



**Landelijke
Kwaliteitsmonitor
Kinderopvang**

Kwaliteit van de babyopvang in Nederland

Gecombineerde metingen 2017-2019



Pauline Slot

IJsbrand Jepma

Paulien Muller

Bodine Romijn

Celeste Bekkering

Paul Leseman

Juli 2020



Universiteit Utrecht

Sardes



Kwaliteit van de babyopvang in Nederland

Verklarende analyses op basis van de
gecombineerde gegevens metingen 2017-2019

Voorwoord

Hoewel LKK2020 volledig was opgestart, hebben we vanwege de beperkingen rondom corona moeten besluiten om de meting te staken. In nauw overleg met het Ministerie van Sociale Zaken en Werkgelegenheid is ervoor gekozen om een aantal verdiepende analyses te doen op de gegevens die zijn verzameld tussen 2017 en 2019. Dit is het eerste rapport in een reeks van drie en gaat dieper in op de kwaliteit van de babyopvang. We willen Reina Kloosterman en Frank Schalekamp van het Ministerie van Sociale Zaken en Werkgelegenheid bedanken voor de prettige en constructieve samenwerking, ook bij het zoeken naar een alternatieve invulling van de opdracht.

Pauline Slot, Celeste Bekkering, Bodine Romijn & Paul Leseman (Universiteit Utrecht),
IJsbrand Jepma & Paulien Muller (Sardes)

Inhoud

Voorwoord	4
Samenvatting	7
Leeswijzer	9
1. Inleiding	10
1.1 Landelijke Kwaliteitsmonitor Kinderopvang (LKK)	10
1.2 Kwaliteit van de opvang voor baby's en beleidsontwikkelingen	10
2. Kwaliteit van de baby opvang	12
2.1. Introductie en leeswijzer	12
2.2. Samenvatting en beschouwing van de resultaten	12
2.3. Methodische aanpak en verantwoording	15
2.4. Proceskwaliteit.....	15
2.4.1. CLASS Infant	15
2.4.2 NCKO-interactievaardigheden	18
2.4.3. Omgevingschaal ITERS-R	19
2.4.4 Activiteitsaanbod.....	21
2.4.5. Samenvattend beeld van de proceskwaliteit.....	23
2.5 Structurele kwaliteit	24
2.5.1 Kenmerken van de medewerkers en de groep	24
2.5.2 Pedagogisch beleid en pedagogisch klimaat.....	27
2.5.4 Samenvattend beeld van de interviewgegevens	29
3. Verklarende analyses babyopvang	30
3.1 Introductie en leeswijzer	30
3.2 Samenvatting en beschouwing van de resultaten	31
3.3 Voorspellers van de kwaliteit in de baby-opvang	34
Referenties.....	39
Colofon.....	41

Samenvatting

De Landelijke Kwaliteitsmonitor Kinderopvang (LKK) **presenteert in dit rapport de resultaten van de gecombineerde metingen van 2017, 2018 en 2019 specifiek voor de opvang van baby's.** Er zijn observaties verricht en interviews uitgevoerd in **62 groepen waar baby's worden opgevangen, waaronder 26 horizontale babygroepen en 36 verticale groepen met baby's.** Voorliggende studie is een verdieping op de kwaliteit van de babyopvang in Nederland. Uit eerdere metingen kwam naar voren dat de kwaliteit voor baby's lager is dan voor peuters. Om die reden wordt er nader ingezoomd op de kwaliteit voor de babyopvang en wordt gekeken welke factoren bijdragen aan een hogere kwaliteit.

