

Een statistische analyse van een instrument dat uitdagend oudergedrag meet.

Een kwantitatieve studie naar de psychometrische gegevens van de CPBQ(4-6).

A statistical analysis of an instrument that measures challenging parental behavior.

A quantitative study on the psychometric data of the CPBQ(4-6).

R.J. van der Laan, S3499685

Orthopedagogiek, Faculteit GMW, Rijksuniversiteit Groningen

Rijksuniversiteit Groningen

PAMA5166: Masterthese

Begeleider: Wendy Post

Tweede beoordelaar: Laura Batstra

18 december, 2023

Aantal woorden: 7404

Abstract

In 2020, one in eleven young adults aged 18 to 34 in the Netherlands had a high risk of developing an anxiety disorder. Challenging Parental Behavior (CPB) in the early years of a child reduces the likelihood of developing anxiety disorders later in life. The Challenging Parental Behavior Questionnaire (CPBQ) measures the CPB of parents with young children. This study assesses the psychometric qualities of the CPBQ(4-6) and examines whether the current subdivision of the CPBQ(4-6) into subscales is psychometrically justified. To answer the research question 'to what extent is the current subdivision of the CPBQ(4-6) into subscales psychometrically justified,' a Mokken scale analysis was conducted. Data from the 3HOEK research project on the influence of parental behavior on the development of young children were utilized. The exploratory analysis reveals that the best-performing scale division consists of two scales of ten and four items, respectively. This study provides multiple indications that the CPBQ(4-6) does not measure what it is intended to measure. It can be concluded that no support has been found to psychometrically justify the current subdivision of the CPBQ(4-6). Further investigation into the poorly performing items and scales is needed, and some items may need to be reconsidered or refined to improve the questionnaire's coherence.

Keywords: Anxiety disorder, Challenging Parental Behavior (CPB), CPBQ (4-6), Mokken Scale Analysis (MSA).

Samenvatting

Eén op de elf jongvolwassenen tussen de 18 en 34 jaar in Nederland had in 2020 een hoog risico op het ontwikkelen van een angststoornis. Uitdagend oudergedrag (CPB) in de vroege levensjaren van het kind vermindert de kans op het ontwikkelen van angststoornissen op latere leeftijd. De Challenging Parental Behaviour Questionnaire (CPBQ) meet het CPB van ouders met jonge kinderen. Dit onderzoek brengt de psychometrische kwaliteiten van de CPBQ(4-6) in kaart en kijkt of de huidige opdeling van de CPBQ(4-6) in subschalen psychometrisch te verantwoorden is. Om de onderzoeksvraag ‘in hoeverre is de huidige opdeling van de CPBQ(4-6) in subschalen psychometrisch te verantwoorden’ te kunnen beantwoorden is een mokkenschaalanalyse uitgevoerd. Hiervoor is gebruik gemaakt van gegevens van het 3HOEK-onderzoeksproject naar de invloed van oudergedrag op de ontwikkeling van het jonge kind. Uit de exploratieve analyse blijkt dat de best mogelijk presterende schaalindeling bestaat uit twee schalen van tien en vier items. Onderhavig onderzoek geeft meerdere aanwijzingen dat de CPBQ(4-6) niet meet wat het zou moeten meten. Geconcludeerd kan worden dat er geen steun gevonden is om de huidige opdeling van de CPBQ(4-6) psychometrisch te verantwoorden. Nader onderzoek naar de slechter scorende items en schalen is nodig en sommige items zullen mogelijk moeten worden heroverwogen of verfijnd om de samenhang van de vragenlijst te verbeteren.

Sleutelwoorden: Angststoornis, Uitdagend oudergedrag (CPB), CPBQ (4-6), Mokkenschaalanalyse (MSA).

Inhoud

Abstract.....	2
Samenvatting.....	3
Inhoud	4
Inleiding	5
Methode	9
Onderzoeksdesign	9
Doelpopulatie, toegankelijke populatie en steekproef	9
Onderzoeksvariabelen en instrumenten	9
Procedure	11
Mokkenschaaanalyse	12
Psychometrische analyse	13
Resultaten	15
Conclusie en discussie	23

Inleiding

Uitdagens oudergedrag (CPB) betreft gedrag van ouders waarmee ze hun kinderen stimuleren om persoonlijke grenzen op te zoeken en te verleggen. Onder het begrip CPB valt: het plagen van het kind, wat ruiger en fysiek zijn tijdens het spelen met het kind, het kind stimuleren om dingen die als eng worden ervaren toch te doen en het kind op sociaal-emotioneel vlak uitdagen (Majdandžić et al., 2010). Uit onderzoek blijkt dat vaders vaker CPB vertonen (Grossmann et al., 2002) en dat beide ouders vaker dit gedrag laten zien bij het opvoeden van jongens dan bij het opvoeden van meisjes (Morrongiello & Dawber, 2000).

Uit onderzoek van Lazarus et al. (2016) blijkt dat CPB van ouders in de vroege levensjaren van het kind de kans op het ontwikkelen van angststoornissen op latere leeftijd vermindert. Er wordt van een angststoornis gesproken wanneer angstklachten met piekeren en lichamelijke verschijnselen, zonder duidelijke aanleiding, langer dan zes maanden vrijwel continu aanwezig zijn en ze hinder of last geven in het dagelijkse leven (American Psychiatric Association, 2013). Uit onderzoek blijkt dat CPB van vaders de kans op het ontwikkelen van een angststoornis op symptoom-niveau mogelijk vermindert. CPB van moeders vermindert mogelijk de kans op het ontwikkelen van angststoornissen zowel op symptoomniveau, als diagnosesniveau (Lazarus et al., 2016). Angststoornissen zijn één van de meest voorkomende stoornissen bij jongvolwassenen (Bodden et al., 2008). Uit cijfers van het RIVM (2020) blijkt dat één op de elf jongvolwassenen tussen de 18 en 34 jaar in Nederland in 2020 een hoog risico had op het ontwikkelen van een angststoornis. Dit is een toename van 7% ten opzichte van 2016. De impact die dit heeft op persoonlijk, sociaal en economisch niveau geeft reden tot meer onderzoek naar de factoren die een rol spelen in de preventie van angststoornissen. Omdat CPB een positief effect kan hebben op het voorkomen van angststoornissen is het van belang dat het construct op een valide en betrouwbare manier gemeten kan worden. Op dit moment is er nog niet voldoende onderzoek gedaan naar de instrumenten die dit kunnen bewerkstelligen.

Omdat zowel de negatieve als de positieve factoren van opvoedingsgedrag van ouders invloed kunnen hebben op de ontwikkeling van angststoornissen, is het van belang beide goed in kaart te brengen. In de huidige literatuur is veel aandacht voor de negatieve factoren. Naast genetische factoren worden te beschermende of controlerende en te negatieve ouders vaak gekoppeld aan het ontwikkelen van angststoornissen op latere leeftijd (Rapee et al., 2009). Te beschermende ouders worden in de literatuur beschreven als ouders die hun kinderen fysiek beperken door risicovolle situaties zoveel mogelijk uit de weg te gaan. Daarnaast is vaak sprake van overmatige betrokkenheid en controle, waardoor het kind te

afhankelijk kan worden van de ouder (Kiel & Buss, 2011). Uit onderzoek van Houtepen et al. (2019) blijkt dat overmatige controle kan leiden tot internaliserende problematiek in de adolescentie. Te negatieve ouders worden in de literatuur beschreven als ouders die inconsistent zijn en hun kinderen weinig warmte geven, bijvoorbeeld door afwijzing, kritiek en vijandigheid. Ook bieden te negatieve ouders hun kinderen weinig autonomie en is er sprake van overmatige controle (Gulley et al., 2014). Een negatieve opvoedingsstijl wordt geassocieerd met een grotere kans op het ontwikkelen van angststoornissen (Knappe et al., 2009; Nditch & Varela, 2012).

