



UvA-DARE (Digital Academic Repository)

Ontwikkeling meetinstrument pedagogische kwaliteit gastouderopvang

Gevers Deynoot-Schaub, M.J.J.M.; Helmerhorst, K.O.W.; Fukkink, R.G.; Bollen, I.

Publication date

2014

Document Version

Final published version

[Link to publication](#)

Citation for published version (APA):

Gevers Deynoot-Schaub, M. J. J. M., Helmerhorst, K. O. W., Fukkink, R. G., & Bollen, I. (2014). *Ontwikkeling meetinstrument pedagogische kwaliteit gastouderopvang*. Nederlands Consortium Kinderopvang Onderzoek / Kohnstamm Instituut. <http://www.rijksoverheid.nl/bestanden/documenten-en-publicaties/rapporten/2014/12/08/ontwikkeling-meetinstrument-pedagogische-kwaliteit-gastouderopvang/ncko-rapport-szw-ontwikkeling-go-instrument-definitief.pdf>

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Ontwikkeling Meetinstrument

Pedagogische Kwaliteit Gastouderopvang

M.J.J.M. Gevers Deynoot-Schaub, K.O.W. Helmerhorst, R.G. Fukkink, & I. Bollen

Subsidiegever:

Ministerie van Sociale Zaken en Werkgelegenheid

Uitvoerder:

**Nederlands Consortium Kinderopvang Onderzoek
Kohnstamm Instituut**



KOHNSTAMM
INSTITUUT
Kohnstamm

2014

INHOUDSOPGAVE

	Voorwoord	3
1	Inleiding	4
2	Achtergrond	6
	2.1 Het NCKO-kwaliteitsmodel	6
	2.2 Het NCKO-instrumentarium voor de kinderopvang	8
	2.3 Het BSO-instrumentarium voor de buitenschoolse opvang	9
	2.4 Uitgangspunten van het Gastouder-instrument	10
3	Doel van de validatiestudie	19
4	Methode	21
	4.1 Steekproef	21
	4.2 Procedure	22
	4.3 Te valideren instrumenten: Gastouder-instrument	23
	4.4 Instrumenten voor validering	24
	4.5 Concrete verwachtingen	26
5	Resultaten	28
	5.1 GO-interactieschalen	28
	5.2 Algemene proceskwaliteit: FCCERS-R	32
	5.3 Ontwikkelingsgericht werken voor 2-4 jarigen	34
	5.4 Samenvatting van de resultaten	36
6	Conclusie en discussie	37
	Referenties	41

VOORWOORD

Het in dit rapport beschreven onderzoek: ontwikkeling Meetinstrument Pedagogische Kwaliteit Gastouderopvang was onderdeel van een groter project waarin ook een inventarisatiestudie voor de gastouderopvang is uitgevoerd. Deze studie is in verschillende fasen en met verschillende groepen respondenten uitgevoerd die wij op deze plaats graag willen bedanken voor hun deelname. In de eerste plaats gaat onze dank uit naar de gastouders. De validering van onderdelen van het Gastouderinstrument zou onmogelijk geweest zijn zonder deelname van de gastouders, ouders en gastouderbureau's. De gastouders danken we speciaal voor de hulp bij de organisatie rondom de bezoeken van de observatoren en voor het zeer bereidwillig toelaten van de verschillende observatoren tijdens de bezoeken. Daarnaast maakten zij tijd vrij voor een vraaggesprek aan het einde van de observatie en voor het invullen van een uitgebreide vragenlijst. De ouders van de kinderen die bij de gastouders worden opgevangen, willen we bedanken voor de toestemming die zij verleenden voor het filmen van hun kind(eren) tijdens ons bezoek.

We willen graag de volgende medewerkers bedanken voor hun bijdrage aan de organisatie en dataverzameling van dit onderzoek: Lisette Vermeulen, Bennie Mooren, Jessica Muda, Anne Hendriks, Belinda Ommering, Roy Gerst, Kristin van Bekkum, Bo van der Poel en Jessica Fokkinga. Zij gingen in alle vroegte op pad om te observeren of te filmen bij de gastouders en/of hebben vele video-opnames van gastouders en kinderen te beoordeeld. Ook Marianne Boogaard willen wij bedanken voor haar deskundige hulp bij de vertaling van het instrument.

Tot slot danken wij de leden van de klankbordgroep voor hun bijdrage: Helene Arons (Brancheorganisatie), Suzanne Plaisier (Boink), Carlien Langelaan (VGOB), Sebastiaan Dekkers (4Kids) Josette Hoex (NJI), Femke Kolsteren en Yvette Derks (PGV Nederland) en Belkiz Tukas, Alice van Gent en Willeke van der Werf (directie Kinderopvang SWZ).

Dit project werd uitgevoerd door het Nederlands Consortium Kinderopvang Onderzoek (NCKO) en het Kohnstamm Instituut. Deze studie is financieel mogelijk gemaakt door het Ministerie van Sociale Zaken en Werkgelegenheid.

I. INLEIDING

In 2012 maakte in Nederland 15,9% van de kinderen tussen 0 en 12 jaar gebruik van gastouderopvang (Brancheorganisatie Kinderopvang, 2013).

Een landelijk beeld van de pedagogische kwaliteit van de gastouderopvang ontbreekt. Dergelijke peilingen met gevalideerde meetinstrumenten zijn inmiddels wel uitgevoerd voor de kwaliteit van de kinderdagopvang voor 0- tot 4 jarigen. In 2005 werd een eerste landelijke kwaliteitsmeting uitgevoerd (Vermeer et al., 2005), in 2008 werd de tweede peiling uitgevoerd (De Kruif, Riksen-Walraven, Gevers Deynoot-Schaub, Helmerhorst, Tavecchio, & Fukkink, 2009; Helmerhorst, Riksen-Walraven, Gevers Deynoot-Schaub, Tavecchio, & Fukkink, 2014) en in 2012 werd de pedagogische kwaliteit voor de derde keer gemeten (Fukkink, Gevers Deynoot-Schaub, Helmerhorst, Bollen, & Riksen-Walraven, 2013). In 2011 is een eerste peiling van de pedagogische kwaliteit van de buitenschoolse opvang (4-12 jaar) uitgevoerd (Boogaard & Van Daalen-Kapteijns, 2012). Aan de kwaliteitspeilingen in zowel de dagopvang als de buitenschoolse opvang is een grondige fase van instrumentontwikkeling en –validatie voorafgegaan. Het gebruikte instrumentarium voor beide typen peilingen sluit goed op elkaar aan omdat er bij beide gebruik is gemaakt van hetzelfde wetenschappelijke model.

Om de pedagogische kwaliteit van de gastouderopvang (0-12 jaar) te bevorderen en/of op een hoog peil te houden is het allereerst van belang om die kwaliteit te kunnen *meten*. Een goed meetinstrument maakt het mogelijk om ook de kwaliteit van dit type opvang te ‘monitoren’ en waar nodig te verbeteren. Als de directie Kinderopvang van het ministerie van SZW in de toekomst een landelijk kwaliteitsmeting wil laten uitvoeren in de gastouderopvang is er een valide en betrouwbaar meetinstrument nodig. De directie Kinderopvang heeft om die reden opdracht gegeven tot het ontwikkelen van een dergelijk instrument. Daarbij is uitdrukkelijk de wens geformuleerd dat het meetinstrument voor de gastouderopvang aansluit bij het al beschikbare instrumentarium voor andere vormen van de kinderopvang, zoals dat is ontwikkeld door het NCKO (KDV: 0-4 jaar) en het Kohnstamm Instituut (BSO: 4-12 jaar). Daarnaast moest het mogelijk zijn om met het instrument voor de gastouderopvang na te gaan of de kwaliteit van de gastouderopvang voldoet voor de leeftijdsgroepen van 2-4 jaar, gelet op ontwikkelingsgericht werken.

Uiteindelijk doel is om de pedagogische kwaliteit van het gehele kinderopvangveld op een vergelijkbare manier te kunnen meten. Met de ontwikkeling van een meetinstrument voor de gastouderopvang is het instrumentarium voor kwaliteitspeilingen voor de gehele kinderopvangsector compleet.

In dit rapport wordt verslag gedaan van de ontwikkeling en validering van een instrument om de pedagogische kwaliteit van de gastouderopvang op een verantwoorde wijze te meten. In het tweede hoofdstuk wordt de achtergrond van het NCKO-kwaliteitsmodel, het NCKO- en BSO-instrumentarium en de uitgangspunten van het Gastouder-instrument besproken. Hoofdstuk 3 beschrijft het doel van de validatiestudie. Achtereenvolgens komen in hoofdstuk 4 en 5 de Methode en de Resultaten aan de orde en hoofdstuk 6 besluit met een conclusie en discussie.

2. ACHTERGROND

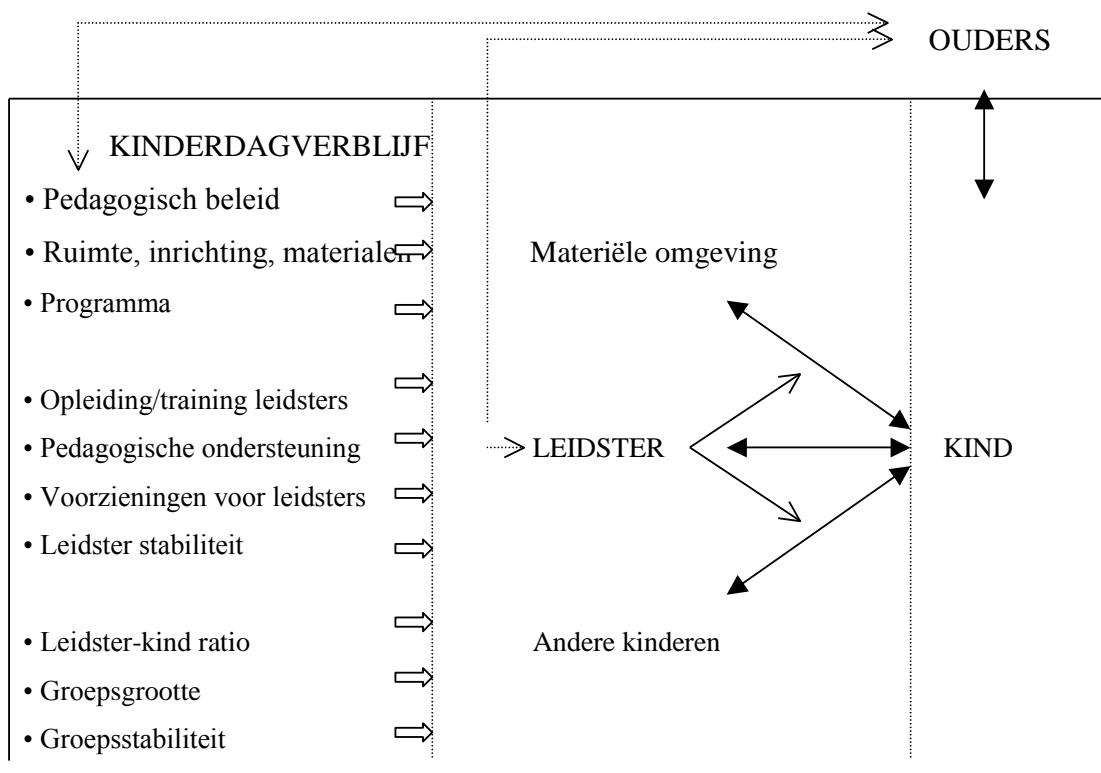
2.1. Het NCKO-kwaliteitsmodel

Het ontwikkelen van een betrouwbaar en valide meetinstrument vereist allereerst duidelijkheid met betrekking tot het te meten *begrip*: wat verstaan we onder ‘pedagogische kwaliteit’? Zijn er meerdere kwaliteitsaspecten te onderscheiden, hoe verhouden die zich tot elkaar en verdienen sommige aspecten bij de meting meer aandacht dan andere? Om deze vragen te beantwoorden heeft het NCKO een *kwaliteitsmodel* voor de kinderopvang ontwikkeld dat uitgebreid is beschreven in hoofdstuk 5 van het boek *De kwaliteit van de Nederlandse kinderopvang* (Riksen-Walraven, 2004). Het NCKO-kwaliteitsmodel is, net zoals bij de ontwikkeling van het meetinstrument voor de kinderdagopvang (De Kruif, Vermeer, Fukkink, Riksen-Walraven, Tavecchio, van IJzendoorn & van Zeijl, 2007) en het meetinstrument voor de buitenschoolse opvang (Boogaard & Fukkink, 2009), in deze studie als uitgangspunt gehanteerd bij de ontwikkeling van het meetinstrument voor de gastouderopvang. In het NCKO-kwaliteitsmodel voor de kinderdagopvang zijn verschillende factoren onderscheiden die direct of indirect van invloed zijn op de ontwikkeling van het kind in een kinderdagverblijf (zie Figuur 2.1). In dit model worden twee typen kwaliteitsindicatoren onderscheiden. Op de eerste plaats gaat het om indicatoren van de *proceskwaliteit* van de opvang, ofwel de kwaliteit van het feitelijke zorg- en opvoedingsproces en van de directe ervaringen van de kinderen. Daarnaast worden er *structurele* kwaliteitsindicatoren onderscheiden, dat wil zeggen reguleerbare kenmerken van de kinderopvang, zoals groeps grootte, leidster-kind ratio en opleiding/training van de pedagogisch medewerkers, die van invloed zijn op de proceskwaliteit en beschouwd kunnen worden als voorwaarden voor het realiseren van een hoge proceskwaliteit. In een eerdere studie van het NCKO werd aangetoond dat de in het model opgenomen proces- en structurele kwaliteitsindicatoren tezamen volgens betrokken partijen (ouders, leidinggevenden, pedagogisch medewerkers en externe deskundigen) een genuanceerd en compleet beeld geven van de pedagogische kwaliteit van de opvang in een kinderdagverblijf en dat beide groepen kwaliteitsindicatoren daarom opgenomen moeten worden in een instrument om de pedagogische kwaliteit van de opvang te meten (Vermeer et al., 2007).