Uit de gecombineerde metingen van 2017, 2018 en 2019 blijkt dat **zowel de emotionele als de educatieve kwaliteit is hoger in horizontale babygroepen.** Ook de inrichting van de ruimte, de kwaliteit van het programma en het taalgebruik met kinderen blijken vaker voldoende tot goed in horizontale groepen vergeleken met verticale groepen. Dit ligt anders voor de kwaliteit van het materialen- en activiteiten aanbod. In het algemeen scoort de kwaliteit van de materialen en het activiteiten aanbod lager dan de andere kwaliteitsaspecten, maar naar verhouding is dit aanbod beter in verticale groepen. De inrichting van de groep, zoals het hebben van voldoende, veilig en passend meubilair voor kinderen en een goede indeling van de ruimte voor actief en rustig spel (al dan niet met zachte materialen) blijkt voor baby's dus beter in horizontale groepen, maar de variatie, kwaliteit en toegankelijkheid van spelmateriaal en het activiteiten aanbod blijkt in verticale groepen. Een grotere variatie in leeftijd van de kinderen gaat dus samen met meer gevarieerd aanbod. **Over het algemeen, is de gemiddelde emotionele proceskwaliteit voldoende tot goed en de gemiddelde educatieve kwaliteit is duidelijk lager net als in de andere opvangsoorten voor peuters of schoolgaande kinderen.** Hoewel de educatieve proceskwaliteit gemiddeld op of iets boven de grens ligt van wat volgens de standaards van het meetinstrumentarium als voldoende kan gelden, is er sprake van grotere variatie dan bij de emotionele proceskwaliteit.

Daarnaast is er op basis van de gegevens van de medewerkers gekeken naar het aanbod van exploratie, spel en taal waarbij onderscheid is gemaakt naar aanbod voor baby's jonger dan een jaar en dreumesen tot twee jaar. Naarmate er **meer baby's in de groep zijn, blijken de medewerkers minder mogelijkheden te bieden voor exploratie en spel voor oudere kinderen.** De ondersteuning van exploratie en spel blijken voor baby's hoger dan voor dreumesen in zowel horizontale als verticale groepen. Hoewel voor dreumesen de mogelijkheden voor exploratie beter zijn in verticale groepen. De ondersteuning van de Nederlandse taalontwikkeling voor baby's laat niet veel verschillen zien tussen horizontale en verticale groepen, maar voor dreumesen is er meer aanbod in verticale groepen. Mogelijkheden voor exploratie en spel werden dagelijks aangeboden terwijl de aandacht voor taal meerdere malen per week voor kwam.

Er is gesproken met pedagogisch medewerkers en hieruit blijken de grootste verschillen tussen horizontale en verticale groepen de groepssamenstelling te betreffen. Zo hebben verticale groepen een groter aandeel kinderen van laagopgeleide ouders, met een migratieachtergrond en met een zorgbehoefte. **De werkbeleving van pedagogisch medewerkers is overwegend positief.** Zij ervaren gemiddeld een prettig collegiaal werkklimaat en hebben het gevoel betrokken te worden bij beslissingen binnen de organisatie. De ervaren werkstress is over het algemeen laag. **Het aanbod aan professionaliseringsactiviteiten is gemiddeld matig,** met matig-frequent pedagogisch-inhoudelijk overleg met collega's en leidinggevenden. Medewerkers van verticale groepen blijken vaker een vast contract te hebben dan medewerkers van horizontale groepen. De overige structurele kwaliteitskenmerken, zoals opleidingsniveau van de medewerkers, groepsgrootte en kind-staf ratio, zijn in overeenstemming met de wettelijke regels. Babygroepen zijn gemiddeld iets kleiner (11 kinderen vergeleken met 12 in verticale groepen) en gemiddeld zijn er 4 kinderen per medewerker vergeleken met 5 kinderen voor verticale groepen.

