarbinnen jongeren een lagere mentale gezondheid toegeschreven wordt, doordat ze zich in het algemeen meer bewust zijn van het bestaan van mentale problematiek. Het gevolg is dat jongeren vaker naar een psycholoog worden doorverwezen (Universiteit van Amsterdam, 2024). Epidemiologische studies laten zien dat, ondanks dat mentale gezondheid vaker een onderwerp van gesprek is, de aanwezigheid van psychische klachten onder jongeren in de afgelopen dertig jaar nauwelijks veranderd is (Universiteit van Amsterdam, 2024). Op basis van bovenstaande wordt duidelijk dat er sprake is van verschillen in interpretatie. Volgens Goodman en collegae (2018) is het waar dat mensen met psychische klachten over het algemeen aanzienlijke beperkingen in welzijn ervaren, maar is het verband tussen deze klachten en bepaalde dimensies van welzijn niet lineair te associëren. Stellen dat de mentale gezondheid niet achteruitgaat omdat de prevalentie van psychische klachten gelijk is gebleven, zou dan duiden op een fact-value distinctie. Anders gezegd is wel gebleken dat psychische klachten onder jongeren nagenoeg gelijk zijn gebleven (feit), maar deze informatie is ontoereikend om daar de waarde aan te hangen dat de mentale gezondheid van jongeren niet achteruit gaat. Tegenstrijdige overtuigingen tonen, ongeacht wie gelijk heeft, het belang aan van een instrument waarmee zo objectief mogelijk de latente variabele mentale gezondheid kan worden gemeten.

De wereldgezondheidsorganisatie (WHO) definieert mentale gezondheid als een staat van mentaal welzijn die het voor mensen mogelijk maakt met stress om te gaan, hun potentieel te verwezenlijken, onderwijs te volgen en arbeid te verrichten, en een maatschappelijke bijdrage te leveren (World Health Organization: WHO, 2019). Sociale wetenschappers onderzoeken welzijn sinds midden 20ste eeuw. Gezondheid en welzijn werden geruime tijd gelijkgesteld aan de

afwezigheid van een ziekte, stoornis of problemen, maar hedendaags onderzoek suggereert dat welzijn eerder bestaat uit de aanwezigheid van capaciteiten, vaardigheden, sterktes en andere positieve eigenschappen (Magyar & Keyes, 2019). Vanuit dit uitgangspunt ontwikkelde Keyes (2002) een meetinstrument, de Mental Health Continuum - Long Form (MHC-LF), met 40 items, gericht op het in kaart brengen van positieve mentale gezondheid van participanten. Andere veelgebruikte instrumenten om mentale gezondheid te meten zijn, volgens Breedvelt en collegae (2019), de DASS-21 (Lovibond & Lovibond, 1995), HADS (Zigmond & Snaith, 1983), PANAS (Watson et al., 1988) en GHQ-12 (Goldberg, 1973). De MHC-LF onderscheidt zich van deze meetinstrumenten door zich enkel te richten op positieve mentale gezondheid, meetbaar gemaakt door de drie aspecten van welbevinden te scoren: emotioneel welbevinden, psychisch welbevinden en sociaal welbevinden (Keyes, 2002). Het instrument werd later, ten behoeve van de efficiëntie, verkort tot een versie van 14 items, de Short Form (MHC-SF) (Keyes, 2005). De antwoordmogelijkheden van de MHC-SF zijn zo geformuleerd, dat de door de respondent ervaren frequentie van elk symptoom van mentale gezondheid gemeten kan worden (Keyes, 2018). Het doel is een duidelijke standaard te bieden voor het onderverdelen van participanten in verschillende categorieën van positieve geestelijke gezondheid (Keyes, 2018). De uitkomst van de vragenlijst levert een score op die participanten onderverdelen in drie niveaus: goede, gematigde of zwakke mentale gezondheid (Keyes, 2005). Het instrument is voor het eerst vertaald naar het Nederlands door Lamers en collegae (2010) en staat volgens therapieland.nl, een platform speciaal ontwikkeld om professionals te ondersteunen in het selecteren van de juiste behandelingen en het bijhouden van de voortgang van hun cliënten, in de top 10 van meest gebruikte GGZ-vragenlijsten (Nisha, 2024). Volgens Westerhof en Keyes (2008) is de MHC-SF het meest complete instrument om geestelijke gezondheid te meten. Het Nationaal Regieorgaan Onderwijsonderzoek (NRO) ziet in de MHC-SF een geschikte vragenlijst om te gebruiken voor het volgen van het brede welbevinden van middelbare scholieren over tijd (Kennisrotonde, 2024).

De kwaliteit van uitspraken die gedaan worden op basis van de MSH-SF is direct afhankelijk van de psychometrische kwaliteiten. Luijten en collegae (2019) analyseerden onderzoekgegevens, afkomstig van Nederlandse scholieren tussen de 11 en 17 jaar, waaruit een excellente betrouwbaarheid van de volledige schaal ($\alpha = .91$) en een goede betrouwbaarheid voor elke subschaal apart ($\alpha > .80$) bleek. De volledige schaal liet een vergelijkbaar beeld zien over de

verschillende leeftijden en tussen jongens en meisjes (Luijten et al., 2019). Een systematische review met betrekking tot MHC-SF, waarbinnen onderzoeken zijn geselecteerd op het uitvoeren van factoranalyse aan de hand van exploratieve factoranalyse, confirmatieve factoranalyse of door middel van een exploratief structureel vergelijkingsmodel (ESEM), laat zien dat de 46 betrokken onderzoeken consistent wezen op voldoende structurele validiteit en een goede tot uitstekende interne consistentie betrouwbaarheid. Regelmatig wezen analyses op invariantie voor leeftijdsgroepen, geslacht en cultuur of nationaliteit. Factoranalyse gaf geen eenduidige steun voor één structuur, maar de meeste onderzoeken vonden steun voor een goede fit van het 3 factoren model. ESEM binnen 26 onderzoeken vond ook de meeste steun voor dit model. Andere onderzochte modellen zijn het 2 factoren model, het hierarchische model en het bi-factor model (Iasiello et al., 2022).