Ook focussen veel onderzoeken zich op het verband tussen opvoedingsgedrag van ouders en de algehele ontwikkeling van hun kinderen. De wijze en kwaliteit van interacties tussen ouders en hun kinderen kan bijdragen aan het welzijn van het kind (Bandura, 1997). Daarnaast is er in de literatuur veel aandacht voor de vier opvoedingsstijlen van Baumrind (1966) en de mate van sociale support afhankelijk van de opvoedingsstijl (Larzere et al., 2013; Pinquart, 2017). Baumrind onderscheidt vier opvoedingsstijlen: autoritair, autoritatief, permissief en verwaarlozend. Een autoritaire opvoedingsstijl wordt gekenmerkt door hoge niveaus van controle en eisen aan het kind. Ouders leggen strenge regels op en verwachten gehoorzaamheid. Een autoritatieve opvoedingsstijl wordt gekenmerkt door hoge niveaus van responsiviteit en eisen. Ouders hebben duidelijke regels, maar er is ruimte voor discussie en ouders houden rekening met de behoeften van het kind. Een permissieve opvoedingsstijl wordt gekenmerkt door een hoog niveau van responsiviteit, maar lage eisen aan het kind. Ouders zijn toegeeflijk en geven het kind veel vrijheid. Een verwaarlozende opvoedingsstijl wordt gekenmerkt door lage niveaus van responsiviteit en eisen. Ouders tonen weinig betrokkenheid en bieden weinig emotionele steun aan het kind. De opvoedingsstijl van ouders is gecorreleerd aan de kans op het ontwikkelen van angststoornissen in het latere leven. Hoge responsiviteit en warmte vermindert de kans op het ontwikkelen van angststoornissen, daarentegen verhoogt lage responsiviteit en weinig emotionele steun de kans op het ontwikkelen van angststoornissen (Erozkan, 2012).

Voor de bovenstaande factoren die een rol spelen in de ontwikkeling van angststoornissen in het latere leven zijn meerdere onderzoeksinstrumenten ontwikkeld om het construct betrouwbaar en valide te kunnen meten (Rajan et al., 2019). De manier waarop het construct CPB gemeten kan worden is in de huidige literatuur nog onderbelicht. Voor zover bekend is de Challenging Parental Behaviour Questionnaire (CPBQ) het enige instrument waarmee CPB van ouders met jonge kinderen gemeten kan worden. De CPBQ is opgedeeld in vijf verschillende subschalen: Plagen, Stoeien, Stimuleren tot grenzen verleggen, Sociale

durf en Competitie. Deze subschalen zijn gebaseerd op de verschillende dimensies binnen het begrip CPB. De CPBQ is een Nederlandse vragenlijst en is beschikbaar voor meerdere leeftijdsgroepen, maar de versie voor kinderen van vier tot zes jaar is nog niet psychometrisch gevalideerd (Lazarus et al., 2016). Het is van belang dat er meer aandacht wordt besteed aan het meten van CPB, omdat het een ander aspect van opvoeden betreft dan de huidige theorieën, die vooral gefocust zijn op responsiviteit, controle en warmte in de interactie. Het stimuleren van kinderen om hun grenzen op te zoeken en verleggen geeft daarnaast de mogelijkheid aan ouders om concrete handelingen uit te voeren die de kans op het ontwikkelen van een angststoornis kunnen verlagen.

Het doel van dit onderzoek is om de psychometrische kwaliteiten van de CPBQ(4-6) in kaart te brengen, zodat vastgesteld kan worden of de vragenlijst geschikt is om in de praktijk gebruikt te worden. Het onderzoek tracht dit te doen door de onderzoeksvraag ‘In hoeverre is de huidige opdeling van de CPBQ(4-6) in subschalen psychometrisch te verantwoorden?’ te beantwoorden. Deze vraag zal worden beantwoord door verschillende psychometrische kenmerken van de CPBQ(4-6) te analyseren. Door middel van een mokkenschaaanalyse (MSA) zal worden gekeken naar de sterkte en unidimensionaliteit van de subschalen. Sterkte is de mate waarin items binnen de subschalen passen. Unidimensionaliteit is de mate waarin de subschalen hetzelfde onderliggende construct, in dit geval CPB, meten. Sterkte en unidimensionaliteit zijn belangrijke aspecten om rekening mee te houden omdat ze bijdragen aan de constructvaliditeit van een vragenlijst. Constructvaliditeit verwijst naar de mate waarin een meetinstrument meet wat het beoogt te meten. Betrouwbaarheid is een voorwaarde voor constructvaliditeit. De resultaten van een betrouwbare vragenlijst zijn consistent over herhaalde metingen onder dezelfde omstandigheden. Naast de sterkte en unidimensionaliteit wordt onderzocht of er steun te vinden is voor de huidige opdeling van de CPBQ(4-6) in de vijf verschillende subschalen. Als laatste wordt onderzocht of er een verschil is in psychometrische gegevens tussen vaders en moeders. Dit houdt in dat er gekeken wordt of het instrument CPB voor vaders op eenzelfde manier meet als voor moeders. Als er geen verschil is in de manier waarop CPB gemeten wordt tussen de subgroepen, zouden de conclusies over de psychometrische kwaliteit bij vaders en moeders hetzelfde zijn. De vragenlijst is in dat geval steekproefonafhankelijk. Er is een verschil in de mate waarin en manier waarop vaders en moeders CPB tonen naar hun kinderen (Bögels & Perotti, 2011; Grossmann et al., 2002). Vanwege dit verschil is het aannemelijk dat de CPBQ(4-6) het construct voor vaders en moeders verschillend meet. Wanneer de psychometrische gegevens verschillen tussen de subgroepen, kan kennis van

deze variatie nuttig zijn voor het interpreteren van individuele testresultaten en gemiddelde schaalscores. Hierom is het vanwege praktische implicaties met betrekking tot de bruikbaarheid van de CPBQ(4-6) van belang om te onderzoeken of er een verschil is in psychometrische kwaliteit van de subgroepen. Aangezien de psychometrische kwaliteiten van de versies van de CPBQ voor kinderen jonger dan vier jaar voldoende zijn gebleken (Lazarus et al., 2016; Möller et al., 2015), wordt verwacht dat er ook steun gevonden wordt voor de huidige opdeling van de CPBQ(4-6).

In dit onderzoek wordt getracht de onderstaande bijbehorende deelvragen te beantwoorden:

- ‘In hoeverre wordt er met een exploratieve analyse een samenstelling van subschalen gevonden die overeenkomt met de oorspronkelijke subschalen van de CPBQ(4-6)?’
- ‘In hoeverre blijkt uit de mokkenschaalanalyse dat een schaal met alle items van de CPBQ(4-6) voldoende construct valide is?’
- ‘In hoeverre vormen de oorspronkelijke subschalen van de CPBQ(4-6) een meetinstrument met voldoende constructvaliditeit?’
- ‘In hoeverre zijn de psychometrische kwaliteiten tussen vaders en moeders voor de best mogelijk presterende subschalen verschillend (steekproefonafhankelijk)?’

De laatste vraag zal beantwoord worden voor de beste schalen voortkomend uit de exploratieve analyse.

Methodie

Onderzoeksdesign

In dit onderzoek wordt gebruikt gemaakt van kwantitatieve onderzoeksmethoden, namelijk een survey-onderzoek en een statistische analyse. Er zijn gestandaardiseerde enquêtes gebruikt om de gegevens te verzamelen. De verzamelde data maakt deel uit van een groter onderzoek van het 3HOEK-onderzoeksproject naar de invloed van oudergedrag op de ontwikkeling van het jonge kind. De data voor het huidige onderzoek is verzameld door middel van een eenmalige vragenlijst. Binnen het onderzoek van 3HOEK is CPB ook onderzocht door middel van observationeel onderzoek.

Doelpopulatie, toegankelijke populatie en steekproef

De doelpopulatie in dit onderzoek bestaat uit ouders met een Nederlandse achtergrond en de toegankelijke populatie bestaat uit ouders in Rotterdam. De inclusiecriteria voor het onderzoek waren: het hebben van een driejarig kind, het samen opvoeden van het kind en het beiden beschikken over een Nederlandse achtergrond, wat inhoudt dat de ouders in Nederland geboren zijn. De steekproef bestaat uit 104 moeders en 104 vaders uit de regio Rotterdam. De vaders waren ten tijde van het onderzoek gemiddeld 36.64 jaar oud ($SD = 5.26$) en de moeders 33.28 jaar oud ($SD = 4.66$). In 53 van 104 gezinnen beschikt minstens één van de ouders over een opleidingsniveau van HBO of hoger. In de andere 51 gezinnen beschikt geen van de ouders over een opleidingsniveau van HBO of hoger (Verweij et al., 2021).