Uit verklarende analyses blijkt dat een aantal factoren relevant zijn voor de kwaliteit van de babyopvang. Een **hogere vooropleiding** van pedagogisch medewerkers blijkt samen te gaan met een **hogere emotionele en educatieve proceskwaliteit.** Wanneer er **meer baby's in een groep** zitten, blijkt de **emotionele en educatieve proceskwaliteit eveneens hoger.** Het hebben van **meer werkervaring** blijkt van belang voor de **globale meting van kwaliteit,** waaronder de kwaliteit van de ruimte en meubilering, de aanwezige materialen en het activiteitsaanbod, brede interactiekwaliteit en aandacht voor gesprekken en taal in het bijzonder en een programma. Daarnaast is het zo dat het aantal werkuren per week van belang. Naarmate medewerkers **meer uren** werkten per week, is de **emotionele en educatieve proceskwaliteit** hoger, evenals het door de medewerker gerapporteerde **taalaanbod.** Mogelijk hebben medewerkers die meer uren werken, de kinderen beter leren kennen en hebben ze meer ervaring opgedaan in het werken met jonge kinderen. Daardoor kunnen ze misschien beter aansluiten bij de behoeften van kinderen. Ook **continue professionalisering,** zoals cursussen, coaching, teamoverleg, conferentiebezoek en frequent pedagogisch-inhoudelijk overleg, is van belang voor de globale kwaliteit. Echter, voor de meer specifieke aspecten van de interactiekwaliteit zijn misschien andere vormen van professionalisering nodig. Met de Wet IKK moeten op termijn alle medewerkers specifieke scholing volgen gericht op het werken met baby's. Toekomstige metingen kunnen inzicht geven in de effecten van deze scholing op de kwaliteit.

Internationaal is er nog niet veel onderzoek gedaan naar de kwaliteit voor baby's. Over het algemeen scoort Nederland **net zo hoog als Australië en vergelijkbaar met de Verenigde Staten** op de ITERS-R. De kwaliteit blijkt **hoger dan** gemeten in **Noorwegen en Portugal.** Dit

geldt eveneens voor de CLASS metingen. Ook hier scoort Nederland **hoger dan Vlaanderen, Portugal en de Verenigde Staten**, vooral met betrekking tot de emotionele proceskwaliteit.

Hoewel de kwaliteit in Nederland gemiddeld hoog is, roept het wel de vraag op of verticale groepen optimaal tegemoet kunnen komen aan de unieke ontwikkelingsbehoeften van baby's. Voor baby's blijken horizontale groepen goed uit te pakken op verschillende aspecten. Zo is de kwaliteit van de interacties hoger en is de ruimte beter ingericht op de behoeften van deze jonge baby's. Echter, de kwaliteit van spelmateriaal en het activiteitsaanbod blijkt beter in verticale groepen. Dit duidt er mogelijk op dat in de inrichting van de ruimte er vooral aandacht is voor veiligheid, maar minder voor het stimuleren van exploratie en de ontwikkeling van baby's. Ook binnen horizontale en verticale groepen blijkt er een verschil te zijn tussen de ervaringen voor baby's (tot ca 1/1½ jaar) en ervaringen voor dreumesen (tussen ca 1/1½ - 2/2½ jaar), waarbij oudere kinderen er kariger vanaf komen in horizontale babygroepen of wanneer er veel baby's aanwezig zijn in de groep. Mogelijk duidt dit erop dat kinderen in deze leeftijd- en ontwikkelingsfase beter tot hun recht kunnen komen in specifieke leeftijdsgroepen (bijvoorbeeld, baby-, dreumes- en peutergroepen) waarin aandacht is voor de behoeften van het kind.

Hoewel recente beleidsontwikkelingen zoals de Wet IKK inzetten op kwaliteitsverbetering (o.a. door aanvullende eisen aan scholing van medewerkers en een kleinere ratio voor het opvangen van baby's) is het nog de vraag of dit in de praktijk het gewenste effect zal hebben. Het is voor kinderopvangorganisaties kostbaarder om baby's op te vangen in horizontale groepen. Het is denkbaar dat dit doorgerekend wordt aan ouders via een hogere uurprijs, wat ervoor kan zorgen dat deze vorm van opvang mogelijk niet voor alle ouders toegankelijk is. Een andere consequentie kan zijn dat er meer baby's worden opgevangen in verticale groepen waarvan duidelijk is dat de geboden kwaliteit lager is.