Een alternatieve manier om de schaalbaarheid van de MHC-SF te analyseren is aan de hand van Mokken Schaal Analyse (MSA) (Mokken, 1971). Het model van Mokken is minder restrictief, omdat het uitgaat van minder strenge assumpties. MSA werkt met ordinale data, waarbij een ordinale schaal voor persoonsmetingen op basis van de observeerbare testscore (en, wanneer aan de voorwaarden van dubbele monotoniciteit wordt voldaan, itemscores) wordt geïmpliceerd (Sijtsma & Van Der Ark, 2016). MSA toont zich, in tegenstelling tot factoranalyse, robuust bij zowel een niet normaal verdeelde latente variabele als bij niet normale itemverdelingen (Kappenburg -ten Holt, 2014). Om inzicht te krijgen in de frequentie waarin MSA in publicaties is toegepast op de MHC-SF en dit te vergelijken met de frequentie van factoranalyse, zijn binnen EBSCOhost alle databases geselecteerd en is gezocht naar de combinatie MHC-SF en (1) CFA of confirmatory factor analysis, (2) EFA of exploratory factor analysis, (3) SEM Structural Equation Modeling en (4) MSA of Mokken Scale Analysis. Hierbij zijn achtereenvolgens 167, 59, 84 en 0 hits gevonden. Het toepassen van MSA op de MHC-SF is binnen deze databases dus nog niet gepubliceerd, in tegenstelling tot de hierboven beschreven factoranalyses, waarover binnen veel onderzoeken iets is beschreven. MSA zou daarom nieuwe inzichten kunnen brengen omdat het meer informatie geeft over de schaalbaarheid door middel van schaalbaarheidscoëfficiënten (Mokken, 1971). Het voordeel van non-parametrische modellen zoals MSA, ten opzichte van parametrische modellen zoals factoranalyse, is dat het gebruik maakt van minder restrictieve assumpties (Sijtsma & Van Der Ark, 2016).

Bewijs voor het drie factor model vanuit factoranalyse zou reden kunnen zijn voor het uitvoeren van een confirmatieve Mokkenschaaanalyse, maar juist omdat een Mokkenschaaanalyse wel rekening houdt met ongelijke frequentieverdelingen van items en de uitkomsten daardoor een ander beeld zouden kunnen laten zien dan bij factoranalyse, kiezen we ervoor binnen dit onderzoek een exploratieve analyse uit te voeren.

Ook is het van belang om te onderzoeken of de variabele leeftijd de uitkomsten van de analyse beïnvloedt. Bij onderzoek met factoranalyse werd aannemelijk geacht dat leeftijd geen invloed had op de resultaten van de analyse (Luijten et al., 2019). Het is van belang te onderzoeken of dit binnen ons onderzoek met uitvoering van MSA ook aannemelijk is. Uit cijfers van het Rijksinstituut voor volksgezondheid en milieu (RIVM), gericht op mentale gezondheid, blijkt dat 84,2 % van de mensen van 18 jaar en ouder tevreden (score 7 of hoger op een schaal van 1 tot 10) is over hun leven, waarbij de groep van 65 t/m 74 jaar met 89,9 % het hoogste percentage laat zien en de groep van 18 t/m 24 jaar met 76,1 % het laagste percentage (RIVM, 2023). Het zou kunnen zijn dat ook binnen de door ons gebruikte data dergelijke verschillen worden gevonden in totaalscores tussen groepen. De vraag is echter of die totaalscores ook op een valide manier vergeleken kunnen worden tussen de groepen. Het is daarom interessant om MSA toe te passen op de beide subgroepen om daardoor inzichtelijk te krijgen of en in welke mate er verschillen zijn in de gevonden schaalindeling. Hierbij hebben we er in het kader van haalbaarheid voor gekozen om de leeftijdscategorieën die op basis van deze cijfers naar verwachting het verst uit elkaar liggen te selecteren. Daarom wordt de MSA van de groep van 18 t/m 24 jaar vergeleken met de MSA van de groep van 65 jaar en ouder.

Ons onderzoek richt zich, aan de hand van Paneldata, afkomstig van LISS panel, op de Nederlandse bevolking. We richten ons daarbij op de volgende vragen:

1. Welke schaal of schalen worden gevonden bij het uitvoeren van een exploratieve Mokkenschaaanalyse? Is dit in lijn met de structuur zoals deze is voorgesteld in de literatuur?
2. Welke schaal of schalen worden gevonden, wanneer we enkel de participanten in de leeftijd van 18 t/m 24 jaar selecteren? En welke schaal of schalen worden gevonden bij het selecteren van participanten in de leeftijd van 65 jaar en ouder? Worden vergelijkbare schalen gevonden voor de leeftijdsgroepen en de gehele steekproef?

In de methode zal uitgebreid beschreven worden hoe het onderzoek is vormgegeven en uitgevoerd. Vervolgens zullen resultaten beknopt en systematisch worden weergegeven en waar mogelijk visueel worden ondersteund. In de conclusie en discussie worden de onderzoeksvragen beantwoord en worden mogelijke aanbevelingen gedaan met betrekking tot de items van MHC-SF. Ook wordt gereflecteerd op de sterke en zwakke punten van dit onderzoek en worden aanbevelingen voor vervolgonderzoek gedaan.

Methode

Design

Binnen dit onderzoek wordt een secundaire data-analyse uitgevoerd. Het betreft een correlationeel onderzoek, gericht op de schaalbaarheid van het MHC-SF. Er wordt gebruik gemaakt van bestaande data van het LISS panel. LISS is een afkorting van Langlopende Internet Studies voor de Sociale wetenschappen. In het LISS panel zitten personen die een representatieve afspiegeling vormen van alle lagen van de bevolking van Nederland. Panelleden vullen online vragenlijsten in ten behoeve van onderzoeken. Er zijn naar schatting 5000 Nederlandse huishoudens die in het LISS panel participeren. Deze huishoudens worden door Centerdata en het Centraal Bureau voor de Statistiek (CBS) uitgekozen (LISS panel, 2024) en vertegenwoordigen op basis van hun adres en telefoonnummer een belangrijk deel van de Nederlandse bevolking (LISS panel, 2024a). Het panel is probability-based, wat betekent dat het gebaseerd is op een reële kanssteekproef (LISS Panel - Centerdata NL, 2022). Er is dus sprake van gerandomiseerde selectie.