Het perspectief op pedagogische kwaliteit met aandacht voor proceskwaliteit en structurele kwaliteit is door Riksen-Walraven (2004) aangeduid als een *dubbelfocus-benadering*. Zowel de belangrijkste aspecten van de proceskwaliteit als de structurele factoren, die deze proceskwaliteit

distaal - ‘op afstand’ - beïnvloeden, worden immers in kaart gebracht. Samen geven zij een goed beeld van de pedagogische kwaliteit van kinderdagverblijven. Een goede opleiding van de pedagogisch medewerker, bijvoorbeeld, draagt bij aan een juiste omgang met de kinderen, maar wanneer de pedagogische medewerker op een te grote groep werkt, is het voor de medewerker moeilijker om die juiste omgang te realiseren dan op een kleine groep, hoe zeer bepaalde basisvaardigheden ook worden beheerst. Kortom, aandacht voor de proceskwaliteit op de groep enerzijds en de randvoorwaarden daarvoor anderzijds vullen elkaar aan en dragen elk bij aan een goede analyse van de pedagogische kwaliteit.

Figuur 2.1 Het NCKO kwaliteitsmodel



Noot: \longleftrightarrow Proximale processen: ontwikkeling kind in interactie met de directe omgeving
 \Rightarrow Invloed van structurele opvangkenmerken op proximale processen

(Bron: Riksen-Walraven, 2004)

2.2 Het NCKO-instrumentarium voor de kinderdagopvang

Het NCKO heeft op basis van het NCKO-kwaliteitsmodel een instrument ontwikkeld voor de kinderdagopvang (0-4 –jaar) om de pedagogische kwaliteit van de Nederlandse kinderdagverblijven te meten. Dit instrument bestaat uit drie onderdelen.

(a) Algemene proceskwaliteit

Allereerst wordt de zogenoemde **algemene proceskwaliteit** gemeten met behulp van vijf schalen die afkomstig zijn uit de ITERS-R (Infant/Toddler Environment Rating Scale Revised; Harms, Cryer, & Clifford, 2003) en de ECERS-R (Early Childhood Environment Rating Scale Revised; Harms, Clifford, & Cryer, 1998), namelijk (1) Ruimte en meubilering, (2) Taal, (3) Activiteiten, (4) Interacties, en (5) Programma. Tevens wordt een totaalscore voor kwaliteit berekend. Omdat de vijf schalen en de totaalscore ook in de eerdere landelijke metingen zijn gebruikt, kunnen resultaten van achtereenvolgende peilingen met elkaar vergeleken worden.

(b) Interactievaardigheden

Het tweede onderdeel van het NCKO-instrument wordt gevormd door een zestal observatieschalen voor de interactievaardigheden van pedagogisch medewerkers. Deze schalen vormen een belangrijke aanvulling op de ITERS-R/ECERS-R omdat zij een meer gedetailleerd beeld van de **kern van de proceskwaliteit** geven. Zoals beschreven in het NCKO-kwaliteitsmodel wordt deze kern van de proceskwaliteit gevormd door de vaardigheden van pedagogisch medewerkers in hun omgang met de kinderen. De pedagogisch medewerker speelt een sleutelrol in het welbevinden en de ontwikkeling van de kinderen, niet alleen door haar rechtstreekse interacties met de kinderen, maar ook door haar invloed op de interacties tussen de kinderen onderling en hun omgang met het aanwezige spel materiaal. Daardoor draagt zij in hoge mate bij aan het realiseren van de in de Wet kinderopvang genoemde vier pedagogische doelen: het bieden van veiligheid en het bevorderen van de persoonlijke, de sociale en de morele competentie (het zich eigen maken van normen en waarden) van de kinderen. Het NCKO-kwaliteitsmodel – dat wordt onderschreven door ouders, leidinggevendenden, pedagogisch medewerkers en pedagogische experts (De Kruif et al., 2007; Fukkink, Tavecchio, De Kruif, Vermeer & Van Zeijl, 2005) – onderscheidt zes interactievaardigheden waarover pedagogisch medewerkers zouden moeten beschikken:

1. Sensitieve responsiviteit
2. Respect voor de autonomie van het kind
3. Structureren en grenzen stellen

4. Praten en uitleggen
5. Ontwikkelingsstimulering
6. Begeleiden van interacties tussen kinderen

Het NCKO heeft observatieschalen ontwikkeld en gevalideerd waarmee, op basis van videomateriaal, de zes interactievaardigheden van pedagogisch medewerkers in kinderdagverblijven op wetenschappelijk verantwoorde wijze kunnen worden beoordeeld (De Kruif et al., 2007; Helmerhorst, Riksen-Walraven, Vermeer, Fukkink & Tavecchio, 2014).

(c) Structurele kwaliteit

Het derde onderdeel van het NCKO-instrument betreft **structurele kwaliteitskenmerken**, oftewel de kenmerken van de opvang en van de pedagogisch medewerkers die van invloed kunnen zijn op de proceskwaliteit, zoals de groepsgrootte, de staf-kind ratio en de opleiding en werkervaring van de pedagogisch medewerkers. De structurele kwaliteitskenmerken zijn de randvoorwaarden voor proceskwaliteit.

2.3 Het BSO-instrumentarium voor de buitenschoolse opvang

Het meetinstrument pedagogische kwaliteit van buitenschoolse opvang (bso 4 t/m 12 jaar) (Boogaard & Fukkink, 2009), dat gebruikt is voor de eerste landelijke peiling in 2011 (Boogaard & Van Daalen-Kapteijns, 2012), is eveneens gebaseerd op het NCKO-kwaliteitsmodel en kent eenzelfde opbouw als het NCKO-instrument in drie deelinstrumenten:

(a) Algemene proceskwaliteit

De Observatielijst pedagogische kwaliteit brengt de **algemene proceskwaliteit** van de buitenschoolse opvang in beeld. Het instrument bevat in totaal 33 items, verdeeld over acht schalen: (1) Groepen, (2) Ruimte en inrichting, (3) Dagindeling, spelmaterialen en activiteiten, (4) Interacties, (5) Pedagogisch beleid en organisatie, (6) Samenwerking (met ouders, basisscholen en andere instellingen), (7) Kinderen met speciale behoeften en (8) Groepsfunctioneren. De schaal (3) 'Dagindeling, spelmaterialen en activiteiten' sluit goed aan bij de schalen Activiteiten, Taal en Programma uit het NCKO-instrumentarium. Naast het aspect 'evenwichtige dagindeling' zijn in deze schaal de activiteiten op verschillende ontwikkelingsgebieden opgenomen (naast 'beweging en motoriek', 'creatieve ontwikkeling' en 'ruimtelijk inzicht en technische ontwikkeling', gaat het daarbij ook om activiteiten op het gebied van 'taal en denken').

(b) Interactievaardigheden

Met het tweede deelinstrument wordt ingezoomd op de interactievaardigheden van de pedagogisch medewerkers in de buitenschoolse opvang als **kern van de proceskwaliteit**. Het gaat daarbij om zes vaardigheden die gebaseerd zijn op de NCKO-observatieschalen voor kwaliteit van de interactievaardigheden. De schalen zijn aangepast aan de behoeften voor de oudere leeftijdsgroep in de buitenschoolse opvang.

1. Emotionele ondersteuning (sensitiviteit en responsiviteit)
2. Respect voor autonomie
3. Regie en leiding
4. Praten en uitleggen
5. Ontwikkelingsstimulering
6. Begeleiden van sociale interacties tussen kinderen

Op basis van videomateriaal, kunnen de zes interactievaardigheden van pedagogisch medewerkers in de BSO kunnen worden beoordeeld.

N.B. Twee van de hierboven genoemde vaardigheden (1. Emotionele ondersteuning en 3. Regie en leiding) wijken in naam af van nummer 1 en 3 van de interactievaardigheden uit het NCKO-instrument (zie pagina 8) en het Gastouderinstrument (zie pagina 13), maar niet qua concept.

(a) Structurele kwaliteit: structurele, organisatorische en achtergrondkenmerken

Met behulp van een vragenlijst en een deel van de Observatielijst pedagogische kwaliteit zijn de **structurele, organisatorische en achtergrondkenmerken** van de buitenschoolse opvang worden geïnventariseerd. In de vragenlijst komen onder meer aan de orde: typering van de samenwerking met basisscholen en andere instellingen, de kindpopulatie, opleiding en werkervaring van de teamleden, omvang van de ruimtes. De structurele aspecten groepsgrootte en staf-kind ratio zijn tevens meegenomen in het observatie-instrument.

2.4 Uitgangspunten en inhoud van het meetinstrument Gastouderopvang

Een meetinstrument voor de pedagogische kwaliteit van de gastouderopvang zou zich, net zoals de meet instrumenten voor de kinderdagopvang als de buitenschoolse opvang, moeten richten op zowel de proceskwaliteit als op de structurele kwaliteitskenmerken. Bij meting van de proceskwaliteit zou

aandacht moeten worden besteed aan de algemene proces kwaliteit en de zes interactievaardigheden van gastouders (sensitieve responsiviteit, respect voor de autonomie van de kinderen, structureren en grenzen stellen, praten en uitleggen, ontwikkelingsstimulering en het begeleiden van interacties tussen kinderen). Als belangrijke structurele kwaliteitskenmerken werden in het NCKO-model tien kenmerken genoemd, waaronder pedagogisch beleid, pedagogische ondersteuning, staf-kind ratio, groepsgrootte, en staf- en groepsstabiliteit. Voor het vaststellen van de kwaliteit van de interactievaardigheden en de algemene proceskwaliteit zijn reeds instrumenten beschikbaar die kunnen worden toegepast. Het betreft het bestaande FCCERS-R instrument, dat speciaal is vertaald en aangepast aan de Nederlandse situatie. Voor het meten van de interactievaardigheden zijn bestaande instrumenten, de NCKO-observatieschalen en de BSO-interactieschalen, aangepast en gebruikt. In de volgende onderdelen van deze paragraaf worden de instrumenten nader toegelicht en in onderstaand schema is te zien hoe deze zich verhouden tot de instrumenten voor de kinderopvang en buitenschoolse opvang.

Voor het meten van de structurele kwaliteitskenmerken behoeften geen nieuwe instrumenten ontwikkeld te worden, omdat betrouwbare en valide meting hiervan relatief eenvoudig is en er voldoende geschikte instrumenten voorhanden zijn.

Voor het meetinstrument voor de gastouderopvang worden dezelfde kaders en inhoud - met aanpassingen voor de gastoudercontext - gebruikt als in het NCKO- en BSO-instrumentarium.