Leeswijzer

Het rapport is als volgt opgebouwd. Het eerste hoofdstuk schetst kort de achtergrond van dit rapport. Hoofdstuk 2 beschrijft de kwaliteit voor baby's aan de hand van drie observatie instrumenten en de gegevens uit de interviews met pedagogisch medewerkers. Hoofdstuk 3, tot slot, rapporteert de resultaten van de verklarende analyses.

1. Inleiding

1.1 Landelijke Kwaliteitsmonitor Kinderopvang (LKK)

Van 2017 tot en met 2019 is de Landelijke Kwaliteitsmonitor Kinderopvang (LKK) uitgevoerd door het consortium van de Universiteit Utrecht en Sardes. Dit rapport bevat de gecombineerde bevindingen van deze kwaliteitsmetingen in de babyopvang in Nederland.

Er zijn verschillende instrumenten gebruikt om de kwaliteit te meten waaronder de ITERS-R en de NCKO-schalen, vanwege vergelijkbaarheid met eerder onderzoek, evenals met de internationaal veelgebruikte CLASS. Daarnaast zijn medewerkers gevraagd naar hun achtergrond en werkbeleving en kenmerken van de groep.

1.2 Kwaliteit van de opvang voor baby's en beleidsontwikkelingen

Het aantal kinderen dat naar de kinderdagopvang gaat, is de afgelopen jaren gestegen blijkens de cijfers van het eerste kwartaal 2020¹. Recent onderzoek onder ruim 400 ouders toont aan dat het gebruik van kinderopvang varieert van 26-52% afhankelijk van het opleidingsniveau van ouders², wat redelijk vergelijkbaar is met een eerdere rapportage met gegevens uit 2014³. Redenen om geen gebruik te maken van formele opvang voor baby's hebben deels te maken met de (gepercipieerde) hoge kosten, maar nog meer met overwegingen dat kinderopvang niet goed zou zijn voor jonge kinderen. In Nederland domineert het beeld dat jonge kinderen het beste af zijn bij de ouders thuis en dat opvang in een kindercentrum (te)veel stress met zich meebrengt.

De eerste jaren van een baby zijn cruciaal voor de verdere ontwikkeling. Erkenning van het belang van deze jaren blijkt uit bijvoorbeeld het programma Kansrijke Start waarbij nationaal en lokaal aandacht wordt besteed aan de eerste 1000 dagen van een kind. Hierin staat de kwaliteit van de interactie tussen het kind en de omgeving centraal. Dit betreft allereerst de interactie met ouders of verzorgers in de thuisomgeving, maar omvat ook de kwaliteit van de geboden opvang en zorg door een pedagogische medewerker. Een hoge kwaliteit van kinderdagopvang komt ten goede aan een optimale ontwikkeling van kinderen, zoals een internationale review studie aantoont⁴.

Recente beleidsontwikkelingen, zoals de graduele implementatie van de Wet Innovatie Kwaliteit Kinderopvang (IKK) per 2018 is gericht op een verhoging van de kwaliteit, onder meer voor de babyopvang. Zo wordt er ingezet op een aanvullende scholingseis voor het werken met baby's (per 2018 met een ingroeiperiode tot 2025) en is de kind-staf ratio per 2019 aangepast van vier naar drie baby's per medewerker. Daardoor zijn de kosten van een aparte

¹ TK 1^e kwartaalrapportage kinderopvang 2020

² Kok, Plantenga, Schwartz, & ter Weel (2020). Naar een stabiel stelsel. SEO

³ Roeters & Bucx (2016). Het gebruik van kinderopvang door ouders met lagere inkomens. SCP

⁴ Melhuish et al. (2015). A review of research on the effects of early childhood education and care (ECEC) upon child development. EU-FP7 CARE

babygroep toegenomen. Hoewel exacte cijfers ontbreken over de ontwikkelingen in de verhouding tussen aparte babygroepen en verticale groepen, bestaat de indruk dat de IKK-maatregel ertoe heeft geleid dat baby's vaker worden opgenomen in een verticale groep. De vraag is of dat een gewenste ontwikkeling is. Het is nog te vroeg om deze maatregel te evalueren, maar het is zinvol om te bekijken wat de stand van zaken is in de kwaliteit van de babyopvang.