Ethische kwesties

Er is gebruikgemaakt van vrij toegankelijke open data, waarin persoonsgegevens gepseudonimiseerd zijn. Via de website www.lissdata.nl kunnen geplubliceerde datasets worden gedownload voor wetenschappelijke doeleinden, na het tekenen van een gegevensverklaring. Stichting Centerdata is de instantie die verantwoordelijk is voor de verwerking, opslag en beveiliging van de door LISS Panel opgehaalde data. De voor dit onderzoek gebruikte dataset is volledig geanonimiseerd, wat maakt dat de data uit het databestand waarmee is gewerkt voor deze thesis niet is terug te voeren naar participanten, waardoor de privacy van de participanten geborgd is. In 2020 is door onderzoekers van de Universiteit Twente (UT) de door ons gebruikte data verzameld. Daar is vooraf toestemming voor aangevraagd en deze is afgegeven door de

ethische commissie van de UT. De door ons gedownload datasets zijn bewaard op een harde schijf die beheerd wordt door de Rijksuniversiteit Groningen. Bovenstaande opgesomd is reden geweest om voorafgaand aan dit onderzoek geen aanvraag in te dienen bij de ethische commissie van de faculteit Gedrags- en Maatschappijwetenschappen van de Rijksuniversiteit Groningen. In de privacyverklaring (Privacy - Centerdata NL, 2022) staan de rechten van participanten beschreven. De procedure voor het benaderen en informeren van deelnemers staat beschreven op de website (Ethics - LISS Panel, 2024), waar beschreven staat dat participanten enkel kunnen deelnemen als ze tijdens de procedure om lid te worden van het Liss Panel meerdere keren instemmen met de voorwaarden en expliciete geïnformeerde toestemming geven voor het gebruik van de verzamelde gegevens. Centerdata hecht waarde aan zorgvuldige behandeling van onderzoeksgegevens die zijn verzameld via het LISS-panel. Om de privacy en integriteit van panelleden te beschermen, ondergaat elke vragenlijst een interne beoordeling, waarbij zowel het taalgebruik als de inhoud centraal staan, voordat deze aan het panel wordt voorgelegd. Bij het indienen van vragen die door respondenten als bedreigend, beledigend of zeer onaangenaam kunnen worden ervaren, behoudt Centerdata het recht om deze vragen niet aan panelleden voor te leggen, evenals voor vragen die het risico met zich meebrengen dat panelleden worden geïdentificeerd. Dit interne beoordelingsproces staat los van een ethische beoordeling die wordt uitgevoerd door een (externe) ethische commissie (Ethics - LISS Panel, 2024).

Populatie, steekproef en procedure

In mei 2020 zijn 3571 LISS panelleden benaderd voor het invullen van de vragenlijst. De panelleden zijn op basis van hun adres en telefoonnummer gerandomiseerd geselecteerde mensen uit alle lagen van de populatie van de Nederlandse bevolking. De data van 2724 participanten is door Centerdata in 2020 verzameld, waarmee de lijst door 76,3 procent van de benaderde panelleden is beantwoord. De data is verzameld door het afnemen van online vragenlijsten, aangeboden aan panelleden van het LISS panel (Centerdata, 2021). De panelleden kregen een vergoeding voor het invullen van de vragenlijst, waarbij uitgegaan wordt van de gemiddelde tijd die een panellid nodig is om de vragenlijst in te vullen (LISS panel, 2024a). Hiernaast worden panelleden die anders niet zouden kunnen deelnemen voorzien van een computer en internet (LISS Panel - Centerdata NL, 2022). Bij 5 deelnemers was sprake van het ontbreken van data. Deze zijn, omdat MSA niet kan worden uitgevoerd met ontbrekende data, niet meegenomen in dit onderzoek. Het totaal aantal participanten met volledige data komt uit op

2719. Deze groep is opgedeeld in vier groepen voor het invullen van vier varianten van de MHC-SF. De groep die de door ons geanalyseerde versie van de vragenlijst heeft ingevuld bestaat daardoor uit 650 respondenten. Deze groep is op te delen in 8 groepen, geselecteerd op leeftijd en geslacht, waarbij de leeftijdsgroepen van 18 t/m 29 jaar, van 30 t/m 49 jaar, van 50 t/m 64 jaar en gelijk aan of ouder dan 65 jaar zijn. Omdat binnen ons onderzoek de leeftijdsgroep 18 t/m 24 jaar onderzocht wordt, bestaat deze groep uit minder respondenten ($N=86$) dan de groep van 65 jaar en ouder ($N=152$). Binnen de leeftijdsgroepen zijn de mannen en vrouwen apart gegroepeerd, resulterend in 8 ongeveer even grote groepen.

Variabelen en instrumenten

De Nederlandse versie van de MHC-SF is een instrument waarmee door middel van 14 items de latente variabele mentale gezondheid wordt gemeten. Elk item is een vervolg op het begin van de vraag “In de afgelopen maand, hoe vaak had u het gevoel...”. Het instrument bestaat uit drie subschalen. Item 1 t/m 3 geven informatie over emotioneel welbevinden ((1) gelukkig zijn;(2) levensinteresse;(3) tevredenheid) item 4 t/m 8 over sociaal welbevinden ((4) bijdragen aan de samenleving;(5) deel zijn van de maatschappij ;(6) verbetert samenleving voor mensen;(7) oordeel goedheid mens;(8) begrip hoe maatschappij werkt) en item 9 t/m 14 ((9) eigenwaarde;(10) verantwoordelijkheden;(11) relaties;(12) stimulering persoonlijke groei;(13) zelfverzekerdheid ;(14) betekenisgeving) over psychisch welbevinden. De antwoorden die gegeven kunnen worden binnen een 6-keuze Likert-schaal zijn: (0) nooit, (1) één of twee keer per maand, (2) ongeveer 1 keer per week, (3) 2 of 3 keer per week, (4) bijna elke dag of (5) elke dag. Voor alle items geldt dat een hogere score leidt tot een hogere totaalscore op de latente variabele mentaal welbevinden, waarvan de somscore van alle items in theorie zou kunnen variëren van 0 tot 70. Meerdere zeer recente onderzoeken vinden steun voor goede betrouwbaarheid en validiteit in uiteenlopende populaties. Zo laat onderzoek onder 400 Iraanse adolescenten een hoge betrouwbaarheid (Cronbachs Alfa (α)=0,88 en McDonalds Omega(ω)=0,84) en een acceptabele validiteit zien, zowel convergent als divergent (Khazaei et al., 2022). Ander recent onderzoek onder 3669 adolescenten in Zweden voert bewijs aan voor een hoge betrouwbaarheid met $\alpha=0,92$ en $\omega=0,88$ voor de totale schaal (Söderqvist & Larm, 2021). Een representatieve steekproef van de Nederlandse bevolking wijst op een hoge betrouwbaarheid ($\alpha=0,91$) en ondersteuning van constructvaliditeit (De beurs et al., 2022).