(a) Algemene proceskwaliteit

Voor het vaststellen van de **algemene proceskwaliteit** in de gastouderopvang is gekozen voor het bestaande meetinstrument Family Child Care Environment Rating Scale Revised Edition (FCCERS-R; Harms, Cryer, & Clifford, 2007). Dit instrument is van dezelfde ontwikkelaars als de internationaal erkende en toegepaste meetinstrumenten voor de kinderdagopvang (0- tot 4-jarigen), namelijk de Infant/Toddler Environment Rating Scale (ITERS, voor kinderen van 0 tot 2.5 jaar) en de Early Childhood Environment Rating Scale (ECERS, voor kinderen van 2.5 tot 5 jaar). De voornaamste reden voor de keuze van dit instrument is dat het equivalent van de FCCERS-R, de ITERS-R en ECERS-R, al enkele jaren door het NCKO worden gebruikt om de algemene proceskwaliteit van de Nederlandse kinderdagopvang in landelijke kwaliteitsmetingen vast te stellen. Dit maakt een vergelijking van de algemene proceskwaliteit van de gastouderopvang en de dagopvang mogelijk. Bovendien is het instrument, na training, relatief makkelijk te hanteren (met het oog op vertaling naar het veld). De FCCERS-R meten, net als de instrumenten voor de dagopvang (ITERS-R/ECERS-R) en de buitenschoolse opvang (Observatielijst pedagogische kwaliteit leefomgeving), een aantal aspecten

van de proceskwaliteit van de gastouderopvang, dat wil zeggen de kwaliteit van de dagelijkse ervaringen van de kinderen bij de gastouder. Daarnaast levert de FCCERS-R ook een totaalscore op voor de *overall* kwaliteit van de opvang.

Op basis van een observatie van minimaal 2.5 uur wordt voor 19 verschillende items een score toegekend op een 7-puntsschaal. De items zijn ondergebracht in 6: Ruimte en Meubilering (6 items), Dagelijkse Individuele Verzorging (1 item), Luisteren en Praten (3 items), Activiteiten (11 items), Interacties (4 items) en Programma (4 items). De subschalen Dagelijkse Individuele Verzorging (met uitzondering van 1 item) en Ouders en Staf werden, in overeenstemming met het gebruik van de ITERS-R en ECERS-R voor de dagopvang, in het gastouder-meetinstrument niet meegenomen.

De betrouwbaarheid van de oorspronkelijke FCCERS-R is door de ontwikkelaars onderzocht en voldoende bevonden (zie Harms, Cryer, & Clifford, 2007, blz 2-6).

Hieronder wordt per subschaal een overzicht gegeven van alle items die in het meetinstrument voor de gastouderopvang zijn opgenomen:

Ruimte en Meubilering

1. Binnenruimte
2. Meubilering voor dagelijkse verzorging en spel
3. Voorzieningen voor ontspanning en comfort
4. Indeling van de ruimte
5. Aankleding van de ruimte
6. Ruimte voor privacy

Dagelijkse individuele verzorging

7. Brengen en halen

Luisteren en praten

8. Kinderen helpen taal te begrijpen
9. Kinderen helpen taal te gebruiken
10. Het gebruik van boeken

Activiteiten

11. Fijne motoriek
12. Knutselen
13. Muziek en beweging
14. Blokken

15. Rollenspel

16. Rekenen/cijfers
17. Natuur/wetenschap
18. Zand en waterspel
19. Bevorderen van het accepteren van verscheidenheid
20. Gebruik van tv, video en/of computer
21. Actief lichamenlijk spel

Interacties

22. Toezicht op spel en ontwikkeling
23. Gastouder-kind interacties
24. Discipline (handhaven orde)
25. Interacties tussen kinderen

Programma

26. Dagschema/dagindeling
27. Vrij spel
28. Groepstijd
29. Voorzieningen voor kinderen met een handicap

(b) Interactievaardigheden

Voor het in kaart brengen van **de kern van algemene proceskwaliteit**, de interactievaardigheden van de gastouders, zijn de interactieschalen uit NCKO- en BSO-instrumentarium in het instrument, met enige aanpassingen aan de nieuwe doelgroep, voor de gastouderopvang opgenomen. Met de

observatieschalen kunnen de zes interactievaardigheden (sensitieve responsiviteit, respect voor de autonomie van het kind, structureren en grenzen stellen, praten en uitleggen, ontwikkelingsstimulering en begeleiden van interacties tussen kinderen) met behulp van videofragmenten van verschillende opvoed- en verzorgingssituaties ook in de gastouderopvang in kaart worden gebracht. De NCKO-interactieschalen (zie beschrijving hieronder) vormen het vertrekpunt van de gastouderinteractieschalen, omdat de meerderheid van de kinderen die wordt opgevangen bij gastouders in de leeftijd van 0-4 jaar is. De NCKO-interactieschalen zijn aangevuld met relevante aspecten voor 5-12 jarigen uit de BSO-interactieschalen. Op deze manier meet het instrument de pedagogische kwaliteit van de gastouderopvang voor de complete leeftijdsrange van 0-12 jaar.

Beschrijving NCKO-interactievaardigheden

Sensitieve responsiviteit

Sensitieve responsiviteit verwijst naar de mate waarin de opvoeder adequaat ingaat op signalen die aangeven dat het kind zich niet goed voelt of behoefte heeft aan contact of aan een andere vorm van emotionele ondersteuning. Sensitieve responsiviteit vereist dat de opvoeder oog heeft voor de toestand en de signalen van het kind, dat ze deze goed weet te interpreteren en dat ze er tijdig en adequaat op reageert, zodat het kind zich begrepen, geaccepteerd en veilig voelt.

Een opvoeder die *hoog* scoort op deze schaal geeft emotionele steun aan alle kinderen die deze steun op een bepaald moment nodig hebben, zowel tijdens stressvolle als niet stressvolle situaties. Zij herkent de signalen van de kinderen en reageert hier over het algemeen tijdig en adequaat op. De opvoeder toont belangstelling voor de activiteiten van de kinderen en laat zien dat zij zich bewust is van de behoeften, stemmingen en mogelijkheden van de kinderen. Zij vormt een veilige haven voor de kinderen.

Een opvoeder die *laag* scoort op deze schaal slaagt er niet in emotionele steun aan de kinderen te geven wanneer zij die nodig hebben. Ofwel zij herkent de signalen van de kinderen niet, ofwel zij herkent de signalen, maar reageert niet of heel langzaam of inadequaat. Haar respons op de signalen van de kinderen, als die er al is, past over het algemeen niet bij de situatie, de emoties of het ontwikkelingsniveau van de kinderen. De opvoeder kan onverschillig of afstandelijk gedrag vertonen.

Respect voor autonomie

Respect voor de autonomie van een kind verwijst naar de mate waarin een opvoeder de individualiteit, motieven, en perspectieven van het kind erkent en respecteert.

Een opvoeder die *hoog* scoort op deze schaal geeft hen zoveel mogelijk de kans om dingen zelf te doen en zelf uit te proberen. Ze stimuleert de kinderen om problemen zelf op te lossen en

respecteert ook hun eigen oplossingen en ideeën. Ze commandeert kinderen niet, maar vraagt hen om iets voor haar te doen of om mee te helpen.

Een opvoeder die *laag* scoort op deze schaal is intrusief, dringt haar eigen ideeën en plannen op en houdt onvoldoende rekening met (of negeert) de intenties of drang naar zelfstandigheid van de kinderen. Een gebrek aan respect voor de autonomie uit zich in het abrupt ingrijpen in de activiteiten van kinderen, het afpakken van dingen, het zonder aankondiging of toelichting optillen of verplaatsen van kinderen

Structureren en grenzen stellen

Structureren en grenzen stellen verwijst naar de vaardigheid van een leidster om aan kinderen duidelijk te maken wat er van hen verwacht wordt en er tevens voor te zorgen dat zij zich daar ook aan houden.

Een opvoeder die *hoog* scoort op deze schaal, maakt de kinderen goed duidelijk wat er van hen verwacht wordt door situaties, activiteiten en taken zodanig te structureren dat zij deze kunnen overzien. Zij zorgt ervoor dat de kinderen zich aan de regels houden door het stellen van voldoende grenzen. Zij doet dat ook tijdig, zodat kinderen niet het gevoel krijgen dat zij voortdurend ‘in overtreding’ zijn; bij het stellen van grenzen is zij bovendien consistent en doortastend. Als kinderen negatief gedrag vertonen, reageert zij adequaat, d.w.z. in overeenstemming met de aard van de overtreding en het ontwikkelingsniveau van de kinderen. Bovendien besteedt ze aandacht aan de consequenties van het negatieve gedrag voor anderen door het geven van uitleg en het aanbieden van alternatieven.

Een opvoeder die *laag* scoort op deze schaal slaagt er niet in de kinderen duidelijk te maken wat er van hen verwacht wordt. Zij structureert situaties, activiteiten en taken onvoldoende. Het is voor de kinderen niet duidelijk aan welke regels zij zich dienen te houden; het kan zijn dat er teveel of te weinig regels zijn, een inconsistent gebruik van regels, of regels die niet in overeenstemming zijn met het ontwikkelingsniveau van de kinderen. Als kinderen negatief gedrag vertonen, reageert zij inadequaat, d.w.z. dat haar reacties niet in overeenstemming zijn met de aard van de overtreding of het ontwikkelingsniveau van de kinderen. Er kan sprake zijn van zeer negatieve reacties, zoals schreeuwen of fysiek straffen. Ook een opvoeder die overdreven veel controle op het gedrag van de kinderen probeert uit te oefenen, scoort laag op deze schaal.

Praten en uitleggen

Deze vaardigheid betreft de kwaliteit van de verbale interacties tussen opvoeder en kind, waarbij het niet alleen gaat om de frequentie van die interacties maar ook om de vorm en inhoud ervan.

Een opvoeder die *hoog* scoort op deze schaal, praat vaak met de kinderen, stimuleert de kinderen om hun gedachten en gevoelens te verwoorden en stemt de timing en inhoud van haar verbale interacties af op het begripsniveau van de kinderen en de activiteiten en interesses die de kinderen op dat moment bezig houden.

Een opvoeder die *laag* scoort op deze schaal, praat nauwelijks of niet en reageert nauwelijks of niet op verbale of non-verbale initiatieven van de kinderen. De opvoeder doet geen moeite om de kinderen te verstaan en stimuleert het taalgebruik van kinderen niet. Ook een opvoeder die zich voornamelijk negatief uit tegen de kinderen krijgt een lage score op deze schaal.

Ontwikkelingsstimulering

Ontwikkelingsstimulering heeft vooral betrekking op de stimulering van de persoonlijke competentie van kinderen. Het gaat hier om de “extra” dingen die de leidster doet om motorische, cognitieve en taalontwikkeling en de creativiteit van de kinderen te stimuleren. Zij kan dat bijvoorbeeld doen door de aandacht van kinderen op bepaalde dingen te richten, door nieuwe activiteiten of nieuw spel materiaal aan te bieden of door op nieuwe mogelijkheden van het spel materiaal te wijzen. Ze kan taal- of bewegingsspelletjes initiëren of samen met de kinderen muziek maken of naar muziek luisteren.

Een opvoeder die *hoog* scoort op deze schaal, ziet ontwikkelingsstimulering als één van haar doelen in haar interacties met kinderen. Deze opvoeder grijpt dagelijkse situaties aan om de ontwikkeling te stimuleren, maar biedt ook uitdagende activiteiten en materialen aan die passen bij het ontwikkelingsniveau van de kinderen. Een opvoeder die goed is in het stimuleren van de ontwikkeling van kinderen, biedt niet alleen veel “extra” stimulering, maar stemt die stimulering ook af op de interesse, het ontwikkelingsniveau en de toestand van de kinderen, waardoor zij overstimulering van de kinderen voorkomt.

Een opvoeder die *laag* scoort op deze schaal, ziet ontwikkelingsstimulering niet als een prioriteit in haar interacties met kinderen. Zij beperkt haar interacties met de kinderen tot verzorgingssituaties en situaties waarin kinderen aangeven hen nodig te hebben, en laat de kinderen vooral vrij spelen. Een opvoeder die exploratie belemmert, ontmoedigt, beperkt of consequent niet afstemt op de interesse en het ontwikkelingsniveau van de kinderen krijgt ook een lage score op deze schaal.