Onderzoeksvariabelen en instrumenten

De data zijn verzameld aan de hand een verkorte versie van de CPBQ(4-6). Deze vragenlijst meet de variabele CPB; de mate waarin een ouder een kind stimuleert tot sociaal-emotioneel en fysiek risicovol gedrag of tot gedrag dat buiten de comfort-zone van het kind gaat (Majdandžić et al., 2018). Naast CPB wordt in het onderzoek rekening gehouden met geslachtsvariabele. De vragenlijst waarmee de data zijn verzameld bevat 20 items verdeeld over vijf subschalen: *Plagen*, *Stoeien*, *Stimuleren tot grenzen verleggen*, *Sociale durf en Competitie*. Iedere subschaal bevat vier items. Plagen is bijvoorbeeld door het kind voor de gek te houden of te laten schrikken. Stoeien is het wat ruiger om gaan met het kind, door bijvoorbeeld kussengevechten of in de rondte draaien. Stimuleren tot grenzen verleggen is het kind verbaal uitdagen om spannende dingen te doen, bijvoorbeeld nieuwe dingen proberen of angsten opzij te zetten. Sociale durf is het kind sociale situaties zelf op laten lossen, bijvoorbeeld door vreemden aan te spreken of het kind aan te moedigen om voor zichzelf op

te komen. Competitie is het aanmoedigen van het kind om zijn best te doen in spel en sport (Majdandžić et al., 2015). Ouders worden gevraagd om elk item te beoordelen op een 5-punten-Likert schaal van helemaal niet van toepassing (1) tot helemaal van toepassing (5). Aangezien de data zijn verzameld met een, door de onderzoekers, aangepaste versie van de CPBQ(4-6), waarin er nog 20 items over zijn gebleven van de oorspronkelijke 39, is ervoor gekozen om alle 20 items mee te nemen in de analyse. Meegenomen is dat er geen theoretische overwegingen zijn om een item uit de set te verwijderen. Er zijn geen psychometrische gegevens bekend over de verkorte versie van de CPBQ(4-6).

In Tabel 1 is een overzicht weergegeven van welk item welk gedrag meet.

Tabel 1

Item met label en omschrijving van gedrag

Item	Label	Subschaal	Omschrijving
1	Grapjes	Plagen	Ik haal grapjes uit met mijn kind.
2	Stoeien	Stoeien	Ik stoei met mijn kind.
3	Spannende dingen	Grenzen verleggen	Ik moedig mijn kind aan om spannende dingen te doen, zoals van een hoge glijbaan afgaan.
4	Ruzie	Sociale durf	Als mijn kind naar me toe komt omdat hij/zij een ruzietje heeft, dan laat ik hem/haar dat zelf oplossen.
5	Tikkertje	Competitie	Bij een spelletje tikkertje met mijn kind laat ik me niet zomaar pakken.
6	Eng vinden	Grenzen verleggen	Als mijn kind iets eng vindt, dan stimuleer ik hem/haar om toch door te gaan.
7	Afstappen	Sociale durf	Ik laat mijn kind zelf op andere kinderen afstappen om te vragen om samen te spelen.
8	Aanmoedigen	Competitie	Ik moedig mijn kind aan de beste te zijn.
9	Voor de gek houden	Plagen	Ik houd mijn kind vrijwel nooit voor de gek.
10	Rondzwaaien	Stoeien	Ik doe wel eens een spelletje met mijn kind waarbij ik hem/haar in de rondte zwaai.
11	Nieuwe dingen	Grenzen verleggen	Als ik iets zie dat nieuw en spannend voor mijn kind is ga ik er direct met hem/haar op af.

12	Vragen	Sociale durf	Als mijn kind op de wip of schommel wil maar er zit een ander kind op, dan laat ik hem/haar zelf vragen of hij/zij erop mag.
13	Uitdagen	Competitie	Ik daag mijn kind uit tot een wedstrijdje, bijvoorbeeld wie het hardst kan.
14	Plagen	Plagen	Ik plaag mijn kind regelmatig.
15	Gooien	Stoeien	Ik gooi mijn kind wel eens in de lucht en vang hem/haar weer op.
16	Opkomen	Sociale durf	Ik stimuleer mijn kind om voor zichzelf op te komen.
17	Wedstrijdjes	Competitie	Ik moedig mijn kind aan om wedstrijdjes te doen met andere kinderen.
18	Schrikken	Plagen	Ik laat mijn kind voor de grap schrikken, bijvoorbeeld door onverwachts tevoorschijn te komen.
19	Speeltuin	Grenzen verleggen	In de speeltuin stimuleer ik mijn kind om alle speeltoestellen uit te proberen.
20	Dansen	Stoeien	Ik houd ervan om wild met mijn kind te dansen.

Procedure

De data zijn verzameld door onderzoekers van het 3HOEK-onderzoeksproject, die onderzoek deden naar de rol van vaders en moeders in de ontwikkeling van jonge kinderen. Voor het huidige onderzoek is de verzamelde data verstrekt om te analyseren. De participanten zijn geworven door een team van student-assistenten bij speeltuinen, zwembaden, festivals en bibliotheken. De geworven gezinnen zijn vervolgens in de periode tussen mei 2018 en januari 2020 bezocht door twee getrainde observatoren. Het eerste half uur van het bezoek is besteed aan invullen van vragenlijsten door ouders, daarna vonden er observaties plaats. De ouders hebben de vragenlijst zelfstandig op een tablet ingevuld. Daarnaast hebben alle ouders schriftelijke toestemming verleend om deel te nemen aan het onderzoek.

Data-analyse

In de volgende paragrafen volgt een globale uitleg van de mokschaalanalyse en vervolgens wordt beschreven hoe de analyses in dit onderzoek zijn uitgevoerd.

Mokkenschaalanalyse

Door middel van een MSA (Mokken, 1971) wordt de data geanalyseerd. Een MSA is een vorm van non-parametrische Item Response Theorie (IRT). IRT is een statistische methode voor het beoordelen van de relatie tussen onderliggende kenmerken van personen, ook wel een latente trek genoemd en de manier waarop de latente trek tot uiting komt, bijvoorbeeld in een testscore (Bock & Gibbons, 2021). De items binnen een mokkenschaal zijn steekproefonafhankelijk, wat inhoudt dat de psychometrische gegevens van de items onafhankelijk zijn van de steekproef die gebruikt wordt. Een steekproefonafhankelijk meetinstrument zou bij verschillende steekproeven tot dezelfde conclusies moeten komen over de psychometrische gegevens. Mocht een meetinstrument twee subgroepen op een verschillende manier meten, dan is er mogelijk sprake van differential item functioning (DIF) (Ozberk et al., 2021; Visser et al., 2015).

De MSA kan zowel gebruikt worden om de dimensionaliteit van een vragenlijst te onderzoeken, als de sterkte van de items en schalen te meten (Sijtsma & Van der Ark, 2017). Een mokkenschaal gaat uit van de volgende assumpties: unidimensionaliteit, lokale onafhankelijkheid, monotonie en dubbele monotonie. Unidimensionaliteit houdt in dat de items één onderliggend construct meten. Lokale onafhankelijkheid houdt in dat de score op één item geen informatie biedt over een ander item. Monotonie houdt in dat de scores op de items niet dalend gerelateerd zijn aan het onderliggende construct. Dubbele monotonie houdt in dat de itemresponsecurves elkaar niet kruisen. Dit is van belang omdat het veronderstelt dat de volgorde van itemmoeilijkheid geldt ongeacht de score op de latente trek (Roorda, et al., 2011; Sijtsma & van der Ark., 2017).

Crit-waardes geven informatie over in welke mate een item de assumpties schendt. Een crit-waarde van onder de 40 geeft aan dat er slechts kleine schendingen zijn, die waarschijnlijk te wijten zijn aan steekproeffluctatie. Een crit-waarde tussen de 40 en 80 geeft aan dat er twijfel is over de schending van de assumpties. Een crit-waarde van hoger dan 80 is een sterke aanwijzing dat de assumpties geschonden worden (Molenaar & Sijtsma, 2000a). Rho-waardes kunnen, mits de dubbele monotonie geldt, gezien worden als indicatie voor de betrouwbaarheid van de schaal. Een schaal met een rho-waarde tussen de 60 en 80 wordt gezien als een schaal met een lage betrouwbaarheid, terwijl een schaal met een rho-waarde van hoger dan 80 gezien wordt als een betrouwbare schaal (Zijlstra et al., 2023).