Analyseplan

In dit onderzoek is een exploratieve Mokkenschaalanalyse uitgevoerd met behulp van het mokken pakket in R (R Core Team, 2024). Er zijn twee opties binnen het mokken pakket voor een dergelijke analyse. Wij hebben gekozen voor het meest gebruikte en meest onderzochte algoritme: de Automated Item Selection Procedure (AISP). Dit algoritme stelt schalen samen op een iteratieve wijze. Kort gezegd zoekt het AISP items bij elkaar die sterk met elkaar samen hangen en aan een aantal voorwaarden voldoen (zie Sijtsma & van der Ark (2016) voor meer details). Een van die voorwaarden is dat de samenhang tussen items sterker moet zijn dan een vooraf gespecificeerde waarde (c). Als er geen items meer zijn die sterk met elkaar samen hangen en aan de voorwaarden voldoen, wordt een tweede schaal opgestart en begint het proces opnieuw. Dit net zo lang tot er geen items meer aan schalen kunnen worden toegevoegd. In dit onderzoek zijn we met name benieuwd welke schaal of schalen worden gevonden wanneer de waarde van c toeneemt. Samenhang binnen kader van MSA wordt gemeten met de H -waarde. Deze kan worden berekend op itempaar (H_{ij}), item (H_i) en schaalniveau (H). De H -waarde kan variëren van 0 tot 1. Een H -waarde tussen .3 en .4 staat hierbij voor een zwakke (maar nog steeds acceptabele) schaal, een H -waarde van .4 tot .5 voor een matige en een H -waarde van .5 tot 1 voor een sterke schaal (Sijtsma & Van Der Ark, 2016). We beginnen de procedure met een c -waarde van 0,2 en verhogen deze waarde bij herhaling van de procedure met 0,05. We herhalen de procedure tot er geen aanleiding is om aan te nemen dat dit tot nieuwe inzichten zal leiden met betrekking tot de gevonden schalen.

Om resultaten van een MSA serieus te kunnen nemen, moet er voldaan worden aan een aantal assumpties, namelijk unidimensionaliteit, monotoniciteit en lokale onafhankelijkheid. Wanneer alle items binnen dezelfde schaal vallen binnen de AISP bij een waarde van c gelijk aan of groter dan 0,3, dan is het aannemelijk dat aan de assumptie voor unidimensionaliteit is voldaan, omdat alle items dan in voldoende mate dezelfde latente variabele lijken te meten. Monotoniciteit wil zeggen dat er een aanhoudend verband is tussen totaalscore en itemscore. Wanneer de totaalscore toeneemt moet de kans op een hogere itemscore toenemen of in ieder geval gelijk blijven. We onderzoeken dit door middel van een monotoniciteits-check (Van Der Ark, 2007). We inspecteren de item-plots visueel en onderzoeken of lokale schendingen van monotoniciteit ook statistische significant zijn, door te kijken naar de Crit-coëfficiënt. De Crit-waarde kwantificeert de mate waarin monotoniciteit per item geschonden wordt, waarbij een waarde beneden de 40 indiceert dat er geen sterke aanwijzingen zijn voor schendingen. Bij een

waarde tussen 40 en 80 is het bewijs voor schendingen onduidelijk. Een waarde boven de 80 maakt aannemelijk dat het item de assumptie schendt (Crisan et al., 2021). Lokale onafhankelijkheid stelt dat items onafhankelijk van elkaar zijn wanneer je ze los zou zien van de latente variabele, waardoor bij een selectie van mensen met dezelfde uitkomst op de latente variabele de relatie tussen items verdwijnt. We onderzoeken dit door de conditional association procedure te volgen, waarmee items die lokale afhankelijkheid tonen worden opgespoord (zie Straat et al., 2016 voor meer details). We herhalen de volledige procedure zoals hierboven beschreven voor de subgroepen met participanten in de leeftijd van (1) 18 t/m 24 jaar en (2) 65 jaar en ouder, om zodoende te onderzoeken of de compositie van de schalen en de kenmerken van de schalen grote mate van overeenkomst vertonen en daarmee te onderzoeken in hoeverre er aanleiding is om te veronderstellen dat er mogelijk sprake is van een effect van leeftijd op schaalbaarheid. Daarnaast willen we weten of de resultaten voor subgroepen aanleiding geven om te veronderstellen dat aan alle assumpties is voldaan.

Voor de volledigheid starten we met een beschrijving van meer algemene statistische gegevens van de data en berekenen we ook Cronbachs Alfa als maat voor betrouwbaarheid.

Resultaten

Beschrijvende statistiek

Per item is het gemiddelde, de standaardafwijking en de mediaan van de itemscore berekend voor de gehele groep en de leeftijdsgroepen van 18 t/m 24 jaar (hierna adolescenten genoemd) en 65 jaar en ouder (hierna ouderen genoemd) (zie tabel 1). Aangaande het aantal respondenten wordt zichtbaar dat de vragenlijst door 650 respondenten is beantwoord, waarvan 86 adolescenten en 152 ouderen. De laagste gemiddelde itemscore voor de hele groep is 1,74, waarbij de mediaan 2 is (item 6). De hoogste gemiddelde itemscore is 3,79, waarbij de mediaan 4 is (item 2). Ook binnen de subgroepen wordt er gemiddeld het hoogst gescoord op item 2 en het laagst op item 6. Binnen de steekproef denken respondenten aanzienlijk minder vaak positief over de manier waarop de samenleving zich ontwikkelt dan over de mate waarin ze het gevoel hadden interesse in het leven te hebben. Voor de gehele groep geldt dat binnen item 3 (tevredenheid), met een standaarddeviatie van 1,03, antwoorden gemiddeld het minst afwijken van de gemiddelde score. Item 12 (stimulering persoonlijke groei) heeft met 1,56 de hoogste standaarddeviatie en dus de hoogste gemiddelde afwijking ten opzichte van het gemiddelde. De totale score ligt bij ouderen met 39,78 lager dan de totale score van de gehele groep, die uitkomt

op 41,31. De groep adolescenten scoort gemiddeld iets hoger dan de gehele groep met 41,64. Ouderen laten in de totaalscores gemiddeld een hogere spreiding zien met een standaardafwijking van 13,66. Opvallend is dat binnen item 12 adolescenten veel hoger scoren dan ouderen (1,80 tegenover 2,72), waarbij de mediaan bij adolescenten op 3 ligt en bij ouderen op 1. Ook binnen items 5, 9 en 13 zien we dat adolescenten gemiddeld aanmerkelijk hoger scoren dan ouderen ($>0,2$). Bij item 2, 3 en 10 zien we juist dat ouderen gemiddeld aanmerkelijk hoger scoren dan adolescenten.