Begeleiden van interacties

Begeleiden van interacties verwijst naar de vaardigheid van een leidster om positieve interacties tussen kinderen te begeleiden en bevorderen.

Een opvoeder die *hoog* scoort op deze schaal gaat op een adequate manier om met positieve interacties van kinderen. Ze merkt spontane positieve interacties op en beoordeelt deze positief. Bovendien probeert ze positieve interacties te stimuleren door veel uitleg te geven, actief geschikte situaties te creëren om positieve interacties te bevorderen en zelf het goede voorbeeld te geven.

Een opvoeder die *laag* scoort op deze schaal begeleidt de kinderen niet of nauwelijks in hun interacties met andere kinderen. Ze besteedt nauwelijks of geen aandacht aan positieve interacties en stimuleert de kinderen ook niet in het tonen van positieve interacties. Een opvoeder die positieve interacties wel opmerkt maar stopt of afremt krijgt ook een lage score op deze schaal.

(a) Structurele kwaliteit

Ook de **structurele kenmerken** vormen een onderdeel in het instrument voor de gastouderopvang: kenmerken van de gastouderopvang en de gastouders die van invloed kunnen zijn op de proceskwaliteit, zoals groeps grootte, de gastouder-kind ratio en de opleiding en werkervaring van de gastouders worden opgenomen in het instrument. Deze aspecten worden met behulp van vragenlijsten of observaties gemeten.

Drie meetinstrumenten kinderopvang (0-12 jaar)

In onderstaand schema hebben we de instrumenten van de kinderdagopvang (0-4 jaar), de buitenschoolse opvang (bso, 4-12 jaar) en de gastouderopvang (0-12 jaar) naast elkaar gezet.

Schema 1: Overzicht van de vergelijkbaarheid van de drie meetinstrumenten (KDV, BSO en GO)

Deelinstrumenten	KDV	BSO	Gastouderopvang
Algemene proceskwaliteit (observatielijst)	ITERS-R/ECERS-R - Meubilering en inrichting - Taal - Activiteiten - Programma - Interacties	Observatielijst pedagogische kwaliteit leefomgeving - Groepen - Ruimte en inrichting - Dagindeling, spelmaterialen en activiteiten (per ontwikkelingsgebied) - Interacties - Groepsfunctioneren. - Pedagogisch beleid en organisatie - Samenwerking (met ouders, basisscholen en andere instellingen) - Kinderen met speciale behoeften	FCCERS-R - Ruimte en inrichting - Taal - Activiteiten - Programma en structuur - Interacties
Interactievaardigheden (Observatieschalen voor de beoordeling van videofragmenten)	NCKO-Interactieschalen - Sensitieve responsiviteit - Respect voor autonomie - Structuur en grenzen - Praten en uitleggen - Ontwikkelingsstimulering - Begeleiden van interacties	BSO/NCKO Interactieschalen - Emotionele ondersteuning - Respect voor autonomie - Regie en leiding - Praten en uitleggen - Ontwikkelingsstimulering - Begeleiden van (sociale) interacties	GO-Interactieschalen - Sensitieve responsiviteit - Respect voor autonomie - Structuur en grenzen - Praten en uitleggen - Ontwikkelingsstimulering - Begeleiden van interacties
Structurele kwaliteit (vragenlijst en observaties)	- Groepsgrootte - Staf-kindratio - Leeftijdsopbouw groep - Opleiding - Werkervaring - Leeftijd - etc	- Groepsgrootte - Staf-kindratio - Leeftijdsopbouw groep - Opleiding - Werkervaring - Leeftijd - etc	- Groepsgrootte - Staf-kindratio - Leeftijdsopbouw groep - Opleiding - Werkervaring - Leeftijd - Pedagogisch netwerk - Aantal uren opvang - Flexibiliteit van de opvang

			- <i>Aanwezige familieleden etc.</i>
--	--	--	--------------------------------------

Met het hierboven beschreven instrument voor de gastouderopvang kan tevens worden nagegaan of de kwaliteit van de gastouderopvang voldoet voor de leeftijdsgroepen van 2-4 jaar gelet op ontwikkelingsgericht werken. Dat is mogelijk doordat het beoogde instrument indicatoren bevat waarvan in ontwikkelingspsychologisch onderzoek is aangetoond dat zij bijdragen aan de ontwikkeling en het welzijn van kinderen tussen 0-12 jaar, dus ook aan de specifieke ontwikkeling en het welzijn van de leeftijdsgroep 2-4 jaar. Deze indicatoren, beschreven in NCKO-kwaliteitsmodel (Riksen-Walraven, 2004), hebben betrekking op kwaliteitsaspecten zoals de aanwezigheid van adequaat spel- en ontwikkelingsmateriaal en het in voldoende mate voorkomen van rijke en stimulerende interacties tussen kinderen en opvoeder.

Verder kan met het beoogde instrument voor de gastouderopvang specifiek worden ingezoomd op de meer 'educatieve vaardigheden' van de gastouder zoals praten en uitleggen, ontwikkelingsstimulering en het begeleiden van interacties tussen kinderen. Vooral de vaardigheid 'ontwikkelingsstimulering' geeft inzicht in de mate waarin gastouders ontwikkelingsgericht kunnen werken. Daarnaast bevatten verschillende onderdelen uit de observatielijst FCCERS-R zoals (1) Ruimte en meubilering, (2) Taal, (3) Activiteiten, (4) Interacties en (5) Programma, aspecten die relevant zijn om het ontwikkelingsgericht werken voor 2-4 jarigen van de gastouder in kaart te brengen.

3 DOEL VAN DE BETROUWBAARHEIDS- EN VALIDATIESTUDIE

De in dit hoofdstuk beschreven validatiestudie had tot doel het toetsen van de twee belangrijkste psychometrische kwaliteitskenmerken van de gastouder-interactieschalen (GO-interactieschalen) en de FCCERS-R, namelijk de betrouwbaarheid en de validiteit.

De *betroouwbaarheid* van een meetinstrument zegt iets over de kwaliteit van de meting, met name de mate waarin de meting consistent en herhaalbaar is. In dit onderzoek hebben we drie veelgebruikte betrouwbaarheidsindicatoren bepaald, namelijk:

1. De *inter-observer* betrouwbaarheid van de meetinstrumenten ofwel de mate waarin verschillende (getrainde) observatoren zelfde scores toekennen op de schalen op basis van dezelfde observatie-episode.
2. De *test-hertest* betrouwbaarheid van de meetinstrumenten ofwel de mate waarin observatie op verschillende meettijdstippen dezelfde schaalscores oplevert. Hierbij wordt ervan uitgegaan dat het gemeten kenmerk (bv. de sensitieve responsiviteit van de gastouder of de binnenruimte) relatief stabiel is over tijd, zodat inconsistentie van de observatie-uitkomsten over tijd inderdaad duidt op een gebrek aan precisie van het meetinstrument, en niet op een werkelijke verandering in het geobserveerde fenomeen. Daarom moeten de twee tijdstippen waarop gemeten wordt om de test-hertest betrouwbaarheid van een meetinstrument te bepalen niet al te ver uiteen liggen in de tijd. In deze validatiestudie is, vanwege een strakke en complexe tijdsplanning van het project, gekozen voor twee opeenvolgende metingen met een tussentijd van twee weken bij een klein aantal gastouders.
3. De *interne consistentie* van de meetinstrumenten, ofwel de mate waarin de schalen onderling samenhangen. Het uitgangspunt is hierbij dat schalen die verschillende aspecten van hetzelfde begrip meten (bv. interactievaardigheden van de gastouder of algemene proceskwaliteit) onderling significant zullen samenhangen. Omdat het gaat om gerelateerde maar verschillende aspecten is de samenhang echter naar verwachting niet sterk. Daarnaast werd ook een significante doch matige samenhang verwacht tussen de interactieschalen enerzijds en de algemene proceskwaliteit anderzijds, omdat beide schalen indicatoren pretenderen te zijn van eenzelfde begrip, namelijk de pedagogische kwaliteit van de gastouderopvang.

De *validiteit* van een meetinstrument geeft aan of het instrument inderdaad datgene lijkt te meten wat het bedoelt te meten. In deze validatiestudie hebben we ons gericht op convergente validiteit als

belangrijk empirisch criterium voor nieuwe meetinstrumenten of meetinstrumenten met een nieuwe toepassing.

De *convergente validiteit* van de meetinstrumenten is de mate waarin de scores op de schalen samenhangen (“convergeren”) met scores die verkregen zijn met behulp van andere instrumenten die hetzelfde pretenderen te meten. De convergente validiteit van de GO-interactieschalen is bepaald door de scores op deze schalen te relateren aan inhoudelijk vergelijkbare scores die op basis van dezelfde observatie-episodes zijn toegekend met behulp van één ander meetinstrument. Om van convergente validiteit te kunnen spreken moeten de correlaties tussen de GO-interactieschalen en het betreffende valideringsinstrument significant zijn. Erg hoge correlaties worden echter niet verwacht, omdat de GO-interactieschalen juist ontworpen zijn om net iets andere kenmerken te meten dan het valideringsinstrument; zeer hoge correlaties zouden erop duiden dat de GO-interactieschalen hetzelfde meten als bestaande instrumenten. Om de convergente validiteit van de FCCERS-R te bepalen hebben we de scores van de gastouder op deze schalen gerelateerd aan inhoudelijk vergelijkbare scores die op basis van dezelfde observatie-episodes toegekend zijn met behulp van een ander instrument voor beoordeling van de algemene proceskwaliteit. De meetinstrumenten die gebruikt zijn bij de validering van de GO-interactieschalen en de FCCERS-R staan beschreven in de methode sectie van dit hoofdstuk (4.4) en een gedetailleerd overzicht van de getoetste hypothesen is te vinden in Tabel 4.1 voor de GO-interactieschalen en voor de FCCERS-R.

4 METHODE

In de validatiestudie is een onderzoeksopzet gehanteerd waarbij een deel van de gastouders tweemaal zijn bezocht (test-hertest design). In de eerste ronde is bij de kwaliteit van de gastouderopvang gemeten met het nieuwe gastouder-instrumentarium (de GO-interactieschalen en de FCCERS-R), aangevuld met een bestaand instrumenten om ter validering een vergelijking te kunnen maken tussen de bestaande instrumenten en het nieuwe instrument. In de tweede ronde is bij een deel van de gastouders de kwaliteit opnieuw gemeten, nu alleen met de GO-interactieschalen en de FCCERS-R, om de stabiliteit (c.q. test-hertest betrouwbaarheid) van het nieuwe instrument te kunnen bepalen.

4.1 Steekproef

In februari tot en met september van 2014 zijn er 30 kleine en 21 grote gastouderbureaus in en rondom de Randstad telefonisch benaderd met de vraag of hun gastouders mee willen werken aan de ontwikkeling van een meetinstrument voor de gastouderopvang. Wanneer gastouderbureaus hier positief over gestemd waren, kregen zij via de email een brief met aanvullende informatie over het onderzoek om aan de gastouders te sturen. In de weken daarna werd er contact onderhouden met de meewerkende gastouderbureaus om op deze manier contactgegevens te verzamelen van bereidwillige gastouders.

Bij bereidheid tot medewerking van de gastouder werd er telefonisch contact gezocht met de gastouders en werd het project nogmaals toegelicht en werden eventuele vragen van de gastouder beantwoord. Vervolgens werd er een afspraak gemaakt voor een bezoek door 2 medewerkers van het Kohnstamm Instituut. Er zijn 50 gastouders benaderd om mee te werken aan het project en uiteindelijk daarvan zijn 40 gastouders bezocht. Redenen waarom afspraken met gastouders niet door konden gaan waren in de meeste gevallen organisatorisch van aard.

Op één bezoek na hebben alle bezoeken bij de gastouder thuis plaatsgevonden. Eén bezoek aan de gastouder vond plaats bij de ouders thuis waar de gastouder de kinderen opving. Voor elk kind bij de deelnemende gastouder werd een toestemmingsbrief voor de ouders aan de gastouder gemaild. In deze brief werden het doel van het project en de algemene procedure beschreven. Daarnaast werd toestemming gevraagd aan de ouders om het kind te mogen filmen. In totaal namen 40 gastouders en 160 kinderen deel aan het project.