De items binnen de subschalen zijn voldoende sterk als ze met elkaar samenhangen en één construct meten. De sterkte coëfficiënt van een schaal wordt uitgedrukt in de waarde H en de sterkte coëfficiënt van items in H_i . De volgende vuistregel wordt gebruikt: $H: 0.3 \leq H \leq$

0.4 is een zwakke schaal, $0.4 \leq H \leq 0.5$ is een gemiddeld sterke schaal en $H \geq 0.5$ is een sterke schaal. Dezelfde waarden worden aangehouden voor het uitdrukken van de sterkte van H_i . De gemiddelde scores van een persoon op de items geven een schatting van de positie van een persoon op de latente trek. Hoe hoger de score, hoe hoger de schatting van de positie van de persoon op de latente trek. De gemiddelde scores op de items geven een schatting van de positie van een item. Hoe hoger de gemiddelde score op een item, hoe makkelijker het item wordt gevonden. Een moeilijk item houdt in dat er een hoge mate van de latente trek nodig is om hoog op het item te scoren. Voor een makkelijker item is een mindere mate van de latente trek nodig (Molenaar & Sijtsma, 2000a).

Psychometrische analyse

Het softwareprogramma MSPWIN5 (Molenaar & Sijtsma 2000b) is gebruikt om de MSA uit te voeren. Door middel van een exploratieve en confirmatieve MSA zijn de deelvragen beantwoord.

De eerste deelvraag is beantwoord door middel van de ‘search normal’ functie in MSPWIN5 een exploratieve analyse van 20 de items uit te voeren. Deze procedure plaatst de geselecteerde set items door middel van een stapsgewijze bottom-up zoekprocedure in één of meerdere schalen. Op deze manier ontstaan de best mogelijk presterende subschalen en kan gekeken worden of deze subschalen overeen komen met de huidige opdeling van de CPBQ(4-6) in vijf subschalen.

De tweede deelvraag is beantwoord door middel van de ‘test’ functie in MSPWIN5 een confirmatieve analyse van de 20 items uit te voeren. Deze procedure plaatst alle items in één schaal. Op basis van de H-waarden, H_i -waarden, rho-waarden en crit-waarden van de gemaakte schaal zijn er conclusies getrokken over de betrouwbaarheid en validiteit als meetinstrument voor CPB. Een schaal met een H-waarde van 0.3 of hoger wordt als een voldoende sterke schaal beschouwd.

De derde deelvraag is beantwoord door middel van de ‘test’ functie in MSPWIN5 voor elke subschaal een afzonderlijke confirmatieve analyse uit te voeren. Op basis van de H-waarden, H_i -waarden, rho-waarden en crit-waarden zijn er conclusies getrokken over de individuele items en de betrouwbaarheid en validiteit van de afzonderlijke subschalen als meetinstrument voor de onderliggende constructen van CPB. In dit onderzoek worden subschalen waarin alle vier H_i -waarden boven de 0.3 scoren als voldoende sterk beschouwd.

De vierde deelvraag is beantwoord door middel van de ‘test’ functie in MSPWIN5, afzonderlijk voor vaders en moeders een confirmatieve analyse uit te voeren. De analyses zijn

uitgevoerd voor de best mogelijk presterende schalen van de exploratieve analyse van de eerste deelvraag. Er is gekozen voor deze schalen omdat de steekproefonafhankelijkheid zo onderzocht wordt voor de schalen waarin de items het meest met elkaar samenhangen. Ten eerste is er door middel van de H-waarde en de rho-waarde gekeken of de subschalen even goed bij beide subgroepen horen. Ten tweede is er door middel van de H_i -scores gekeken of de itemscores op een vergelijkbare manier zijn verdeeld. Ten derde is er door middel van de gemiddelde scores gekeken of er items zijn die voor de ene subgroep moeilijker zijn dan voor de andere subgroep. De crit-waardes van equal item step ordening (EISO) zijn gebruikt om te controleren of er geen schendingen zijn in de DIF van de items.

Als laatste is er met het softwareprogramma IBM SPSS Statistics (versie 28) een gepaarde t-test uitgevoerd om te kijken of er een significant verschil is tussen de gemiddelde totaalscores van de twee groepen. Er is gekozen voor een gepaarde t-test omdat de scores van de vaders en moeders, doordat ze hetzelfde kind opvoeden, afhankelijk van elkaar zijn.

Resultaten

Uit de exploratieve analyse van de 20 items van de CPBQ(4-6) komen 4 verschillende subschalen voort van 10, 4, 3 en 2 items. Uit de analyse bleek dat item 9 afgewezen werd doordat het een negatieve H_1 -coëfficiënt had met enkele van de andere items. Hierdoor is item 9 in de analyses die hierop volgen niet meegenomen. Met uitzondering van de confirmatieve analyse van de subschaal plagen.

Tabel 2 toont van de eerste best mogelijk presterende subschaal uit de exploratieve analyse de gemiddelde scores, de H_1 -waardes voor elk item en de crit-waardes.

Tabel 2

De eerste best mogelijk presterende subschaal.

Item	Label	Gemiddelde	ItemH	Crit-waarde*
1	Grapjes	4.26	0.37	
3	Spannende dingen	4.16	0.39	17
2	Stoeien	4.06	0.37	16
10	Rondzwaaien	3.71	0.31	11
13	Uitdagen	3.59	0.39	21
15	Gooien	3.00	0.41	16
18	Schrikken	2.94	0.42	5
14	Plagen	2.73	0.34	7
8	Aanmoedigen	2.63	0.32	22
17	Wedstrijdjes	2.58	0.37	19

Noot. H-waarde schaal = 0.37, Rho = 0.83.

*Voor check dubbele monotonie.

De H-waarde van 0.37 geeft aan dat de items met elkaar samenhangen en een zwakke schaal vormen. De crit-waardes zijn allemaal onder de 40, wat aangeeft dat er geen schendingen lijken te zijn in de assumpties van dubbele monotonie. De schaal heeft een rho van 0.83, wat aangeeft dat er sprake is van een redelijk betrouwbare schaal. De schaal is opgebouwd uit drie items van plagen, drie items van stoeien, drie items van competitie en één item uit grenzen verleggen. Inhoudelijk hebben de meeste items te maken met de attitudes van ouders richting winnen en verliezen en sport en spel.

Tabel 3 toont van de tweede best mogelijk presterende subschaal uit de exploratieve analyse de gemiddelde scores, de H_1 -waardes voor elk item en de crit-waardes van de tweede schaal.

Tabel 3*De tweede best mogelijk presterende subschaal.*

Item	Label	Gemiddelde	ItemH	Crit-waarde*
16	Opkomen	4.14	0.32	12
7	Afstappen	4.06	0.46	4
12	Vragen	3.86	0.39	11
4	Ruzie	3.44	0.43	14

Noot. H-waarde schaal = 0.40, Rho = 0.69.

*Voor check dubbele monotonie.

De H-waarde van 0.40 geeft aan dat de items met elkaar samenhangen en een gemiddeld sterke schaal vormen. De crit-waardes zijn allemaal onder de 40, wat aangeeft dat er geen schendingen lijken te zijn van de assumptie van dubbele monotonie. De rho van 0.69 is aan de lage kant, maar dit kan verklaard worden door het geringe aantal items in de schaal. Alle vier de items komen oorspronkelijk uit de ‘sociale durf’ subschaal en lijken deze latente trek dus ook te meten. Tabel 3 toont, aangezien de exploratieve analyse dezelfde vier items in één schaal heeft gestopt, ook de confirmatieve analyse van de schaal sociale durf.

De laatste twee subschalen worden niet meegenomen in de analyse, omdat ze niet groot genoeg geacht worden om betrouwbare conclusies uit te trekken. De creatie van de vier subschalen en het feit dat de oorspronkelijke subschalen voor een groot deel uit elkaar zijn getrokken geven aan dat er geen steun te vinden is voor de huidige opdeling van de CPBQ(4-6) in vijf verschillende subschalen. Enkel de ‘sociale durf’ subschaal blijft intact in de exploratieve analyse. De andere vier subschalen lijken dus niet elk apart een dimensie van het construct CPB te meten.

Tabel 4 toont van de confirmatieve analyse van de subschaal van de negentien overgebleven items van de CPBQ(4-6) de gemiddelden, H_i -waardes en crit-waardes.