Hoewel er internationaal veel onderzoek is gedaan naar de kwaliteit in de kinderopvang, is er weinig bekend over de kwaliteit specifiek gericht op baby's⁵. Veel onderzoek omvat wel de opvang voor baby's – al dan niet in aparte babygroepen –, maar dit is moeilijk te scheiden van de opvang voor dreumesen en peuters. Eerdere resultaten uit LKK-metingen hebben echter al laten zien dat de kwaliteit voor baby's achterblijft op de kwaliteit voor peuters. Ook bleken er verschillen in kwaliteit tussen horizontale en verticale groepen. De laatste liet in 2017 en 2018 een lagere kwaliteit zien in vergelijking met horizontale groepen.

Verticale groepen zijn groepen met kinderen in de leeftijd van tien weken tot vier jaar. Aangezien de kind-staf ratio afhankelijk is van de leeftijd van de kinderen in de betreffende groep, geeft het vormgeven van verticale groepen meer flexibiliteit aan organisaties om de totale groeps grootte en ratio maximaal te benutten. Horizontale groepen, daarentegen, zijn groepen naar leeftijd van kinderen waarbij verschillende grenzen worden gehanteerd op basis van leeftijd of ontwikkelingsfase van een kind. Zo zijn er babygroepen, dreumesgroepen en peutergroepen en combinaties hiervan. Er zijn babygroepen waar een grens van 1 of 1½ jaar wordt gehanteerd of die het kunnen lopen als mijlpaal gebruiken. Dreumesgroepen zijn groepen met kinderen van 1 of 1½ tot 2 - 2 ½ jaar en peutergroepen zijn groepen van twee- en driejarige kinderen. Vaak worden baby's en dreumesen ook in één groep opgevangen waardoor een babygroep jonge kinderen van 10 weken tot 2 jaar betreft. Door deze kleinere leeftijdsrange kan mogelijk beter worden ingespeeld op de behoeften die passen bij de leeftijd van kinderen. Echter, kinderen moeten wel één of soms twee keer een overgang maken naar een nieuwe groep met nieuwe kinderen en medewerkers en zij moeten zich hier opnieuw aan hechten. Tot op heden is er amper systematisch onderzoek gedaan naar de verschillen tussen horizontale en verticale groepen.

Gezien het belang van kwaliteit voor de emotionele veiligheid en de ontwikkeling van kinderen wordt in het huidige rapport de kwaliteit voor baby's onderzocht. Hierbij wordt onderscheid gemaakt naar type groep (horizontaal of verticaal) en waar nodig ook rekening gehouden met de specifieke groepssamenstelling (aandeel baby's). In lijn met de Wet IKK, waarin de ontwikkeling van het kind centraal wordt gesteld, wordt er in het huidige rapport ook gekeken naar de ondersteuning van spel, exploratie en taal in het aanbod voor baby's.

⁵ Vermeer & Groeneveld (2017)

2. Kwaliteit van de baby opvang

2.1. Introductie en leeswijzer

Dit hoofdstuk beschrijft de kwaliteit van de babyopvang voor de eerste drie kwaliteitsmetingen van de LKK tussen 2017 en 2019 waarbij onderscheid is gemaakt tussen horizontale babygroepen en verticale groepen voor kinderen van 0 tot 4 jaar. Beide komen voor in de praktijk, maar het is niet duidelijk hoe de verdeling is tussen beide varianten in de praktijk. Om een goede vergelijking te kunnen maken tussen horizontale en verticale groepen is er in de LKK voor gekozen om een evenwichtige indeling in horizontale babygroepen, horizontale peutergroepen en verticale babypeutergroepen mee te nemen in de steekproef. Dit rapport gaat over de horizontale babygroepen en verticale groepen. Voor de verticale groepen zijn ook gegevens verzameld die betrekking hebben op peuters, maar die zijn geen onderdeel van het huidige rapport.