4.2 Procedure

De gastouders en kinderen werden bezocht door 2 getrainde observatoren. Het bezoek aan de gastouder duurde van ongeveer 7:30 uur tot na de lunch (tussen 12:00 en 13:00 uur) of van 15.00 tot 18.00 uur. Observator 1 nam de FCCERS-R af (zie hoofdstuk 4.3). Hiertoe maakte hij/zij gedurende de ochtend of middag aantekeningen met betrekking tot de dagelijkse ervaringen van de kinderen in de ruimte (inclusief inrichting van de ruimte, beschikbare materialen, programma, routines, etc.). Observator 2 filmde de gastouders en de kinderen, om in een latere fase de interactie tussen de gastouder en de kinderen te kunnen beoordelen met behulp van de GO-interactieschalen (zie hoofdstuk 4.3). Daarnaast nam observator 2 de HOME af (zie hoofdstuk 4.4). Beide observatoren stelden zich afzijdig op om het groepsgebeuren zo min mogelijk te beïnvloeden. De gastouders werd gevraagd om het reguliere programma te volgen.

Iedere gastouder werd gedurende het bezoek in vier verschillende situaties gefilmd (verzorging, eten met de kinderen, spel, en transitie). Iedere filmepisode duurde doorgaans 8 tot 10 minuten. De individuele opnames van de gastouders werden op een later tijdstip door getrainde observatoren beoordeeld. Aan het eind van de observatie hielden beide observatoren een kort vraaggesprek met de gastouder voor nadere informatie die niet via directe observatie te verkrijgen was. De individuele opnames van de gastouders en kinderen werden op een later tijdstip door getrainde onafhankelijke observatoren beoordeeld (de codeur van de GO-interactieschalen was niet dezelfde als een van de observatoren die de gastouder hadden bezocht).

Na ongeveer 2 tot 4 weken werden 8 gastouders een tweede keer bezocht door een observator. De werving van deze her-test groepen was niet gemakkelijk. Veranderde omstandigheden, vakantieperiode en ziekte waren vaak voorkomende redenen om af te zeggen. Om ervoor te zorgen dat de groepssamenstelling zoveel mogelijk hetzelfde zou zijn, vond het tweede bezoek plaats op dezelfde dag van de week als het eerste bezoek. Observator 1 nam wederom de FCCERS-R af en filmde de gastouders en kinderen in dezelfde vier situaties (verzorging, eten met de kinderen, spel, en transitie).

Vóór het eerste bezoek ontving iedere gastouder een vragenlijst waarin vragen waren opgenomen met betrekking tot de gastouder zelf (achtergrondinformatie over opleiding, leeftijd, bestaansduur, etc).

4.3 Te valideren instrumenten: Gastouder-instrument

Gastouder-Interactieschalen (GO-interactieschalen)

De gefilmde gastouder-episodes werden beoordeeld met behulp van de zes GO-interactieschalen (NCKO/Kohnstamm Instituut, 2014): *sensitieve responsiviteit, respect voor autonomie, structureren en grenzen stellen, praten en uitleggen, ontwikkelingsstimulering en begeleiden van interacties tussen kinderen*. Deze interactievaardigheden worden gemeten op een 7-puntschaal, waarbij 7 = *zeer aanwezig* en 1 = *zeer afwezig*.

Vijf observatoren werden getraind om de interactieschalen betrouwbaar te kunnen gebruiken. Drie observatoren volgden 4 trainingssessies die elk gemiddeld 3 uur duurden (in totaal dus 12 uur). In deze sessies werd aandacht besteed aan het betrouwbaar scoren van de interactieschalen en aan de procedures rondom het filmen van de gastouders. Een trainingssessie bestond uit het intensief bestuderen van de schaal die op die dag centraal stond en vervolgens het beoordelen en nabespreken van oefenfragmenten. De training werd afgesloten met twee toetsen. De eerste toets bestond uit 8 fragmenten van 10 minuten uit het NCKO-onderzoek. Bij voldoende mate van overeenkomst met de consensuscores (meer dan 80%) werd een tweede toets afgenomen. De opnames van de tweede toets waren gemaakt bij de gastouders van verschillende opvoed- en verzorgingssituaties. Bij twee al eerder getrainde observatoren op de NCKO-interactieschalen werd alleen de tweede toets afgenomen. De consensuscore (ofwel de interobservator betrouwbaarheid) voor ieder fragment (8) was bepaald door 3 experts. Het bepalen van de mate waarin de observatoren het eens zijn met de consensuscore behoorde tot één van de doelen van de validatiestudie. De resultaten van de toets worden daarom beschreven bij de resultaten van de validatiestudie. In totaal waren vier observatoren (2 nieuwe observatoren + 2 al eerder getrainde observatoren) verantwoordelijk voor het beoordelen van de videofragmenten van de test en hertest.

De totale steekproef bestaat uit 40 gastouders. Voor elke gastouder werd een gemiddelde score per interactieschaal berekend over de gefilmde episodes. Daarnaast werden de scores voor elke schaal gestandaardiseerd en daarna gemiddeld om tot een totaalscore over alle interactieschalen te komen (GO-interactieschalen *Totaalscore*).

FCCERS-R

De Family Child Care Environment Rating Scale Revised Edition (FCCERS-R: Harms, Cryer & Clifford, 2007) werd vertaald vanuit het Engels en aangepast aan de Nederlandse situatie en is ontwikkeld om de algemene proceskwaliteit, oftewel de dagelijkse ervaringen van de kinderen in de

gastouderopvang, te beoordelen. De originele schaal wordt internationaal gebruikt in onderzoeken naar de proceskwaliteit van de gastouderopvang.

De FCCERS-R is bedoeld voor opvanggroepen met kinderen in de leeftijd van 0 tot 12 jaar. Het meetinstrument bestaat uit 7 subschalen, te weten: (a) Ruimte en Meubilering, (b) Dagelijkse Individuele Verzorging, (c) Luisteren en Praten, (d) Activiteiten, (e) Interacties, (f) Programma en (g) Voorzieningen voor Ouders en Staf. De items worden gepresenteerd op een 7-puntsschaal met beschrijvingen voor 1 (inadequaat), 3 (minimaal), 5 (goed) en 7 (uitstekend). *Inadequaat* beschrijft opvang die niet voldoet aan de basale verzorgingsbehoefte, *minimaal* beschrijft opvang die voldoet aan de basale verzorgings- en ontwikkelingsbehoefte, *goed* beschrijft de opvang die verder gaat en gericht is op de ontwikkeling van de kinderen en *uitstekend* beschrijft opvang van hoge kwaliteit met individuele/persoonlijke zorg. De beoordeling wordt gebaseerd op de situatie zoals die geobserveerd of gerapporteerd wordt gedurende een bezoek van minimaal 3 uur en wordt afgesloten met een vraaggesprek met de gastouder. Voor elk item wordt een score gegeven van 1 tot 7. De totaalscore is de gemiddelde score van alle items. Alleen de totaalscore en de subschalen Ruimte en Meubilering, Luisteren en Praten, Interacties en Activiteiten en Programma werden gebruikt in deze studie.

Vier observatoren werden getraind op de FCCERS-R. De training vond plaats vóór de start van de dataverzameling en werd verzorgd door een van de vertalers van de FCCERS-R en tevens expert op de ITERS-R en ECERS-R. Na een algemene introductie participeerde iedere observator in ten minste 4 veldobservaties onder supervisie van de expert trainer. Iedere observatie werd gevolgd door een nabespreking met de trainer, waarna de inter-observator betrouwbaarheid werd bepaald. Een observator werd beschouwd als betrouwbaar als tijdens ieder van drie opeenvolgende bezoeken aan gastouders 80% van de scores (binnen 1 schaalpunt) overeen kwamen met de score van de trainer. Observatoren bezochten tijdens de training gemiddeld 4 gastouders (range 4 – 5). Het gemiddelde betrouwbaarheidspercentage over de drie opeenvolgende bezoeken (>80%) was 82% (range 80% tot 93%).

4.4 Instrumenten voor validering

Instrumenten voor convergente validiteit

HOME

Van de Home Observation for Measurement of the Environment: Inventory Administration Manual (HOME, Cadwell & Bradley, 2003) werden de Child Care Infant/Toddler Home (CC-IT-HOME) en de Child Care Early Childhood Home (CC-EC-HOME) gebruikt om de convergente validiteit van de

gastouder-interactieschalen schalente toetsen. De CC-IT-HOME en de CC-EC-HOME zijn ontwikkeld om de kwaliteit en kwantiteit van stimulering en ondersteuning aan kinderen in de gastouderopvang in kaart te brengen. De CC-IT-HOME is bedoeld voor gastouderopvang voor kinderen van 0-4 jaar en de (CC-EC-HOME) richt zich op gastouderopvang voor kinderen van 4-6 jaar. Als de meerderheid van de kinderen 0-4 jaar was dan werd de CC-IT-HOME gebruikt en als de meerderheid van de kinderen 4-6 jaar was, werd de CC-EC-HOME gebruikt.

De complete CC-IT-HOME bestaat uit 6 subschalen (totaal 43 items die gescoord worden op een dichotome schaal; 0 = niet aanwezig, 1 = wel aanwezig), te weten: Responsiviteit, Acceptatie, Organisatie, Materialen voor Leeractiviteiten, Betrokkenheid en Variëteit. De totaalscore is de optelsom van alle scores. Uit de CC-IT-Home zijn 4 items niet meegenomen omdat deze betrekking hadden op aspecten die niet relevant worden geacht voor deze validatie studie of waarvan via een andere manier informatie werd verkregen. De totaalscore is gebruikt voor de validering. De interne consistentie (Cronbach's α , een maat voor onderlinge samenhang) van de totaalscore van de CC-IT-HOME schaal was adequaat (Cronbach's $\alpha = .64$). De afzonderlijke subschalen lieten relatief lage scores zien doordat de schalen relatief kort zijn en door relatief hoge scores waardoor er vrijwel geen variantie is. Een samengestelde HOME-score voor Interacties is voor deze studie nog bepaald op basis van alle items die betrekking hebben op de omgang tussen de gastouder en de kinderen; deze score maakt het mogelijk de convergente validiteit te bepalen met de Interactie-schaal uit het FCCERS-instrument. De interne consistentie van deze aparte schaal was adequaat (Cronbach's $\alpha = .64$).

De complete CC-EC-HOME bestaat uit 7 subschalen (totaal 58 items die gescoord worden op een dichotome schaal; 0 = niet aanwezig, 1 = wel aanwezig), te weten: Materialen voor Leeractiviteiten, Taalstimulering, Fysiek straffen, Responsiviteit, Stimulering van schoolvaardigheden, Variëteit en Acceptatie. De totaalscore is de optelsom van alle scores (0 of 1). Uit de CC-EC-Home zijn de subschaal 'Modelgedrag' (7 items) en 6 items niet meegenomen omdat deze betrekking hadden op aspecten die niet relevant worden geacht voor deze validatiestudie of waarvan via een andere manier informatie werd verkregen. De totaalscore is gebruikt in de validering. De interne consistentie hiervan was goed (Cronbach's $\alpha = .81$). De afzonderlijke subschalen lieten relatief lage scores zien doordat de schalen relatief kort zijn en door relatief hoge scores waardoor er weinig variantie is. Ook voor de CC-EC-HOME is aanvullend een samengestelde subschaal Interacties bepaald op basis van alle items die specifiek betrekking hebben op de omgang van de gastouder met de kinderen (Cronbach's $\alpha = .73$).

Vier observatoren volgden twee introductie bijeenkomsten over de CC-IT-HOME en de CC-EC-HOME en voerden vervolgens in tweetallen, twee proefobservaties bij gastouders uit. De proefobservaties werden met de vier observatoren onder supervisie van een expert nabesproken.

4.5 Concrete verwachtingen

Omdat we ervan uitgaan dat de zes GO-interactieschalen verschillende aspecten van hetzelfde begrip meten, namelijk “interactievaardigheden”, werd verwacht dat de scores van de gastouders op de verschillende schalen onderling significant zouden samenhangen; omdat het om *verschillende* aspecten van hetzelfde begrip gaat, werd echter geen erg hoge correlatie verwacht. Ten slotte werd een significante doch matige samenhang verwacht tussen de GO-interactieschalen enerzijds en de FCCERS-R anderzijds, omdat beide schalen indicatoren zijn van eenzelfde begrip, namelijk de pedagogische kwaliteit van de gastouderopvang.