Tabel 4*De confirmatieve analyse van de items van de CPBQ (4-6)*

Item	Label	Gemiddelde	ItemH	Crit-waarde*
1	Grapjes	4.26	0.29	22
3	Spannende dingen	4.16	0.36	39
16	Opkomen	4.14	0.27	30
2	Stoeien	4.06	0.28	18
7	Afstappen	4.06	0.26	17

12	Vragen	3.86	0.23	32
19	Speeltuin	3.82	0.27	22
10	Rondzwaaien	3.71	0.29	21
6	Eng vinden	3.59	0.26	25
13	Uitdagen	3.59	0.34	9
4	Ruzie	3.44	0.26	16
20	Dansen	3.29	0.26	18
5	Tikkertje	3.28	0.21	44
11	Nieuwe dingen	3.15	0.27	19
15	Gooien	3.00	0.31	14
18	Schrikken	2.94	0.36	30
14	Plagen	2.73	0.28	22
8	Aanmoedigen	2.63	0.29	17
17	Wedstrijdjes	2.58	0.33	13

Noot. H-waarde schaal = 0.29, Rho = 0.87.

*Voor check dubbele monotonie.

De H-waarde van 0.29 geeft aan dat de items niet sterk genoeg met elkaar samenhangen om een schaal te vormen. Hieruit valt op te maken dat de unidimensionaliteit van de schaal niet voldoende is. Er kan dus niet geconcludeerd worden dat de items slechts één onderliggend construct meten. De crit-waardes zijn op één waarde van 44 na allemaal onder de 40 wat aangeeft dat er geen schendingen lijken te zijn van de assumptie van dubbele monotonie. De rho van de schaal is 0.87, wat erop duidt dat de schaal wel betrouwbaar is.

Tabel 5 toont van de confirmatieve analyse van de subschaal plagen de gemiddelden, H_i -waardes en crit-waardes.

Tabel 5

De confirmatieve analyse van de subschaal plagen

Item	Label	Gemiddelde	ItemH	Crit-waarde*
1	Grapjes	4.26	0.32	
9	Voor de gek houden	3.49	0.23	34
18	Schrikken	2.94	0.34	23
14	Plagen	2.73	0.40	14

Noot. H-waarde schaal = 0.32, Rho = 0.63.

*Voor check dubbele monotonie.

De H-waarde van 0.32 geeft aan dat de items met elkaar samenhangen en een zwakke schaal vormen. Item 9 heeft de laagste H_i -waarde en past daarmee niet goed in de schaal. Inhoudelijk verschilt item 9 niet veel van de andere items. Wel is het mogelijk dat de verwoording ‘vrijwel nooit’ voor verwarring heeft gezorgd, waardoor het item door ouders verschillend is geïnterpreteerd. Zonder item 9 gaat de H-waarde van de schaal van 0.32 naar 0.42. De crit-waardes zijn allemaal onder de 40, wat aangeeft dat er geen schendingen lijken te zijn van de assumptie van dubbele monotonie. De rho van de schaal is met een waarde van 0.63 aan de lage kant, maar deze lage betrouwbaarheid kan verklaard worden door het geringe aantal items van de schaal.

Tabel 6 toont van de confirmatieve analyse van de subschaal stoeien de gemiddelden, H_i -waardes en crit-waardes.

Tabel 6

De confirmatieve analyse van de subschaal stoeien

Item	Label	Gemiddelde	ItemH	Crit-waarde*
2	Stoeien	4.06	0.33	23
10	Rondzwaaien	3.71	0.48	
20	Dansen	3.29	0.26	67
15	Gooien	3.00	0.44	59

Noot. H-waarde schaal = 0.32, Rho = 0.63.

*Voor check dubbele monotonie.

De H-waarde van 0.38 geeft aan dat de items met elkaar samenhangen en een zwakke schaal vormen. Item 20 heeft de laagste H_i -waarde en past daarmee het minst in de schaal. Het is mogelijk dat dit veroorzaakt wordt door de inhoud van het item. Item 2, 10 en 15 betreffen alle drie handelingen waarin ruig omgaan met het kind centraal staat. Het is mogelijk dat in de interpretatie van ouders het ruige aspect bij ‘wild dansen’ ontbreekt, waardoor het eerder een onderliggend construct dat te maken heeft met speelsheid zou meten. Zonder item 20 gaat de H-waarde van de schaal van 0.38 naar 0.50. De crit-waardes van item 15 en item 20 zijn aan de hoge kant, wat aangeeft dat er twijfel is over de schending van de assumpties van dubbele monotonie. Om deze reden is de rho van de schaal mogelijk geen indicatie voor de betrouwbaarheid. De rho is met een waarde van 0.69 aan de lage kant, maar deze lage betrouwbaarheid kan wederom verklaard worden door het geringe aantal items van de schaal.

Tabel 7 toont van de confirmatieve analyse van de subschaal grenzen verleggen de gemiddelden, H_i -waardes en crit-waardes.

Tabel 7

De confirmatieve analyse van de subschaal grenzen verleggen

Item	Label	Gemiddelde	ItemH	Crit-waarde*
3	Spannende dingen	4.16	0.43	24
19	Speeltuinen	3.82	0.38	
6	Eng vinden	3.59	0.34	28
11	Nieuwe dingen	3.15	0.37	

Noot. H-waarde schaal = 0.38, Rho = 0.67.

*Voor check dubbele monotonie.

De H-waarde van 0.38 geeft aan dat de items met elkaar samenhangen en een zwakke schaal vormen. De crit-waardes van de schaal zijn allemaal onder de 40, wat aangeeft dat er geen schendingen lijken te zijn van de assumptie van dubbele monotonie. De rho van de schaal is met 0.67 aan de lage kant, maar deze lage betrouwbaarheid kan wederom verklaard worden door het geringe aantal items van de schaal.

Tabel 8 toont van de confirmatieve analyse van de subschaal competitie de gemiddelden, H_i -waardes en crit-waardes

Tabel 8

De confirmatieve analyse van de subschaal competitie

Item	Label	Gemiddelde	ItemH	Crit-waarde*
13	Uitdagen	3.59	0.40	48
5	Tikkertje	3.28	0.15	95
8	Aanmoedigen	2.63	0.37	28
17	Wedstrijdjes	2.58	0.41	59

Noot. H-waarde schaal = 0.34, Rho = 0.64.

*Voor check dubbele monotonie.

De H-waarde van 0.34 geeft aan dat de items met elkaar samenhangen en een zwakke schaal vormen. Item 5 heeft de laagste H_i -waarde en past daarmee niet goed in de schaal. Inhoudelijk verschilt item 5 ook van de andere drie items, die allen meer nadruk leggen op het 'de beste zijn' en wedstrijdjes doen. Item 5 lijkt inhoudelijk meer te gaan over het doorzettingsvermogen van het kind. De crit-waarde van item 5 is dusdanig hoog, dat er een sterke aanwijzing is dat de dubbele monotonie wordt geschonden. Aangezien er een sterke

aanwijzing is dat de assumptie van dubbele monotonie geschonden wordt is de rho geen indicatie van de betrouwbaarheid van de schaal. Zonder item 5 gaat de H-waarde van de schaal van 0.34 naar 0.50 en lijkt de assumptie van dubbele monotonie ook niet meer geschonden te worden.

Over het algemeen zijn de H-waardes van de vijf subschalen voldoende, maar wel aan de lage kant. De sterkste schaal is de sociale durf schaal met een H-waarde van 0.40. De subschalen plagen, stoeien en competitie hebben allen een item die niet voldoet aan de minimum H_i -waarde van 0.30. Bij de stoeien schaal is er sprake van twijfels van de schending van dubbele monotonie en bij de competitie schaal zijn er aanwijzingen dat de assumptie van dubbele monotonie wordt geschonden.

Tabel 9 toont van de exploratieve analyse van de eerste best mogelijk presterende subschaal voor beide subgroepen de gemiddelde scores, de H_i -waardes voor elk item en de crit-waardes.

Tabel 9

De eerste best mogelijk presterende subschaal voor beide subgroepen.

Item	Label	Moeders		Vaders		Crit-waarde*
		Gem	ItemH	Gem	ItemH	
1	Grapjes	4.53	0.30	3.99	0.33	47
3	Spannende dingen	4.26	0.33	4.07	0.41	50
2	Stoeien	4.47	0.22	3.66	0.37	54
10	Rondzwaaien	3.92	0.25	3.50	0.30	40
13	Uitdagen	3.67	0.37	3.50	0.38	26
15	Gooien	3.54	0.37	2.46	0.35	63
18	Schrikken	3.11	0.40	2.76	0.41	37
14	Plagen	2.92	0.29	2.54	0.33	22
8	Aanmoedigen	2.76	0.25	2.50	0.34	25
17	Wedstrijdjes	2.67	0.31	2.50	0.41	42

Noot. H-waarde schaal moeders = 0.31, Rho = 0.80. H-waarde schaal vaders = 0.36, Rho = 0.83.