Tabel 1

Beschrijving van de statistische kenmerken van de items.

	Alle leeftijden (N=650)			18 t/m 24 jaar (n=86)			> 65 jaar (n=152)		
Item	gem.	st.afw.	med.	gem.	st.afw.	med.	gem.	st.afw.	med.
1	3,42	1,14	4	3,47	1,12	4	3,44	1,23	4
2	3,79	1,13	4	3,72	1,15	4	3,94	1,21	4
3	3,68	1,03	4	3,59	1,08	4	3,88	1,02	4
4	2,38	1,46	3	2,27	1,36	3	2,16	1,52	2
5	2,19	1,58	2	2,33	1,59	2	1,91	1,62	2
6	1,74	1,39	2	1,85	1,38	2	1,71	1,40	1,5
7	2,64	1,37	3	2,72	1,31	3	2,61	1,44	3
8	2,75	1,41	3	2,71	1,43	3	2,59	1,43	3
9	3,06	1,31	3	3,14	1,27	3	2,80	1,42	3
10	3,69	1,14	4	3,41	1,31	4	3,65	1,26	4
11	3,58	1,25	4	3,64	1,23	4	3,51	1,30	4
12	2,35	1,56	2	2,72	1,48	3	1,8	1,70	1
13	2,94	1,31	3	2,98	1,23	3	2,68	1,53	3
14	3,10	1,45	3	3,10	1,32	3	3,09	1,64	4
Totaal	41,31	12,34		41,64	12,08		39,78	13,66	

Noot: gem. = gemiddelde score, st.afw. = standaardafwijking, med. = mediaan

Hoofdbevindingen: Dimensionaliteit

Bij het uitvoeren van de exploratieve analyse bleek dat binnen de groep met jongeren op item 2 door niemand het antwoord 0 is gegeven. Het is daardoor niet mogelijk om dit item binnen ons onderzoek mee te nemen. Daarom is de analyse zonder item 2 uitgevoerd. De exploratieve

analyse (tabel 2) laat zien dat tot een c-waarde van 0,30 alle items binnen dezelfde schaal vallen. Bij een c-waarde van 0,35 vormen item 6 en 7 bij de adolescenten een tweede schaal en valt item 12 buiten de schaal. Voor de ouderen en alle participanten samen, vallen alle items binnen dezelfde schaal. Vanaf een c-waarde van 0,40 vormen item 6, 7 en 8 een tweede schaal voor de gehele steekproef. Bij de adolescenten blijft de schaalbaarheid ten opzichte van 0,35 gelijk. Bij de ouderen zien we nog steeds alle items in dezelfde schaal. Bij een c-waarde van 0,45 zien we voor de volledige steekproef dat item 6, 7 en 8 de tweede schaal blijven vormen, dat item 4 en 5 een derde schaal vormen en dat item 12 buiten de schaal valt. Buiten item 12 worden alle items die het emotionele en psychische welzijn meten in dezelfde schaal geplaatst. Bij de ouderen geldt dit ook nog bij een c-waarde van 0,5. De items die het sociale welzijn meten vallen niet in dezelfde schaal als de andere items en vallen op zichzelf ook niet in één schaal, maar in twee aparte schalen. De schaalindeling die gevonden wordt bij de exploratieve analyse duidt bij de totale groep en de leeftijdsgroepen steeds op unidimensionaliteit en de zichtbare patronen geven geen aanleiding om aan te nemen dat een ander in de theorie beschreven model leidt tot een betere schaalbaarheid. Dit wordt kracht bijgezet als je kijkt naar de H-waarde voor de totale schaal voor alle onderzochte groepen, die in alle gevallen groter dan 0,40 is.

Tabel 2

Uitkomsten exploratieve analyse

Item	c-waarde alle leeftijden					c-waarde 18 t/m 24 jaar					c-waarde > 65 jaar				
	0,30	0,35	0,45	0,45	0,5	0,3	0,35	0,4	0,45	0,5	0,3	0,35	0,4	0,45	0,5
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
4	1	1	2	3	3	1	1	1	2	2	1	1	1	2	3
5	1	1	2	3	3	1	1	1	2	2	1	1	1	2	3
6	1	1	2	2	2	1	2	2	2	2	1	1	1	1	2
7	1	1	1	2	2	1	2	2	0	0	1	1	1	1	2
8	1	1	1	2	0	1	1	1	2	0	1	1	1	1	0
9	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
12	1	1	1	0	0	1	0	0	0	0	1	1	1	1	0

13	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1

Assumptie checks

Bij het checken van de assumpties zijn we op basis van de bovenstaande analyse steeds uitgegaan van 1 schaal.

Monotoniciteit

Een visuele analyse van de Item-response curves toont binnen de gehele steekproef lichte schendingen van Monotoniciteit voor de items 1, 5 en 10. Bij de adolescenten en de ouderen zien we vanuit visuele analyse geen schendingen. Uit de statistische analyse met betrekking tot monotoniciteit (Tabel 3) blijken binnen de gehele steekproef 2 schendingen op basis van de z-waarde significant. Deze schendingen zijn geconstateerd voor items 5 en 8. Binnen de subgroepen van adolescenten en ouderen zien we geen schendingen. De hoogst gemeten Crit-coëfficiënt is binnen de gehele steekproef is 33 (item 5). Bij de adolescenten is de hoogst gemeten Crit-coëfficiënt 39, bij item 6. Het is ook opvallend dat item 6 met 0,33 de laagste item-schaalbaarheid coëfficiënt onder adolescenten heeft, waar bij de ouderen dit 0,48 is. Bij de ouderen wordt voor alle items een Crit-coefficient van 0 gemeten. Waarden beneden de 40 indiceren dat er geen sterke aanwijzingen voor schendingen zijn. De gevonden waarden voor de H-waarde op itemniveau (H_i) zijn allen boven de 0,4. Deze uitkomsten maken aannemelijk dat er sprake is van een matige schaalbaarheid en maken aannemelijk dat de items voldoen aan de assumptie van monotoniciteit.