De tabellen 4.1 en 4.2 geven een overzicht van de verwachtingen met betrekking tot de validiteit voor de GO-interactieschalen en de FCCERS-R, inclusief een aanduiding van de schalen die gebruikt werden in deze validatiestudie.

Tabel 4.1
Verwachtingen m.b.t. validiteit GO-interactieschalen

Validatie-instrument	Verwachtingen
<i>Convergente validiteit</i>	
▪ FCCERS-R	<ol style="list-style-type: none">1. Significante positieve correlatie tussen de GO-interactieschalen <i>Totaalscore</i> en de FCCERS-R <i>Totaalscore</i>.2. Significante positieve correlatie tussen de GO-interactieschalen <i>Totaalscore</i> en de FCCERS-R subschaal <i>Interacties</i>3. Significante positieve correlatie tussen GO-interactieschaal <i>Praten en uitleggen</i> enerzijds en FCCERS-R subschalen <i>Taal</i> en <i>Interacties</i> anderzijds.4. Significante positieve correlatie tussen GO-interactieschaal <i>Ontwikkelingsstimulering</i> enerzijds en de FCCERS-R <i>Totaalscore</i>, en de FCCERS-R subschalen <i>Taal</i> en <i>Interactie</i> anderzijds.

Verwachtingen m.b.t. validiteit FCCERS-R

Validatie-instrument	Verwachtingen
<i>Convergente validiteit</i>	
▪ HOME	5. Significante positieve correlatie tussen FCCERS-R <i>Totaalscore</i> en de HOME <i>Totaalscore</i> .
	6. Significante positieve correlatie tussen de FCCERS-R subschaal <i>Interacties</i> en samengestelde subschaal <i>Interacties</i> van de HOME.
▪ GO-IV	7. Significante positieve correlatie tussen FCCERS-R <i>Totaalscore</i> en GO-interactieschaal <i>Totaalscore</i> .
	8. Significante positieve correlatie tussen FCCERS-R subschaal <i>Interacties</i> en GO-Interactieschaal <i>Totaalscore</i> .
	9. Significante en positieve correlatie tussen de FCCERS-R subschaal <i>Taal</i> en GO-interactieschaal <i>Praten en Uitleggen</i> .

5. RESULTATEN

5.1 GO-interactieschalen

Beschrijvende gegevens

In Tabel 5.1 is een overzicht gegeven van de gemiddelde scores voor de interactieschalen op gastouderniveau. De zes interactieschalen laten alle een patroon zien met spreiding zien op de gehanteerde 7-puntsschaal en zowel lage als hoge scores komen voor. De totaalscore benadert een normale verdeling; skewness- en kurtosiswaarden en normaliteitstests wijzen niet op een sterk afwijkende verdeling.

Tabel 5.1

Beschrijvende data GO-interactieschalen voor gastouders (min-max: 1-7; N = 40)

	<i>M</i>	<i>SD</i>	Range	Skewness	Kurtosis
Sensitieve responsiviteit	5.20	0.75	3.33 – 6.50	-0.67	0.06
Respect voor autonomie	4.92	0.81	2.75 – 6.25	-0.74	0.84
Structureren en grenzen stellen	5.02	0.82	3.33 – 6.75	-0.12	-0.17
Praten en uitleggen	4.52	0.82	2.33 – 6.00	-0.15	0.10
Ontwikkelingsstimulering	3.58	1.07	1.67 – 5.75	0.31	-0.96
Begeleiden van interacties	3.12	1.00	1.00 – 5.00	-0.11	-0.51
GO-totaalscore	4.40	0.69	2.67 – 5.75	-0.38	-0.13

Betrouwbaarheid

Inter-observator betrouwbaarheid

De getrainde observatoren scoorden elk binnen de marge van 1 punt van de normscore en bij minder dan 5 procent een afwijking van 2 punten voor de zes schalen: *sensitieve responsiviteit* (respectievelijk 0 of 1 punt afwijking: 98%, 2 punten afwijking: 2%), *respect voor autonomie* (98%, 2%), *structureren en grenzen stellen* (96, 4%), *praten en uitleggen* (96%, 4%), *ontwikkelingsstimulering* (98%, 2%) en *begeleiden van interacties* (96%, 4%).

Interne consistentie

De interne consistentie geeft aan in hoeverre metingen die een gedeeld algemeen kenmerk beogen te meten (bijvoorbeeld: de interactievaardigheden van gastouders) dat coherent doen, hier uitgedrukt als Cronbach's α . De consistentie is acceptabel als waarden $\geq .60$ zijn. De waarden voor Cronbach's α zijn weergegeven in Tabel 5.2.

De zes schalen voor het meten van de interactievaardigheden laten allereerst onderlinge samenhang zien voor de scores voor de vier (lunch, spel, transitie, verschonen) *verschillende gefilmde situaties* (tweede kolom in Tabel 5.2) met een gemiddelde waarde van .72, variërend van .56 tot .80. Dit laat zien dat de gefilmde situaties gezamenlijk een betrouwbaar beeld geven van de afzonderlijke interactievaardigheden; alleen Begeleiden van interacties blijft hier iets bij achter. Overigens geldt voor alle schalen (met uitzondering van Begeleiden van interacties) dat ook bij drie gefilmde situaties al sprake is van een betrouwbaar beeld (de waarden van Cronbach's α liggen rond de .70 bij drie gefilmde situaties).

De gemiddelde score van de zes interactievaardigheden, berekend over alle vier de situaties (derde kolom in Tabel 5.2), hangen onderling samen, zoals verwacht. Cronbach's α voor de zes schalen was .87, wat aangeeft dat de zes schalen gezamenlijk een zeer betrouwbare indicatie geven van het begrip 'interactievaardigheden' van gastouders.

Tabel 5.2

Interne consistentie voor de zes schalen voor het meten van de GO-interactievaardigheden

Interactievaardigheid	Situatie	Totaal (4 situaties)
	<i>Cronbach's α</i>	<i>Cronbach's α</i>
Sensitieve responsiviteit	.77	
Respect voor autonomie	.77	
Structureren en grenzen stellen	.80	
Praten en uitleggen	.70	
Ontwikkelingsstimulering	.72	
Begeleiden van interacties	.56	
Totaal	.79	.87

Noot: Grijs vlak = verwachting geformuleerd, vet gedrukt = verwachting uitgekomen.

De zes interactieschalen hangen, naar verwachting, significant positief met elkaar samen (zie Tabel 5.3). De gedeelde varianties variëren van 4 tot 69 procent, wat laat zien dat de scores variantie delen en daarnaast ook eigen unieke variantie hebben. Dit is ook te zien in de duidelijke verschillen in de gemiddelde scores voor de zes interactieschalen (zie Tabel 5.1).

Tabel 5.3
Correlaties tussen de GO-interactieschalen

	1	2	3	4	5
1. Sensitieve responsiviteit	-				
2. Respect voor autonomie	.83*	-			
3. Structureren en grenzen stellen	.68*	.52*	-		
4. Praten en uitleggen	.76*	.65*	.68*	-	
5. Ontwikkelingsstimulering	.58*	.69*	.50*	.73*	-
6. Begeleiden van interacties	.46*	.25	.47*	.41*	.20

* $p < .01$

Test-hertest betrouwbaarheid

De stabiliteit of test-hertest betrouwbaarheid is bepaald voor een relatief kleine groep gastouders. Voor het bepalen van de test-hertest betrouwbaarheid (of stabiliteit) van de interactieschalen waren de scores beschikbaar van 8 gastouders, die op twee verschillende dagen werden gefilmd met gemiddeld 2-4 weken tussen de twee observatiemomenten ($N = 8$). De analyses zijn vanwege de beperkte steekproefgrootte alleen descriptief en exploratief. Het absolute verschil tussen de totaalscores van het eerste en het tweede bezoek was zo'n halve punt, berekend over alle zes interactievaardigheden. Per vaardigheid varieerde het gemiddelde verschil tussen .03 en .84 punt.

Tabel 5.4
Gemiddelde scores en standaarddeviaties voor de GO-interactieschalen op de twee meetmomenten ($N = 8$)

Interactievaardigheid	Test	Hertest	Verskil
	<i>M</i>	<i>M</i>	
Sensitieve responsiviteit	5.38	5.34	.03
Respect voor autonomie	5.15	4.91	.24
Structureren en grenzen stellen	5.13	4.54	.58
Praten en uitleggen	5.04	4.49	.55
Ontwikkelingsstimulering	4.10	3.40	.71
Begeleiden van interacties	3.55	2.71	.84
Totaal	4.72	4.23	.49

Validiteit

De convergente validiteit van de GO-interactieschalen is getoetst aan de hand van correlaties met (de totaalscore en/of betrouwbare subschalen van) de FCCERS-R. De validiteitscoëfficiënten staan hieronder in Tabel 5.5 weergegeven.

Tabel 5.5
Resultaten validatie GO-interactieschalen

Validatie-instrumenten	IV- Totaal	SR	RA	SG	PU	OS	BI
<i>Convergente validiteit</i>							
▪ FCCERS-R Totaal	.27*	.27*	.19	.13	.17	.31*	.19
▪ FCCERS-R Taal	.15	.15	.11	.06	.23	.37*	.06
▪ FCCERS-R Interacties	.28*	.21	.26	.11	.21	.34*	.17

Noot: Grijs vlak = verwachting geformuleerd, vet gedrukt = verwachting uitgekomen.

* $p < .05$. IV-totaal= Interactievaardigheden Totaal, SR= Sensitieve Responsiviteit; RA= Respect voor Autonomie, SG= structureren en grenzen stellen, PU= Praten en Uitleggen, OS= Ontwikkelingsstimulering, BI= Begeleiden van Interacties.

De totaalscore op de GO-interactievaardigheden hangt volgens verwachting (zie hypothese 1 in Tabel 4.1) statistisch significant samen met de FCCERS-R totaalscore. De correlatie is bescheiden maar significant ($r = .27$), wat laat zien dat beide instrumenten weliswaar deels overlap laten zien, maar ook unieke variantie hebben. De FCCERS-R en de GO-interactieschalen vullen elkaar, zo gezien, aan. De totaalscore op de GO-interactievaardigheden hangt tevens volgens verwachting (zie hypothese 2 in Tabel 4.1) statistisch significant samen met de FCCERS-R subschaal Interacties ($r = .28$).

Tegen de verwachting in (hypothese 3 in Tabel 4.1), zijn de correlaties tussen de GO-interactieschaal Praten en Uitleggen enerzijds en de FCCERS-R subschaal Taal en de Interacties anderzijds, zwak en statistisch niet significant; absoluut gezien is het verschil tussen de validiteitscoëfficiënten in Tabel 5.5 overigens klein.

Tot slot hangt de GO-interactieschaal Ontwikkelingsstimulering volgens verwachting (hypothese 4 in Tabel 4.1) positief en significant samen met de FCCERS-R totaalscore ($r = .34$) en de subschalen Taal ($r = .37$) en Interacties ($r = .34$).

5.2 FCCERS-R

Beschrijvende gegevens

De scores op het FCCERS-instrument liggen verspreid over de 7-puntsschaal en zowel lage als hoge scores komen voor. De totaalscore voor de algemene pedagogische kwaliteit is bij benadering normaal verdeeld; de skewness en kurtosis-waarden en normaliteitstests zijn statistisch niet significant en geven geen indicatie van een afwijkende normaalverdeling. Op subschaalniveau zijn de scores niet allemaal normaal verdeeld (dit geldt met name voor Ruimte en Interacties). In Tabel 5.6 zijn de beschrijvende gegevens opgenomen voor de FCCERS-R.