*Voor EISO

De H-waardes van 0.31 voor de subgroep moeders en 0.36 voor de subgroep vaders geven beide aan dat de items van beide schalen samenhangen en beiden zwakke schalen vormen. Bij de subschaal voor moeders zijn er enkele items die onder het gestelde minimum scoren, terwijl bij de vaders alle items voldoen aan het minimum. Opvallend is dat item 2, 8

en 10 alle drie een lage H_i -waarde hebben, terwijl de H_i -waardes van vaders wel voldoende zijn. Op basis van de H en H_i waardes lijkt de subschaal dus iets beter te passen bij de vaders dan bij de moeders. Bij vier van de tien items is de crit-waarde hoger dan 40 wat aangeeft dat er twijfel is over de schending van de assumpties van equal item step ordening. Er zijn enkele items die niet de volgorde van moeilijkheid volgen. Bij de moeders is de moeilijkheid van item 2 en item 3 omgedraaid. Bij de vaders is de moeilijkheid van item 1 en item 3 omgedraaid en is item 15 het moeilijkste item. Item 15 is hier het meest opvallend, aangezien een gemiddeld moeilijk item, voor vaders het moeilijkste item is. Inhoudelijk zou niet verwacht worden dat vaders het minst op het item zouden scoren. Met een rho van 0.80 voor moeders en een rho van 0.83 voor vaders is er bij beide schalen sprake van een redelijk betrouwbare schaal.

Tabel 10 toont van de tweede best mogelijk presterende subschaal voor beide subgroepen uit de exploratieve analyse de gemiddelde scores, de H_i -waardes voor elk item en de crit-waardes. De tabel houdt volgorde van de moeilijkheid van de items zonder subgroepen aan.

Tabel 10

De tweede best mogelijk presterende subschaal voor beide subgroepen

Item	Label	Moeders		Vaders		Crit-waarde*
		Gem	ItemH	Gem	ItemH	
16	Opkomen	4.13	0.41	4.15	0.24	12
7	Afstappen	4.15	0.52	3.97	0.41	4
12	Vragen	3.91	0.42	3.81	0.38	
4	Ruzie	3.40	0.47	3.48	0.40	

Noot. H-waarde schaal moeders = 0.45, Rho = 0.74. H-waarde schaal vaders = 0.36, Rho = 0.66.

*Voor EISO

De H-waarde van 0.45 voor de subgroep moeders geeft aan dat de schaal samenhangt en een gemiddeld sterke schaal vormt. De H-waarde van 0.36 voor de subgroep vaders geeft ook aan dat de schaal samenhangt, maar een zwakke schaal vormt. Item 16 bij vaders heeft als enige item een H_i -waarde onder het gestelde minimum. Inhoudelijk verschilt het item niet veel van de andere items. Op basis van de H en H_i -waardes lijkt de subschaal dus iets beter te passen bij de moeders dan bij de vaders. De rho van 0.74 en 0.66 zijn voor beide schalen aan de lage kant, maar dit kan verklaard worden door het geringe aantal items. Er is weinig verschil in de volgorde van moeilijkheid van de items bij beide subgroepen. Alleen bij de

moeders is zijn item 7 en item 16 omgedraaid, dit gaat echter maar om een verschil in score van 0.02.

Wanneer door middel van een gepaarde t-toets naar totaalscores van de ouders op de twee subschalen wordt gekeken is te zien dat ouders gemiddeld 49.17 scoren ($SD = 8.29$; min = 26.00; max = 67.00). Moeders scoren gemiddelde hoger ($M = 51.44$; $SD = 7.63$) dan vaders ($M = 46.90$; $SD = 8.33$). Dit verschil is significant: $t(100) = 4.49$, $p = <.001$.

Conclusie en discussie

Conclusie en discussie

In dit onderzoek is door middel van een mokkenschaalanalyse de volgende onderzoeksvraag onderzocht: ‘In hoeverre is de huidige opdeling van de CPBQ(4-6) in subschalen psychometrisch te verantwoorden?’.

De eerste deelvraag luidde: *‘In hoeverre wordt er met een exploratieve analyse een samenstelling van subschalen gevonden die overeenkomt met de oorspronkelijke subschalen van de CPBQ(4-6)?’*. Op basis van de resultaten kan gesteld worden dat er geen steun wordt gevonden voor de opdeling in de huidige vijf subschalen. De exploratieve analyse laat namelijk een andere opdeling zien. Alleen de Sociale durf schaal ($H = 0.40$) blijft intact en lijkt hetzelfde onderliggende construct te meten. Daarnaast wordt er een schaal van tien items bestaande uit de overige vier subschalen gecreëerd. Deze schaal lijkt CPB te meten op een manier die te maken heeft met attitudes richting winst-verlies en sport en spel. De vier subschalen, die niet intact zijn gebleven, lijken elk dus niet het beoogde construct dat bijdraagt aan de CPB. Deze bevinding gaat tegen de verwachting in omdat er wel steun is gevonden voor de huidige opdeling bij de vragenlijst voor jongere kinderen (Majdandžić et al, 2010; Möller et al., 2015). Een mogelijke verklaring voor de gevonden opdeling is dat de items afzonderlijk voor elke subschaal de bijbehorende dimensies van CPB niet voldoende dekken. Mogelijk heeft de exploratieve analyse door deze beperkte dekking een andere subschaal geïdentificeerd die meer verklaard wordt door een ander onderliggend construct.

De tweede deelvraag luidde: *‘In hoeverre blijkt uit de mokkenschaalanalyse dat een schaal met alle items van de CPBQ(4-6) voldoende construct valide is?’*. Uit de resultaten van de confirmatieve analyse van de negentien items blijkt dat de items van de subschalen samen niet één schaal vormen ($H = 0.29$). Er kan dus niet worden gesteld dat de negentien items één construct, CPB, meten. Dit gaat opnieuw tegen de verwachting in vanwege de voldoende gebleken psychometrische kwaliteiten van de versies voor jongeren kinderen. Een mogelijke verklaring voor dat de gevormde schaal niet één onderliggend construct lijkt te meten is dat enkele items niet effectief het beoogde gedrag van ouders meten. Mogelijk is dit het geval voor items 5 en 12, die de laagste H-waardes hebben. Wanneer de items niet goed zijn afgestemd op het beoogde construct, kan dit leiden tot een vertekend beeld. Wanneer deze items verwijderd worden ontstaat er een zwakke schaal ($H = 0.30$). Een andere verklaring is dat normen of opvattingen over opvoeding in de loop van tijd verschuiven (Forkosh & Erstad, 2018). Aangezien de vragenlijst uit 2010 komt, kan het zijn dat sommige

gedragingen in 2023 bijvoorbeeld als minder uitdagend gezien worden dan andere gedragingen. De variaties in hoe ouders tegen bepaalde gedragingen aankijken kan resulteren in een verminderde samenhang van de vragenlijst.

De derde deelvraag luidde: *'In hoeverre vormen de oorspronkelijke subschalen van de CPBQ(4-6) een meetinstrument met voldoende constructvaliditeit?'*. Uit de resultaten blijkt dat alle subschalen voldoen aan het gestelde minimum van de H-waarde van 0.3. Vier van de vijf subschalen vormen een zwakke schaal en één subschaal vormt een gemiddeld sterke subschaal; Plagen ($H = 0.32$), Stoeien ($H = 0.38$), Stimuleren tot grenzen verleggen ($H = 0.38$), Sociale durf ($H = 0.40$) en Competitie ($H = 0.34$). De subschalen plagen, stoeien en competitie bevatten alle drie een item die niet aan het gestelde minimum voldoet. Wanneer de items onder de H-waarde van 0.3 uit de subschalen verwijderd worden, ontstaan er twee sterke schalen, twee gemiddeld sterke schalen en één zwakke schaal: Plagen ($H = 0.42$), Stoeien ($H = 0.50$), Stimuleren tot grenzen verleggen ($H = 0.38$), Sociale durf ($H = 0.40$) en Competitie ($H = 0.50$). De kritieke waardes voor de subschalen plagen, grenzen verleggen en sociale durf zijn in orde. Voor de subschaal stoeien is er sprake van twijfels over de schending van dubbele monotonie. Voor de subschaal competitie zijn er sterke aanwijzingen dat de dubbele monotonie wordt geschonden. Op basis van deze gegevens kan niet geconcludeerd worden dat de oorspronkelijke subschalen een valide en betrouwbaar instrument vormen. Dit gaat opnieuw tegen de verwachting in vanwege de voldoende gebleken psychometrische kwaliteiten van de versies voor jongeren kinderen. Een verklaring voor dat de schalen niet voldoende constructvalide lijken te zijn, kan wederom op itemniveau gevonden worden. Drie van de vijf subschalen bevatten een item die mogelijk niet goed is afgestemd op het beoogde construct, wat tot een vertekend beeld kan leiden. Daarnaast is het ook mogelijk dat de ontwikkeling in opvatting over opvoeding voor een verminderde samenhang heeft gezorgd.