Tabel 3

Kenmerken met betrekking tot monotoniciteit van de items

Item	Gehele steekproef (N=650)			18 t/m 24 jaar (n=86)			65 jaar en ouder (n=152)		
	H_i	crit	#zsig	H_i	crit	#zsig	H_i	crit	#zsig
1	0,4	23	0	0,39	14	0	0,46	0	0
3	0,42	9	0	0,47	0	0	0,47	0	0
4	0,4	13	0	0,41	0	0	0,44	0	0
5	0,41	33	1	0,46	0	0	0,44	0	0
6	0,4	0	0	0,33	39	0	0,48	0	0
7	0,42	3	0	0,35	0	0	0,46	0	0

8	0,41	28	1	0,43	0	0	0,47	0	0
9	0,5	0	0	0,49	0	0	0,55	0	0
10	0,43	12	0	0,42	0	0	0,54	0	0
11	0,47	7	0	0,47	0	0	0,55	0	0
12	0,4	13	0	0,34	0	0	0,45	0	0
13	0,47	8	0	0,54	0	0	0,48	0	0
14	0,49	0	0	0,5	13	0	0,53	0	0

Noot: Hi = H-waarde van het item, crit = Crit-coefficient, #zsig = aantal significante overschrijdingen

Lokale onafhankelijkheid

De Conditional Association Procedure laat binnen de gehele steekproef zien dat de conditionele covariantie tussen itempaar 6 en 9 groter is dan de kritieke waarde, waardoor item 6 gemarkeerd wordt. Uit de verdere analyse blijkt dat er op grond van de gevonden waarden geen redenen zijn om aan te nemen dat item 6 betrokken is bij meerdere afhankelijke paren. Er zijn geen aanwijzingen voor structurele afhankelijkheid. Onder de subgroepen van adolescenten en ouderen zien we geen gemarkeerde items. De resultaten geven onvoldoende aanleiding om te veronderstellen dat niet aan de assumptie van lokale afhankelijkheid is voldaan.

Betrouwbaarheid

Alfa als maat waarmee een inschatting van de betrouwbaarheid wordt gemaakt komt uit op 0,90, waarmee de betrouwbaarheid van het instrument volgens het COTAN beoordelingssysteem, ook voor een test die wordt gebruikt voor belangrijke beslissingen op individueel niveau, goed is. Een waarde van 0,9 of hoger wordt hierbij gezien als een vereiste.

Conclusie en discussie

In dit onderzoek naar de dimensionaliteit van de MHC-SF in de Nederlandse bevolking, is steun gevonden voor de unidimensionaliteit van dit instrument. Selectie op leeftijd toonde, met name op itemniveau, wel wat verschillende waarden tussen de leeftijdsgroepen, maar steeds werd steun gevonden voor unidimensionaliteit. Ook voor de leeftijdsgroepen wijzen de resultaten niet op meer ondersteuning voor een ander in de literatuur beschreven model dan het unidimensionele model.

Opvallend is dat de literatuur met name steun vindt voor een 3-factor model, die vaak een betere fit toonde dan het unidimensionele model, maar Iasiello en collegae (2022) wijzen op het

feit dat de MHC-SF is ontworpen om het breed multidimensionele concept van mentale gezondheid meetbaar te maken. Uit onderzoeken blijkt dat dit er vaak voor zorgt dat binnen factoranalyse een slechtere fit voor de gehele schaal gevonden wordt (Iongo et al., 2020; Reise & Revicki, 2014). Sellbom en Tellegen (2019) waarschuwen naar aanleiding van hun onderzoek naar factoranalyse bij psychologische vragenlijsten voor het te zwaar wegen van het belang van modelfit en geven aan dat ook vooral gekeken moet worden naar de theoretische onderbouwing om te bepalen welk model gebruikt wordt. Een iets hogere H-waarde voor een specifiek model ten opzichte van een ander model zou wat dat betreft ook geen aanleiding hoeven te zijn om een keuze te maken tussen deze modellen. Ze kunnen naast elkaar gebruikt worden.

Als we de bevindingen van dit onderzoek vergelijken met het onderzoek van Goosen (2025), die een confirmatieve analyse heeft uitgevoerd voor de MHC-SF met MSA waar hij uitgang van 3 subschalen, zien we verschillen in conclusies. We beargumenteren dat deze twee onderzoeken echter niet als tegenstrijdig maar als aanvullend kunnen worden beschouwd. Als we de resultaten van deze twee onderzoeken samennemen, kunnen we stellen dat zowel voor het unidimensionele model als voor het 3 factoren model steun gevonden wordt en zijn dus, wanneer ze op de juiste manier ingezet worden, beiden in te zetten. In de aanbevelingen komen we hier nog op terug.

Tijdens het onderzoek zijn we tegen enkele beperkingen aangelopen. Zo bleek item 2 door geen van de adolescenten met 0 te zijn beantwoord, waardoor we dit item hebben moeten excluderen bij de uitvoering van MSA. Er zijn geen redenen om aan te nemen dat het includeren van item 2 tot andere conclusies geleid zou hebben, maar het is met de door ons gebruikte gegevens niet hard te maken. Er is geen ander onderzoek gevonden wat melding heeft gemaakt van niet gekozen antwoordmogelijkheden bij de MHC-SF, maar wel wordt, bijvoorbeeld in een grootschalig Canadees onderzoek afgenomen onder jongeren, melding gemaakt van hoge gemiddelde scores, waarbij de hoogst gemeten gemiddelden werden gevonden bij item 2 en 11 (Orpana et al., 2017). Binnen het onderzoek naar Nederlandse adolescenten van Luijten en collegae (2019) wordt binnen de subschaal emotioneel welbevinden (item 1, 2 en 3) met een waarde van 3,88 (op een schaal van 0 tot 5) een hoog gemiddelde en een relatief lage standaardafwijking van 0,94 gevonden. Dit maakt het, uitgaand van een normale verdeling, niet onwaarschijnlijk dat er weinig of geen participanten zijn met een itemscore van 0. Omdat binnen ander onderzoek naar de psychometrische kwaliteiten vaak gebruik is gemaakt van factoranalyse

is het niet ondenkbaar dat geen melding is gemaakt van niet gekozen antwoordmogelijkheden, waar dit het resultaat wel heeft kunnen beïnvloeden.

Doordat ten tijde van de dataverzameling gekozen is om de participanten te verdelen over de vier varianten van de vragenlijst, heeft dit onderzoek zich moeten beperken tot 650 participanten, waarvan er slechts 86 tussen de 18 en 24 jaar waren. Voor vervolgonderzoek verdient het aanbeveling om een grotere steekproef te onderzoeken, om zo de kans op een niet gescoorde categorie binnen een item te verkleinen en de betrouwbaarheid van de analyse te vergroten.

De kracht van MSA is dat aan minder assumpties voldaan moet worden dan bijvoorbeeld de in veel onderzoeken aangehaalde factoranalyse. In dit opzicht vormt MSA een waardevolle aanvulling op de veelvuldig uitgevoerde factoranalyses. Wanneer aantoonbaar is dat aan alle assumpties voldaan is, zoals in ons onderzoek, levert dit een waardevol aanvullend inzicht met betrekking tot de constructvaliditeit van de MHC-SF op.