Tabel 5.6
Beschrijvende data FCCERS-R (min-max: 1-7; N = 40)

	<i>M</i>	<i>SD</i>	<i>Min-max</i>	<i>Skewness</i>	<i>Kurtosis</i>
Ruimte	3.35	0.82	1.83 – 6.33	1.45	3.82
Taal	4.83	0.97	2.67 – 6.67	-0.40	.004
Activiteiten	3.19	0.74	1.70 – 4.90	0.78	0.52
Interacties	5.71	1.22	1.75 – 7.00	-1.54	3.07
Programma	5.85	0.84	4.00 – 7.00	-0.15	-0.97
Totaal	4.15	0.60	2.31 – 5.59	-0.26	1.33

Betrouwbaarheid

Inter-observator betrouwbaarheid

Een observator werd beschouwd als betrouwbaar als tijdens ieder van drie opeenvolgende bezoeken aan gastouders 80% van de scores (binnen 1 schaalpunt) overeen kwamen met de score van de trainer. Observatoren bezochten tijdens de training gemiddeld vier gastouders (range 4 – 5). Het gemiddelde betrouwbaarheidspercentage over de drie opeenvolgende bezoeken (>80%) was 82% (range 80% tot 93%).

Interne consistentie

De betrouwbaarheid van de totale FCCERS-R is adequaat met Cronbach's $\alpha = .73$ (met in totaal 28 items voor de subschalen Ruimte, Taal, Activiteiten, Interacties, Programma. Deze waarde ligt iets onder de waarde die in Amerikaans onderzoek met de FCCERS-R werd gevonden (.90) (Harms, Cryer, & Clifford, 2007), die dan geldt voor de volledige toets met 38 items (in plaats van 28). De interne consistentie van de subschalen is overigens wisselend. De consistentie is acceptabel

(gedefinieerd als Cronbach's α -waarden $\geq .60$) voor de subschalen Taal, Activiteiten en Interacties. Het item voor Boeken is niet meegenomen bij het berekenen van de betrouwbaarheid van Taal, omdat deze inhoudelijk en statistisch niet goed past bij de andere items van de subschaal. Het boeken-item, dat specifiek over lectuur gaat, sluit niet aan op de twee overige items die beide betrekking hebben op mondelinge communicatie en empirisch blijken de kwaliteit van de mondelinge communicatie en het gebruik maken van boeken twee aparte zaken. Het tv-item is niet meegenomen bij het berekenen van de subschaal Activiteiten, omdat deze relatief vaak niet van toepassing was en dan niet was ingevuld door de observatoren. De betrouwbaarheid is te laag voor de subschalen Ruimte en Programma om als aparte onderdelen van het meetinstrument te fungeren.

De betrouwbare subschalen Taal, Activiteiten en Interacties van de FCCERS-R hangen onderling positief samen: Taal en Activiteiten $r = .33, p < .01$; Taal en Interacties $r = .46, p < .01$; Activiteiten en Interacties $r = .23, n.s.$; Activiteiten en Taal $r = .33, p < .01$.

Test-hertest betrouwbaarheid

Voor het bepalen van de test-hertest betrouwbaarheid (of stabiliteit) van de FCCERS-R waren de scores beschikbaar van 8 gastouders, die op twee verschillende dagen werden bezocht met gemiddeld 2-4 weken tussen de twee observatiemomenten. De analyses zijn, zoals eerder ook al aangegeven, vanwege de beperkte steekproef grootte alleen descriptief en exploratief. Het absolute verschil tussen het eerste en het tweede bezoek bedroeg gemiddeld 0,14 punt voor de totaalscore op het FCCERS-instrument. Voor de subschalen varieerde dit verschil van .05 tot .60 (zie Tabel 5.7).

Tabel 5.7

Gemiddelde scores en standaarddeviaties voor de FCCERS-R subschalen en totaal op de twee meetmomenten (N = 8)

	Test	Hertest	Vershil
	<i>M</i>	<i>M</i>	
Ruimte	3.60	3.23	0.37
Taal	5.20	5.25	0.05
Activiteiten	2.81	3.16	0.35
Interacties	5.56	6.16	0.60
Programma	5.98	6.12	0.14
Totaal	4.11	4.25	0.14

Validiteit

In Tabel 5.9 is te zien dat de FCCERS-R totaalscore, zoals verwacht (hypothese 5 in Tabel 4.1), samenhangt met de HOME totaalscore ($r = .29$). Ook de verwachte samenhang tussen de FCCERS-R subschaal Interacties en de samengestelde subschaal Interacties van de HOME, is significant en positief ($r = .37$) (hypothese 6 in Tabel 4.1). Daarnaast hangt de FCCERS-R totaalscore met de totaalscore GO-interactieschalen positief en significant samen ($r = .27$), zoals verwacht (hypothese 7 uit Tabel 4.1). De FCCERS-R subschaal Interacties hangt, naar verwachting (hypothese 8 in Tabel 4.1), positief en significant samen met de totaalscore GO-interactieschalen ($r = .28$). Ten slotte blijkt de FCCERS-R subschaal Taal, tegen de verwachting in (hypothese 9 in Tabel 4.2), niet significant samen te hangen met de GO-schaal Praten en Uitleggen.

Tabel 5.9
Resultaten validatie FCCERS-R

Validatie-instrumenten	FCCERS-R Totaal	FCCERS-R Interacties	FCCERS-R Taal
<i>Convergente validiteit</i>			
▪ HOME (totaal)	.29*	.39*	.21
▪ HOME Interacties	.17	.37*	.14
▪ GO-IV Totaal	.28*	.29*	.25
▪ GO-IV Praten en uitleggen	.17	.21	.23

Noot: Grijs vlak = verwachting geformuleerd, vet gedrukt = verwachting uitgekomen.

* $p < .05$

5.3 Ontwikkelingsgericht werken voor 2-4 jarigen

Binnen het GO-instrument zijn met name de GO-interactieschalen Praten en uitleggen, Ontwikkelingsstimulering en Begeleiden van interacties gericht op de stimulering van de ontwikkeling, kijkend naar de talige, cognitieve en sociale ontwikkeling van kinderen. Voor deze drie 'educatieve' interactieschalen geldt dat zij deels met elkaar samenhangen maar ook eigen 'unieke' variantie laten zien en dus in empirisch opzicht complementair zijn. De FCCERS-R dekt binnen het

domein van ontwikkelingsgericht werken de kwaliteit van de taalstimulering, de ontwikkelingsgerichte activiteiten die aansluiten op het niveau van het kind, het programma en ten slotte de omgang tussen gastouder en de kinderen. Ook voor de FCCERS-R geldt dat de verbanden tussen de ontwikkelingsgerelateerde subschalen van dit instrument bescheiden zijn en zij vullen elkaar dus aan. De voorspelde en gevonden samenhang tussen de FCCERS-R totaalscore enerzijds en de GO-interactieschaal Ontwikkelingsstimulering anderzijds onderstreept het accent van de FCCERS-R op een ontwikkelingsgerichte en op het kind aansluitende aanpak in de gastouderopvang. Ook de gevonden samenhang van alle betrouwbaar gebleken subschalen van de FCCERS-R met de GO-interactieschaal Ontwikkelingsstimulering laat zien dat beide delen van het instrument de kwaliteit van ontwikkelingsgericht werken – op verschillende maar aansluitende manieren – in kaart brengen. Ook zien we op beide instrumenten duidelijk spreiding in de scores voor het ontwikkelingsgericht werken, wat laat zien dat de instrumenten op dit vlak gevoelig zijn voor verschillen in de geobserveerde praktijk; een brede peiling onder een grote steekproef van gastouders zal meer licht werpen op de variatie in de praktijk op dit punt. Kort samengevat, de GO-interactieschalen en de FCCERS-R bieden samen betrouwbaar en valide inzicht in de kwaliteit van ontwikkelingsgericht werken voor 2-4 jarigen met aandacht voor materialen, het programma en de gastouder-kind-interactie.

5.4 Samenvatting van de resultaten

In Tabel 5.10 worden de hierboven beschreven resultaten van de valideringstudie overzichtelijk weergegeven. Het meetinstrument Gastouderopvang dat bestaat uit twee onderdelen: (1) de GO-interactieschalen en (2) de FCCERS-R (aangepast aan de Nederlandse gastouder context). In de Tabel 5.10 is met plussen en minnen aangegeven of de resultaten van het onderzoek voor beide onderdelen van het meetinstrument Gastouderopvang een eerste aanzet zijn voor positieve dan wel negatieve evidentie voor de betrouwbaarheid en de validiteit. De resultaten voor de zes GO-interactieschalen en de FCCERS-R zijn beide samengevat in één kolom.

Tabel 5.10
Overzicht resultaten valideringsonderzoek

	GO-interactieschalen	FCCERS-R
<i>Betrouwbaarheid</i>		
Inter-observator betrouwbaarheid	+	+
Interne consistentie	+	+
<i>Validiteit</i>		
Convergente validiteit	+	+

+ = Resultaten ondersteunen betrouwbaarheid/validiteit van het instrument

- = Resultaten ondersteunen betrouwbaarheid/validiteit van het instrument niet

6 CONCLUSIE EN DISCUSSIE

Dit rapport beschrijft de achtergrond en de ontwikkeling van het meetinstrument Gastouderopvang, dat bestaat uit twee onderdelen: de GO-interactieschalen en de FCCERS-R (aangepast aan de Nederlandse gastouder context). Daarnaast worden de resultaten van een eerste aanzet in het toetsen van de betrouwbaarheid en de validiteit van beide onderdelen van het meetinstrument weergegeven. De resultaten van deze validatiestudie lieten zien dat de GO-interactieschalen en FCCERS-R, de twee componenten van het instrument voor de gastouderopvang, elkaar aanvullen in empirisch onderzoek. Hieronder worden de belangrijkste bevindingen besproken.

Betrouwbaarheid

Uit de resultaten blijkt dat de *interne consistentie* voor vijf van de zes interactievaardigheden over de vier verschillende situaties goed is. Alleen voor de schaal Begeleiden van interacties blijkt de waarde niet voldoende (Cronbach's $\alpha < .60$). Een mogelijke verklaring voor deze bevinding is dat scores op deze schaal tijdens het verschoonmoment afwijken, aangezien er in deze context vooral veel individuele aandacht is voor het kind dat verschoond wordt en niet voor de begeleiding van sociale contacten. De scores voor het verschoonmoment drukken op die manier de betrouwbaarheid op de schaal Begeleiden van interacties voor het totaal van de vier situaties.

Voor de *test-hertest* betrouwbaarheid geldt dat de analyses vanwege de beperkte steekproefgrootte alleen descriptief en exploratief zijn. Om die reden is nader onderzoek naar de stabiliteit over tijd van zowel de GO-interactieschalen als de FCCERS-R gewenst. Een eerste landelijk kwaliteitspeiling gastouderopvang zou hiervoor een goede mogelijkheid bieden.

De FCCERS-R subschalen Ruimte en Programma lieten een lage interne consistentie zien. De FCCERS-R subschaal Ruimte meet kennelijk in de Nederlandse context zaken die alle onder één noemer geschaard kunnen worden, maar empirisch zwak samenhangen met elkaar in de praktijk. Voor de FCCERS-R subschaal Programma kan de verklaring gezocht worden in het feit dat de schaal klein is met slechts 3 items en bovendien is er door de hoge scores een gebrek aan variantie. Deze bevindingen zijn in overeenstemming met de handleiding bij het Amerikaanse instrument waar ook wordt aangegeven voorzichtig te zijn met het rapporteren van de afzonderlijke subschalen (Harms et al., 2007). De ontwikkelaars van de FCCERS-R adviseren om die reden om de totaalscore te rapporteren en terughoudend te zijn in wetenschappelijk onderzoek met de individuele subschalen.

Validiteit

Tegen de verwachting in is er geen verband gevonden tussen de GO-interactieschaal Praten en Uitleggen enerzijds en de FCCERS-R subschaal Taal en de Interacties anderzijds. Een mogelijke inhoudelijke verklaring is dat de FCCERS-R subschaal Taal expliciet gericht is op uitleggen bij het begrijpen en gebruiken van nieuwe taal, waar de GO-interactieschaal Praten en Uitleggen gericht is op kwantiteit en kwaliteit van het actieve taalgebruik van de gastouder in het algemeen (dus ook zonder de specifiekere insteek om actief kinderen taal helpen te gebruiken of begrijpen).