De vierde deelvraag luidde: *'In hoeverre zijn de psychometrische kwaliteiten tussen vaders en moeders voor de best mogelijk presterende subschalen verschillend (steekproefonafhankelijk)?'*. Op basis van de twee best mogelijk presterende schalen uit de exploratieve analyse is de steekproefonafhankelijkheid onderzocht. Uit de resultaten blijkt dat de H-waardes van de subschalen voor zowel moeders als vaders voldoen aan het gestelde minimum. Gemiddeld genomen zijn de H-waardes van vaders hoger dan de H-waardes van moeders. De eerste subschaal lijkt beter te passen bij vaders ($H = 0.36$, ten opzichte van $H = 0.31$), terwijl de tweede subschaal beter bij moeders past ($H = 0.45$, ten opzichte van $H = 0.36$). Wel zijn er op itemniveau verschillen tussen vaders en moeders en wordt er ook onder het gestelde minimum gescoord. Er zijn kleine verschillen in de volgorde van moeilijkheid

van de items bij de subgroepen. Het grootste verschil is in item 15, waar vaders het laagst op scoren, terwijl dit een gemiddeld moeilijk item zou moeten zijn. Dit is enigszins verontrustend, aangezien de aanwezigheid of afwezigheid van dit item er voor zou kunnen zorgen dat moeders hoger scoren omdat één van de items voor vaders moeilijker is dan voor moeders.

De gepaarde t-test geeft aan dat er een significant verschil is tussen de gemiddelde totaalscore van vaders en moeders. Moeders scoren gemiddeld hoger dan vaders. Verwacht werd dat er een verschil zou zijn tussen de scores van vaders en moeders. Echter werd verwacht dat vaders hoger zouden scoren omdat zij vaker CPB zouden laten zien volgens eerder onderzoek (Grossmann et al., 2002; Majdandžić et al., 2018). Een verklaring voor de hogere gemiddelden van moeders, zou kunnen zijn dat er een ander onderliggend construct wordt gemeten. Deze bevinding wordt ondersteund doordat vaders op de tweede schaal, waar het onderliggende construct duidelijk sociale durf betreft, niet lager scoren dan moeders. Een andere verklaring zou kunnen zijn dat de stellingen in de items niet geschikt zijn om CPB te meten, omdat moeders vaak meer betrokken zijn met de opvoeding van het jonge kind (CBS, 2023). Dit geeft moeders vaker de kans om het gedrag van de stelling te tonen, terwijl vaders het mogelijk in verhouding meer laten zien, wat voor een scheef beeld zou kunnen zorgen.

Bovenstaande resultaten geven meerdere aanwijzingen dat de CPBQ(4-6) mogelijk niet meet wat het zou moeten meten. Ten eerste omdat de items van de CPBQ gezamenlijk niet sterk genoeg samenhangen om een schaal te vormen. Ten tweede vanwege het aantal items dat binnen de subschalen onder het gestelde minimum van de H-waarde scoort. Ten derde omdat moeders significant hogere gemiddelde scores laten zien. Alle resultaten in acht nemend kan gesteld worden dat er geen steun gevonden is om de huidige opdeling van de CPBQ(4-6) psychometrisch te verantwoorden.

Beperkingen van het onderzoek en aanbevelingen voor vervolgonderzoek

Huidig onderzoek kent een aantal beperkingen. Allereerst is de generaliseerbaarheid van het onderzoek in twijfel te trekken. De steekproef bestaat uit ouders die allen wonen in de regio Rotterdam. Dit is van invloed op de externe validiteit van het onderzoek. Het is onduidelijk of de resultaten gegeneraliseerd zouden kunnen worden naar de gehele populatie ouders in Nederland. Hoewel alle participanten van het onderzoek in Nederland wonen, is het mogelijk dat er verschillen zijn in culturele normen in de steekproef, door bijvoorbeeld verschil in religie. Voor ouders van een cultuur waar presteren belangrijker is, is een item als ‘ik moedig mijn kind aan om de beste te zijn bijvoorbeeld makkelijker dan voor ouders waar de cultuur

minder prestatiegericht is. Er is in dit onderzoek geen rekening mee gehouden door bijvoorbeeld de formulering van items aan te passen om ze cultureel relevanter te maken. Voor vervolgonderzoek wordt aanbevolen om het onderzoek te repliceren in andere regio's van Nederland. Wanneer onderzoeken in andere regio's tot dezelfde resultaten leiden, verkleint dat de twijfels over de invloed van cultuur op de externe validiteit. Een andere mogelijkheid om de generaliseerbaarheid van de resultaten te vergroten is door in vervolgonderzoek ook andere analysetechnieken te gebruiken om de resultaten te verifiëren.

Een andere beperking is de interpretatie van de crit-waardes. Recent onderzoek (Crişan et al., 2021) zet twijfels bij de bruikbaarheid van de crit-waardes bij de interpretatie van monotonie en invariant item ordering. Met name in kleinere steekproefgroottes lijkt de power van de crit-waardes voor het detecteren van schendingen van monotonie en het kruisen van item responsfuncties niet voldoende. In grotere steekproefgroottes, zoals 500 of 1000, kwam dit probleem minder naar voren (Crişan et al., 2019). Voor vervolgonderzoek in replicatieonderzoek wordt aanbevolen om een steekproefgrootte van minimaal 500 ouders te gebruiken, zodat de interpretatie van de crit-waardes betrouwbaarder is.

Implicaties voor de praktijk

Er zijn twijfels of de CBPQ(4-6) in de huidige vorm geschikt is om als vragenlijst in de praktijk gebruikt te worden als instrument om de CPB van ouders bij kinderen van 4 tot 6 jaar te meten. Nader onderzoek naar de slechter scorende items is nodig en sommige items zullen mogelijk moeten worden heroverwogen of verfijnd om de samenhang van de vragenlijst te verbeteren.

Literatuur

- American Psychiatric Association. (2013) *Diagnostic and statistical manual of mental disorders*. (5th ed.). <https://doi.org/10.1176/appi.books.9780890425596>
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W H Freeman/Times Books/ Henry Holt & Co.
- Baumrind, D. (1966). Effects of authoritative parental control on child behavior. *Child Development*, 37(4), 887–907. <https://doi.org/10.2307/1126611>
- Bodden, D. H. M., Bogels, S. M., Nauta, M. H., De Hann, E., Ringrose, J., Appelboom, C., Brinkman, A. G., & Appelboom-Geerts, K. C. M. M. J. (2008). Child versus Family Cognitive-Behavioral Therapy in Clinically Anxious Youth: An Efficacy and Partial Effectiveness Study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(12), 1384–D. <https://doi.org/10.1097/CHI.0b013e318189148e>
- Böck, R., & Gibbons, R. D. (2021). *Item Response Theory*. John Wiley & Sons Inc. <https://doi.org/10.1002/9781119716723>
- Bögels, S. M., & Perotti, E. C. (2011). Does father know best? A formal model of the paternal influence on childhood social anxiety. *Journal of Child and Family Studies*, 20(2), 171-181. <https://doi.org/10.1007/s10826-010-9441-0>
- CBS. (2023, 8 maart). *Verdeling werk en zorg tussen vaders en moeders vaak anders dan gewenst*. CBS.nl. Geraadpleegd op 21 november 2023, van <https://www.cbs.nl/nl-nl/nieuws/2023/10/verdeling-werk-en-zorg-tussen-vaders-en-moeders-vaak-anders-dan-gewenst>
- Crişan, Daniela & Tendeiro, Jorge & Meijer, Rob. (2019). The Crit Value as an Effect Size Measure for Violations of Model Assumptions in Mokken Scale Analysis for Binary Data. <https://10.31234/osf.io/8ydmr>.