Samenvattend laat ons onderzoek zien dat de MHC-SF unidimensioneel te gebruiken is. Met betrekking tot het gebruik van de MHC-SF is onze aanbeveling vooral te bedenken wat je met welk doeleinde wilt meten en op grond daarvan te bepalen of je gebruik maakt van subschalen, waar binnen het onderzoek van Goosen (2025) de hoogste schaalbaarheid gemeten wordt, of alleen de totaalscore gebruikt. Het unidimensionele model zou bijvoorbeeld bij grootschalige screening gebruikt kunnen worden om een globale indruk te krijgen van het welzijnsniveau en om individuen te identificeren, maar ook voor globale welzijnsmeting in onderzoek naar de effectiviteit van brede interventies gericht op welzijn. Het 3 factoren model zou juist ingezet kunnen worden voor het maken van meer gedetailleerde welzijnsprofielen van individuen, bij het monitoren van gerichte interventies op een specifiek welzijnsdomein, maar ook bij het classificeren op de door Keyes (2005) beschreven manier, waarop mensen in een categorie van mentaal welbevinden worden geplaatst. Het verdient aanbeveling om meer onderzoek naar de MHC-SF aan de hand van exploratieve en confirmatieve MSA te doen, vooral om te kijken of de bevindingen consistent zijn wanneer de populatie (bijvoorbeeld cultureel, klinisch / niet-klinisch of geselecteerd op specifieke kenmerken als depressie of schizofrenie) verandert. Mocht dit leiden tot vergelijkbare resultaten, dan zal dit de aanname dat zowel een unidimensioneel als een 3 factoren model acceptabele schaalbaarheid laat zien en dat beide toepassingen dus mogelijk zijn versterken. Vervolgonderzoek zou zich dan meer kunnen richten

op de manieren waarop de vragenlijst in verschillende praktijksituaties het best gebruikt zou kunnen worden.

Literatuurlijst

- Arslan, I., Distelbrink, M., Yassine, D., & Verwey-Jonker Instituut. (2024). Mentale gezondheid van jongeren en studenten in Nederland: Een overzicht van trends en verklarende factoren. https://www.verwey-jonker.nl/wp-content/uploads/2024/04/223460_Mentale-gezondheid.pdf
- de Beurs, E., Kosterman, S., Anten, S., Bohlmeijer, E., & Westerhof, G. (2022). Psychometrische evaluatie van de Mental Health Continuum—Short Form (MHC-SF): Constructvaliditeit, responsiviteit voor verandering, normen en T-scores. *Gedragstherapie*, 55(2), 131–155.
- Breedvelt, J. J. F., Zamperoni, V., South, E., Uphoff, E. P., Gilbody, S., Bockting, C. L. H., Churchill, R., & Kousoulis, A. A. (2019). A systematic review of mental health measurement scales for evaluating the effects of mental health prevention interventions. *European Journal Of Public Health*, 30(3), 510–516. <https://doi.org/10.1093/eurpub/ckz233>
- Crişan, D. R., Tendeiro, J. N., & Meijer, R. R. (2021). The Crit coefficient in Mokken scale analysis: a simulation study and an application in quality-of-life research. *Quality Of Life Research*, 31(1), 49–59. <https://doi.org/10.1007/s11136-021-02924-z>
- Centerdata. (2021). *The Dutch Mental Health Continuum Short Form – Revised*. LISS Panel. <https://doi.org/10.17026/dans-25d-n6kr>
- Ethics - LISS panel. (2024, 22 juli). LISS Panel. <https://www.lissdata.nl/ethics>
- Evers, A., Lucassen, W., Meijer, R., Sijtsma, K., & NIP. (2009). *COTAN beoordelingssysteem voor de kwaliteit van tests (geheel herziene versie)*. https://pure.uva.nl/ws/files/896125/79346_312828.pdf
- Goldberg, D. P. (1972). The detection of psychiatric illness by questionnaire: A technique for the identification and assessment of non-psychotic psychiatric illness. Oxford U. Press.
- Goodman, F. R., Doorley, J. D., & Kashdan, T. B. (2018). Well-being and psychopathology: A deep exploration into positive emotions, meaning and purpose in life, and social relationships. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of well-being*. Salt Lake City, UT: DEF Publishers. DOI:nobascholar.com

- Goosen, D. (2025). *Psychometrische eigenschappen van de Nederlandse Mental Health continuum-short form (MHC-SF) na aanpassing van de vragenlijst* [Masterthese]. Rijksuniversiteit Groningen.
- Iasiello, M., Van Agteren, J., Schotanus-Dijkstra, M., Lo, L., Fassnacht, D. B., & Westerhof, G. J. (2022b). Assessing mental wellbeing using the Mental Health Continuum—Short Form: A systematic review and meta-analytic structural equation modelling. *Clinical Psychology Science And Practice*, 29(4), 442–456. <https://doi.org/10.1037/cps0000074>
- Kapenberg -ten Holt, J. (2014). A comparison between factor analysis and item response theory modeling in scale analysis. [Thesis fully internal (DIV), University of Groningen]. s.n.
- Kennes, A., Peeters, S., Janssens, M., Reijnders, J., Lataster, J., & Jacobs, N. (2020). Psychometric Evaluation of the Mental Health Continuum-Short Form (MHC-SF) for Dutch Adolescents. *Journal Of Child And Family Studies*, 29(11), 3276–3286. <https://doi.org/10.1007/s10826-020-01803-4>
- Kennisrotonde (2024). Wat zijn betrouwbare en valide methoden om het welbevinden van vmbo-leerlingen te monitoren?(KR. 2251) https://www.kennisrotonde.nl/sites/kennisrotonde/files/media-files/pdf_voor_website-kennisrotonde-antwoord_vraag-2251.pdf
- Keyes, C. L. M. (2002). The Mental Health Continuum: From Languishing to Flourishing in Life. *Journal Of Health And Social Behavior*, 43(2), 207. <https://doi.org/10.2307/3090197>
- Keyes, C. L. M. (2005). Mental Illness and/or Mental Health? Investigating Axioms of the Complete State Model of Health. *Journal of Consulting and Clinical Psychology*, 73(3), 539–548. <https://doi-org.proxy-ub.rug.nl/https://content.apa.org/journals/ccp/73/3>
- Keyes, C. L. M. (2018). Overview of The Mental Health Continuum Short Form (MHC-SF). 10.13140/RG.2.2.24204.62088.
- Khazaei, M., Holder, M. D., Sirois, F. M., Oades, L. G., & Gallagher, M. W. (2022). Evaluating the psychometric properties of the Mental Health Continuum Short Form (MHC-SF) in Iranian adolescents. *Current Psychology*, 42(21), 17995–18009. <https://doi.org/10.1007/s12144-022-02970-x>
- Lamers, S. M., Westerhof, G. J., Bohlmeijer, E. T., Klooster, P. M. T., & Keyes, C. L. (2010). Evaluating the psychometric properties of the mental health Continuum-Short Form