Toekomstig onderzoek

Het verdient aanbeveling om de stabiliteit van het GO-instrument voor de gastouderopvang nader te onderzoeken in een grotere steekproef en met een ruimere tijdsinterval dan in het kader van deze studie mogelijk was, zodat er uitgebreidere gegevens kunnen worden verzameld over de test-herstest. Ook strekt het tot aanbeveling om de predictieve validiteit (wat voorspelt de pedagogische kwaliteit voor de ontwikkeling van kinderen die naar de gastouderopvang gaan?) en eventueel ook discriminante validiteit (waar hangt het instrument juist niet mee samen?) te onderzoeken van het GO-instrument na deze eerste validatie. Concreet valt te denken aan een uitbreiding in toekomstig peilingsonderzoek met kindgegevens met betrekking tot het algemene welbevinden en de ontwikkeling van kinderen. Verder is het in toekomstig (peilings)onderzoek mogelijk met geavanceerde statistische analyses, die wel een grote steekproef vereisen, de psychometrische eigenschappen van het instrument nader te onderzoeken na de eerste positieve bevindingen uit deze validatiestudie. Internationale studies hebben laten zien dat de Environment Rating Scale-meetinstrumenten met de verschillende kwaliteitsindicatoren een activiteiten/materialen-factor en een interactie/taal of staf-factor onderscheiden (zie ook Cassidy, Hestenes, Hedge, Hestenes, & Mims, 2005; Perlman, Zellman, & Le, 2004; Sakai, Whitebook, Wishard, & Howes, 2003; Tietze, Cryer, Bairrao, Palacios, & Wetzel, 1996). Op een groter databestand kan worden nagegaan of de factoriële structuur van de FCCERS-R die is gevonden in Amerikaans onderzoek, kan worden gerepliceerd voor Nederlandse gastouders (zie Schaack, Le, & Setodji, 2013).

Deze studie heeft tevens duidelijk gemaakt dat het HOME-instrument – dat in deze studie als zogenaamd ‘schaduwinstrument’ fungeerde – beduidend minder geschikt is voor peilingsonderzoek onder Nederlandse gastouders dan de FCCERS-R. Het HOME-instrument levert geen betrouwbare subschalen op (zie ook Groeneveld et al., 2011). Verder lijkt het HOME-instrument onvoldoende gevoelig voor genuanceerde kwaliteitsverschillen door de dichotome scoringswijze (wel of niet

aanwezig) bij alle kwaliteitsindicatoren. Doordat voor een groot aantal indicatoren altijd positieve scores worden verkregen discrimineert het instrument onvoldoende tussen de pedagogische kwaliteit van Nederlandse gastouders en laat het instrument onvoldoende gradaties in pedagogische kwaliteit laat zien. Door de afwijkende structuur, en scoringsformat van de HOME ten opzichte van de meetinstrumenten die gebruikt worden in peilingsonderzoek voor de kinderdagopvang en de buitenschoolse opvang (die op een 7-puntschaal meten), kunnen resultaten die eventueel verkregen worden met de HOME slecht worden vergeleken de resultaten uit de peilingsonderzoeken uit de kinderopvang en BSO.

Tot slot

Met het ontwikkelde instrument voor de pedagogische kwaliteit van de gastouderopvang is er nu een instrument beschikbaar dat aansluit bij het al beschikbare instrumentarium voor de kinderdagopvang en de buitenschoolse opvang, en dat is gevalideerd in de Nederlandse situatie. Het instrument meet zowel de kwaliteit van de leefomgeving als de basale en educatieve interactievaardigheden van de gastouder(opvang), en daarnaast worden ook structurele kenmerken meegenomen. Het GO-instrument maakt een vergelijking mogelijk met diverse opvangtypen in het Nederlandse peilingsonderzoek en een vergelijking met buitenlandse peilingsresultaten. De resultaten van de GO-interactieschalen maken bij een toekomstige peiling een vergelijking mogelijk tussen de dagopvang, peuterspeelzalen (met of zonder voor- en vroegschoolse educatie) en buitenschoolse opvang. Ook voor de interactievaardigheden is binnenkort een buitenlandse vergelijking mogelijk, namelijk met de Noorse kinderopvang (Johansson, 2012).

Verder is het mogelijk om op termijn – en met verwerking van de gegevens van een grote landelijke peiling – een kwaliteitsmonitor te ontwikkelen voor de gastouders om deze opvangsoort een praktisch instrument in handen te geven zoals dat al beschikbaar is voor de kinderdagopvang en buitenschoolse opvang (Gevers Deynoot-Schaub, Fukkink, Riksen-Walraven, de Kruif, Helmerhorst & Tavecchio, 2009; Boogaard, van Daalen-Kapteijns, & Gevers Deynoot- Schaub, 2012).

REFERENTIES

- Boogaard, M., Hoex, J., Van Daalen-Kapteijns, M., & Gevers Deynoot-Schaub, M. (2013). *Pedagogisch kader gastouderopvang*. Amsterdam: Reed Business Education.
- Boogaard, M., & Van Daalen-Kapteijns, M. (2012). *Pedagogische kwaliteit van buitenschoolse opvang in Nederland*. Amsterdam: Kohnstamm Instituut.
- Boogaard, M., & Fukkink, R.G. (2009). *Pedagogische kwaliteiten van de buitenschoolse opvang: ontwikkeling van een meetinstrument*. Amsterdam: SCO-Kohnstamm Instituut & NJi.
- Boogaard, M., Van Daalen-Kapteijns, M., & Gevers Deynoot-Schaub, M.J.J.M. (2012). *De Kwaliteitsmonitor BSO: een instrument waarmee de buitenschoolse opvang de eigen pedagogische kwaliteit in kaart kan brengen*. Amsterdam: SWP.
- Brancheorganisatie Kinderopvang, september 2013 (factsheet kinderopvang)
- Caldwall, B. M. & Bradley, R. H. (2003). *Home observation for measurement of the environment: Administration manual*. Little Rock: University of Arkansas Press,
- Cassidy, D.J., Hestenes, L.L., Hegde, A., Hestenes, S., Mims, S., (2005). Measurement of quality in preschool child care classrooms: An exploratory and confirmatory factor analysis of the early childhood environment rating scale-revised. *Early Childhood Research Quarterly*, 30(3), 345-360
- De Kruif, R.E.L., Vermeer, H.J., Fukkink, R.G., Riksen-Walraven, J.M.A., Tavecchio, L.W.C., van IJzendoorn, & van Zeijl, J. (2007). *De Nationale Studie Pedagogische Kwaliteit Kinderopvang Eindrapport Project 0 en 1*. Amsterdam: NCKO.
- De Kruif, R.E.L., Riksen-Walraven, J.M.A., Gevers Deynoot-Schaub, M.J.J.M., Helmerhorst, K.O.W., Tavecchio, L.W.C., & Fukkink, R.G. (2009). *Pedagogische kwaliteit van de opvang voor 0- tot 4-jarigen in Nederlandse kinderdagverblijven in 2008*. Amsterdam/Nijmegen: NCKO.
- Elicker, J. Ruprecht, K. M., Langill, C., Lewsader, J., Anderson, T., & Brizzi, M. (2013). Indiana Paths to QUALITY: Collaborative evaluation of a new child care quality rating and improvement system. *Early Education and Development*, 24, 42-62.

- Fukkink, R.G., Tavecchio, L., Kruif, R. de, Vermeer, H., & Zeijl, J. van (2005). Criteria voor kwaliteit van kinderopvang: Visies van sleutelfiguren. *Pedagogiek*, 25(4), 243-261.
- Fukkink, R., Gevers Deynoot-Schaub, M.J.J.M., Helmerhorst, K.O.W., Bollen, I., & Riksen-Walraven, J.M.A. (2013). *Pedagogische kwaliteit van de opvang voor 0- tot 4-jarigen in Nederlandse kinderdagverblijven in 2012*. Amsterdam/Nijmegen: NCKO.
- Fukkink, R.G., Helmerhorst, K., Gevers Deynoot-Schaub, M., De Kruif, R.E.L., Tavecchio, L.W.C., & Riksen-Walraven, M. (2011). *Pedagogische kwaliteit van de kinderopvang en de ontwikkeling van jonge kinderen: Een longitudinale studie*. Amsterdam: NCKO.
- Gevers Deynoot-Schaub, M.J.J.M., Fukkink, R.G.W., Riksen-Walraven, M., De Kruif, R., Helmerhorst, K.O.W., Tavecchio, L.W.C. (2009). *De NCKO-Kwaliteitsmonitor: Het instrument waarmee kinderdagverblijven zelf hun pedagogische kwaliteit in kaart kunnen*
- Groeneveld, M., Vermeer, H., & IJzendoorn, M. van (2010). Children's wellbeing and cortisol levels in home-based and center-based child care. *Early Childhood Research Quarterly*, 25, 502-514.
- Groeneveld, M. G., Vermeer, H. J., Van IJzendoorn, M. H., & Linting, M. (2011) Enhancing home-based child care quality through video-feedback intervention: A randomized controlled trial. *Journal of Family Psychology*, 25, 86-96.
- Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early Childhood Environment Rating Scale-Revised*. New York, NY: Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R. M. (2003). *Infant/Toddler Environment Rating Scale-Revised*. New York, NY: Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R. M. (2007). *Family Child Care Environment Rating Scale-Revised*. New York, NY: Teachers College Press.
- Helmerhorst, K. O. W., Riksen-Walraven, J. M., Vermeer, H. J., Fukkink, R. G., & Tavecchio, L. W. C. (2014). Measuring interactive skills of caregivers in child care centers: Development and validation of the caregiver interaction profile scales. *Early Education & Development*, 25, 770-790.

- Helmerhorst, K.O.W., Riksen-Walraven, J. M., Gevers Deynoot-Schaub, M.J.J.M., Tavecchio, L.W.C., & Fukkink, R.G. (2014). Child care quality in the Netherlands over the years: A closer look. *Early Education and Development*. doi: 10.1080/10409289.2014.948784
- Johansson, J. (2012). *Better provision for Norway's children in ECEC. A study of children's well-being and development in ECEC, and new tools for quality evaluation. Application to The Research Council of Norway*. Oslo, Norway: Oslo and Akerhus University College of Applied Sciences.
- NCKO /Kohnstamm Instituut (2014). *Gastouder-interactievaardigheden*. Amsterdam: NCKO/Kohnstamm Instituut.
- Perlman, M., Zellman, G.L., & Le, V.N. (2004) Examining the psychometric properties of the early childhood environment rating scale-revised (ECERS-R). *Early Childhood Research Quarterly*, 19 (3), 398-412.
- Riksen-Walraven, M. (2004). Pedagogische kwaliteit in de kinderopvang: doelstellingen en kwaliteitscriteria. In M.H. Van IJzendoorn, L.W.C. Tavecchio, & M. Riksen-Walraven (Red.), *De kwaliteit van de Nederlandse kinderopvang* (pp.100-123). Amsterdam: Boom
- Sakai, L.M., Whitebook, M., Wishard, A., & Howes, C. (2003). Evaluating the Early Childhood Environment Rating Scale (ECERS): Assessing differences between the first and revised edition. *Early Childhood Research Quarterly*, 18 (4), 427-445.
- Schaack, D., Le, V.N., & Setodji, C.M. (2013). Examining the factor structure of the Family Child Care Environment Rating Scale-Revised. *Early Childhood Research Quarterly*, 28 (4), 936-946.
- Taubert, S., Tietze, W., Förster, C., Lee, H., & Schlecht, D. (2006). Pädagogische Qualität der Tagespflege in Brandenburg. Abschlussbericht. Berlin: Freie Universität Berlin.
- Tietze, W., Cryer, J., Bairrao, D., Palacios, J., & Wetzal, G. (1996). Comparisons of observed process quality in early child care and education programs in five countries. *Early Childhood Research Quarterly*, 11, 447-475.
- Vermeer, H., van IJzendoorn, M.H., de Kruif, R.E.L., Fukkink, R.G., Tavecchio, L.W.C., Riksen-Walraven, J.M.A., & van Zeijl, J. (2005). *Kwaliteit van Nederlandse kinderdagverblijven: Trends in kwaliteit in de jaren 1995-2005*. Leiden/Amsterdam/Nijmegen: NCKO.

Vermeer, H., Van IJzendoorn, M.H., De Kruif, R., Fukkink, R.G., Tavecchio, L.W.C., Riksen-Walraven, M.M.A., & van Zeijl, J. (2008). Child care quality in the Netherlands: Trends in quality over the years 1995-2005. *Journal of Genetic Psychology, 169*(4), 360-385.