- Crişan, D.R., Tendeiro, J.N. & Meijer, R.R. (2022). The *Crit* coefficient in Mokken scale analysis: a simulation study and an application in quality-of-life research. *Qual Life Res* 31, 49–59. <https://doi.org/10.1007/s11136-021-02924->
- Erozkan, A. (2012). Examination of relationship between anxiety sensitivity and parenting styles in adolescents. *Kuram ve Uygulamada Eğitim Bilimleri*, 12(1), 52–57.
- Forkosh Baruch A, Erstad O. (2018). Upbringing in a Digital World: Opportunities and Possibilities. *Technology, Knowledge and Learning*, 23(3), 377-390. <https://doi.org/10.1007/s10758-018-9386-8>
- Grossmann, K., Grossmann, K.E., Fremmer-Bombik, E., Kindler, H., Scheuerer-Englisch, H. and Zimmermann, A.P. (2002). The Uniqueness of the Child–Father Attachment Relationship: Fathers’ Sensitive and Challenging Play as a Pivotal Variable in a 16-year Longitudinal Study. *Social Development*, 11, 301-337. <https://doi.org/10.1111/1467-9507.00202>
- Gulley, L. D., Oppenheimer, C. W., & Hankin, B. L. (2014). Associations among Negative Parenting, Attention Bias to Anger, and Social Anxiety among Youth. *Developmental Psychology*, 50(2), 577–585. <https://doi.org/10.1037/a0033624>
- Houtepen, J. A. B. M., Sijtsema, J. J., Klimstra, T. A., Van der Lem, R., & Bogaerts, S. (2019). Loosening the Reins or Tightening Them? Complex Relationships between Parenting, Effortful Control, and Adolescent Psychopathology. *Child & Youth Care Forum*, 48(1), 127–145. <https://doi.org/10.1007/s10566-018-9477-7>
- IBM Corp. (2020). IBM SPSS Statistics for Windows (Version 27.0) [Computer software]. IBM Corp.
- Kiel, E. J., & Buss, K. A. (2011). Prospective Relations among Fearful Temperament, Protective Parenting, and Social Withdrawal: The Role of Maternal Accuracy in a Moderated Mediation Framework. *Journal of Abnormal Child Psychology*, 39(7), 953-966. <https://doi.org/10.1007/s10802-011-9516-4>

- Knappe, S., Beesdo, K., Fehm, L., Lieb, R., & Wittchen, H. (2009). Associations of familial risk factors with social fears and social phobia: Evidence for the continuum hypothesis in social anxiety disorder? *Journal of Neural Transmission*, *116*, 639 – 648.
<https://doi.org/10.1007/s00702-008-0118-4>
- Lazarus, R. S., Dodd, H. F., Majdandžić, M., de Vente, W., Morris, T., Byrow, Y., Bögels, S.M., & Hudson, J. L. (2016). The relationship between challenging parenting behaviour and childhood anxiety disorders. *Journal of Affective Disorders*, *190*, 784-791.
<https://doi.org/10.1016/j.jad.2015.11.032>
- Larzelere, R. E., Morris, A. S., & Harrist, A. W. (Eds.) (2013). *Authoritative parenting: Synthesizing nurturance and discipline for optimal child development*. American Psychological Association. <https://doi.org/10.1037/13948-000>
- Majdandžić, M., Lazarus, R. S., Oort, F. J., van der Sluis, C., Dodd, H. F., Morris, T. M., de Vente, W., Byrow, Y., Hudson, J. L., & Bögels, S. M. (2018). The structure of challenging parenting behavior and associations with anxiety in Dutch and Australian children. *Journal of Clinical Child and Adolescent Psychology*, *47*(2), 282-295.
<https://doi.org/10.1080/15374416.2017.1381915>
- Majdandžić, M., De Vente, W., & Bögels, S. M. (2010). Challenging parenting behavior questionnaire: 4-6 Year Version (CPBQ4-6). *Research institute of child development and education*. Amsterdam, the Netherlands: University of Amsterdam.
- Majdandžić, Mirjana & Vente, Wieke & Bögels, Susan. (2015). Challenging Parenting Behavior from Infancy to Toddlerhood: Etiology, Measurement, and Differences between Fathers and Mothers. *Infancy*, *21*(4), 423-452.
<https://doi.org/10.1111/infa.12125>
- Molenaar, I.W., Sijtsma, K. (2000). *MSPWIN 5.0 A Program for Mokken Scale Analysis for Polytomous Items*. Iec ProGamma, Groningen.

- Molenaar, I.W., & Sijtsma, K. (2000). *MSP5 for Windows User's Manual*. Groningen, Nederland: Iec ProGAMMA.
- Möller, E. L., Majdandžić, M., & Bögels, S. M. (2015). Parental anxiety, parenting behavior, and infant anxiety: Differential associations for fathers and mothers. *Journal of Child and Family Studies*, 24(9), 2626–2637. <https://doi-org.proxy-ub.rug.nl/10.1007/s10826-014-0065-7>

- Morrongiello, B. A., & Dawber, T. (2000). Mothers' responses to sons and daughters engaging in injury-risk behaviors on a playground: implications for sex differences in injury rates. *Journal of Experimental Child Psychology*, 76(2), 89–103.
<https://doi.org/10.1006/jecp.2000.2572>
- Niditch, L. A., & Varela, R. E. (2012). Perceptions of Parenting, Emotional Self-Efficacy, and Anxiety in Youth: Test of a Mediational Model. *Child & Youth Care Forum*, 41(1), 21-35. <https://doi.org/10.1007/s10566-011-9150-x>
- Özberk, E. H., Özberk, E. B. Ü., Uluç, S., & Öktem, F. (2000). Investigating Invariant Item Ordering in Intelligence Tests: Mokken Scale Analysis of KBIT-2. *International Journal of Assessment Tools in Education*, 8(3), 714–728.
<https://doi.org/10.21449/ijate.858183>
- Rapee, R. M., Schniering, C. A., & Hudson, J. L. (2009). Anxiety disorders during childhood and adolescence: Origins and treatment. *Annual Review of Clinical Psychology*, 5, 311–341. <https://doi-org.proxy-ub.rug.nl/10.1146/annurev.clinpsy.032408.153628>
- Pinquart, M. (2017). Associations of parenting dimensions and styles with internalizing symptoms in children and adolescents: A meta-analysis. *Marriage & Family Review*, 53(7), 613–640. <https://doi.org/10.1080/01494929.2016.1247761>
- Sijtsma, K., & van der Ark, L. A. (2017). A tutorial on how to do a Mokken scale analysis on your test and questionnaire data. *British Journal of Mathematical & Statistical Psychology*, 70(1), 137-158. <https://doi.org/10.1111/bmsp.12078>
- RIVM. (2021, 29 juni). *Risico op angststoornis of depressie onder jongvolwassenen toegenomen*. Geraadpleegd op 22 oktober 2022, van <https://www.rivm.nl/nieuws/hoog-risico-angst-depressie-jongeren>
- Roorda, L. D., Houwink, A., Smits, W., Molenaar, I. W., & Geurts, A. C. (2011). Measuring upper limb capacity in poststroke patients: Development, fit of the monotone homogeneity model, unidimensionality, fit of the double monotonicity model, differential item functioning, internal consistency, and feasibility of the Stroke Upper

Limb Capacity Scale, SULCS. *Archives of Physical Medicine and Rehabilitation*, 92, 214–227.

Verweij, R., Helmerhorst, K., & Keizer, R. (2021). Work-to-family conflict, family-to-work conflict and their relation to perceived parenting and the parent-child relationship before and during the first Covid-19 lockdown. *Journal of Family Research*, 33(3), 734-771. <https://doi.org/10.20377/jfr-636>

Visser, L., Ruiter, S., van der Meulen, B., Ruijsenaars, W., & Timmerman, M. (2015). Low verbal assessment with the Bayley-III. *Research in Developmental Disabilities*, 36, 230-243. <https://doi.org/10.1016/j.ridd.2014.09.014>

Zijlstra, A.E., Post, W.J., Hopman, M. *et al.* (2023). The Best Interests of The Child Self-Report (BIC-S): Psychometric Properties of the Adapted Version of the BIC-S used as a Monitoring Instrument to Measure the Quality of The Children's Rearing Environment From a Children's Rights Perspective. *Child Ind Res* 16, 1341–1361. <https://doi.org/10.1007/s12187-023-10010-8>