- (MHC-SF). *Journal Of Clinical Psychology*, 67(1), 99–110.
<https://doi.org/10.1002/jclp.20741>
- LISS panel. (2024, 17 september). Home - LISS panel. LISS Panel.
<https://www.website.lisspanel.nl/>
- LISS panel. (2024a, juni 20). Veelgestelde vragen - LISS panel. LISS Panel.
<https://www.website.lisspanel.nl/veelgestelde-vragen/>
- LISS panel - Centerdata NL. (2022, 5 september). Centerdata NL. <https://www.centerdata.nl/liss-panel>
- Longo, Y., Jovanović, V., De Carvalho, J. S., & Karaš, D. (2017). The General Factor of Well-Being: Multinational Evidence Using Bifactor ESEM on the Mental Health Continuum—Short Form. *Assessment*, 27(3), 596–606. <https://doi.org/10.1177/1073191117748394>
- Lovibond, P., & Lovibond, S. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research And Therapy*, 33(3), 335–343.
[https://doi.org/10.1016/0005-7967\(94\)00075-u](https://doi.org/10.1016/0005-7967(94)00075-u)
- Luijten, C. C., Kuppens, S., Van de Bongardt, D., & Nieboer, A. P. (2019). Evaluating the psychometric properties of the mental health continuum-short form (MHC-SF) in Dutch adolescents. *Health And Quality Of Life Outcomes*, 17(1).
<https://doi.org/10.1186/s12955-019-1221-y>
- Magyar, J. L., & Keyes, C. L. M. (2019). Defining, measuring, and applying subjective well-being. In *American Psychological Association eBooks* (pp. 389–415).
<https://doi.org/10.1037/0000138-025>
- Mokken, R. J. (1971). A Theory and Procedure of Scale Analysis. In *De Gruyter eBooks*.
<https://doi.org/10.1515/9783110813203>
- Nisha. (2024, 14 mei). De 10 meest gebruikte GGZ-vragenlijsten in ons vragenlijstenportaal. e-Health van Therapieland. <https://therapieland.nl/nieuws/de-10-meest-gebruikte-vragenlijsten-voor-de-ggz/>
- Orpana, H., Vachon, J., Dykxhoorn, J., & Jayaraman, G. (2017). Measuring positive mental health in Canada: construct validation of the Mental Health Continuum—Short Form. *Health Promotion And Chronic Disease Prevention in Canada*, 37(4), 123–130.
<https://doi.org/10.24095/hpcdp.37.4.03>

- Privacy - Centerdata NL. (2022, 31 oktober). Centerdata NL. <https://www.centerdata.nl/privacy-statement>
- R Core Team (2024). *_R: A Language and Environment for Statistical Computing_*. R Foundation for Statistical Computing, Wenen, Oostenrijk. <https://www.R-project.org/>
- Reise, S. P., & Revicki, D. A. (2014). Handbook of Item Response Theory Modeling. In Routledge eBooks. <https://doi.org/10.4324/9781315736013>
- RIVM (2023). Levenstevredenheid. <https://www.rivm.nl/mentale-gezondheid/monitor/algemene-bevolking/levenstevredenheid>
- Straat, J. H., Van Der Ark, L. A., & Sijtsma, K. (2016). Using Conditional Association to Identify Locally Independent Item Sets. *Methodology*, 12(4), 117–123. <https://doi.org/10.1027/1614-2241/a000115>
- Universiteit van Amsterdam. (2023, 12 september). Levi van Dam benoemd tot bijzonder hoogleraar Veerkrachtig Opgroeien. <https://www.uva.nl/shared-content/faculteiten/nl/faculteit-der-maatschappij-en-gedragswetenschappen/nieuws/2023/09/hoogleraarbenoeming-levi-van-dam.html>
- Santini, Z. I., Torres-Sahli, M., Hinrichsen, C., Meilstrup, C., Madsen, K. R., Rayce, S. B., Baker, M. M., Have, M. T., Schotanus-Dijkstra, M., & Koushede, V. (2020). Measuring positive mental health and flourishing in Denmark: validation of the mental health continuum-short form (MHC-SF) and cross-cultural comparison across three countries. *Health And Quality Of Life Outcomes*, 18(1). <https://doi.org/10.1186/s12955-020-01546-2>
- Sellbom, M., & Tellegen, A. (2019). Factor analysis in psychological assessment research: Common pitfalls and recommendations. *Psychological Assessment*, 31(12), 1428–1441. <https://doi.org/10.1037/pas0000623>
- Sijtsma, K., & Van Der Ark, L. A. (2016). A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. *British Journal Of Mathematical And Statistical Psychology*, 70(1), 137–158. <https://doi.org/10.1111/bmsp.12078>
- Söderqvist, F., & Larm, P. (2021). Psychometric evaluation of the mental health continuum – short form in Swedish adolescents. *Current Psychology*, 42(3), 2136–2144. <https://doi.org/10.1007/s12144-021-01626-6>

- Universiteit van Amsterdam. (2024, 31 oktober). Mentale gezondheid van jongeren gaat niet achteruit. <https://www.uva.nl/content/nieuws/persberichten/2024/10/mentale-gezondheid-van-jongeren-gaat-niet-achteruit.html?cb>
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: The PANAS scales. *Journal Of Personality And Social Psychology*, 54(6), 1063–1070. <https://doi.org/10.1037/0022-3514.54.6.1063>
- Westerhof, G. J., & Keyes, C. L. M. (2008). Geestelijke gezondheid is meer dan de afwezigheid van geestelijke ziekte (pp. 808–820). <https://ris.utwente.nl/ws/portalfiles/portal/6814957/MGV-oktober-inhoud08.pdf>
- World Health Organization: WHO. (2019, 19 december). Mental health. https://www.who.int/health-topics/mental-health#tab=tab_1
- Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67(6), 361–370. <https://doi.org/10.1111/j.1600-0447.1983.tb09716.x>