Het hoofdstuk is als volgt opgebouwd. Eerst volgt een samenvatting op hoofdlijnen van de belangrijkste uitkomsten en een beschouwing daarover. Voor de geïnteresseerde lezer worden vervolgens op een meer technische manier de resultaten op basis van de verschillende meetinstrumenten beschreven. Het meer technische gedeelte beschrijft de uitkomsten, maar geeft weinig interpretatie of verklaring. Dat gebeurt wel in het eerste, beschouwende deel.

2.2. Samenvatting en beschouwing van de resultaten

Er zijn verschillende instrumenten gebruikt om de proceskwaliteit voor baby's in kaart te brengen waarbij onderscheid is gemaakt tussen horizontale babygroepen en verticale groepen voor kinderen van 0-4 jaar. Zo is de CLASS Infant gebruikt om de kwaliteit van interacties met baby's te beoordelen. De emotionele kwaliteit, blijkt voldoende tot goed volgens de CLASS standaards. De educatieve kwaliteit, ofwel de mate waarin kinderen ondersteund worden in exploratiegedrag en hun taalontwikkeling, blijkt lager en valt in de middenrange. Er blijken wel grote verschillen tussen horizontale en verticale groepen, waarbij de kwaliteit hoger is in horizontale babygroepen. Deze verschillen zijn substantieel, vooral voor de educatieve kwaliteit. Er blijken eveneens kwaliteitsverschillen op basis van de geobserveerde activiteit waarbij gerichte activiteiten een aanzienlijk hogere emotionele en educatieve kwaliteit laten zien in vergelijking met eet- en drinkmomenten en vrij spel momenten. Deze gerichte activiteiten betroffen 17,5% van de geobserveerde activiteiten en kwamen dus niet heel frequent voor.

De resultaten met de NCKO-schalen laten een vergelijkbaar beeld van de emotionele en educatieve proceskwaliteit zien, maar vertonen geen verschillen tussen horizontale en verticale groepen. Een mogelijke verklaring kan liggen in het feit dat voor beide type groep alle interacties van pedagogisch medewerkers met alle kinderen in de groep (dus van alle leeftijden) zijn meegenomen. Dus wanneer er peuters in de groep waren, dan zijn deze interacties ook meegenomen in de scores voor verticale groepen. Mogelijk heeft dit een iets vertekend beeld gegeven van de kwaliteit van interacties zoals ervaren door baby's. Het LKK rapport van 2018 heeft immers laten zien dat vooral de kwaliteit van interacties met baby's lager is in verticale groepen⁶.

Een andere mogelijke verklaring kan liggen in de specificiteit van het instrument. De NCKO-schalen zijn zo ontwikkeld dat ze toepasbaar zijn in de gehele leeftijdsrange van 0-4 jaar, terwijl de CLASS specifieke leeftijdsversies heeft voor baby's en peuters. Twee NCKO-schalen laten wel iets andere scorepatronen zien voor horizontale en verticale groepen. Zo scoorden horizontale babygroepen vaker onvoldoende op het structureren en grenzen stellen en het begeleiden van interacties tussen kinderen wat er mogelijk op duidt dat deze vaardigheden niet voldoende benut werden of mogelijk minder goed van toepassing zijn voor groepen met alleen baby's. Hoewel de kwaliteitsconcepten in de CLASS dezelfde zijn voor de gehele leeftijdsrange, is de vertaalslag naar het gedrag dat geobserveerd wordt in de CLASS aangepast aan de ontwikkelingsfase van de leeftijdsgroep. In de infant versie kijkt het stimuleren van de taalontwikkeling bijvoorbeeld naar non-verbale en verbale uitingen van interacties (om de beurt praten, gebruik van brabbelen of proto-conversaties), terwijl de toddler versie kijkt naar de open vragen die een medewerker stelt en de mate waarin een medewerker uitingen van kinderen herhaalt en uitbreidt tot volledige zinnen. De CLASS is daardoor mogelijk nauwkeuriger in het in kaart brengen van meer specifieke aspecten van de interacties met kinderen en de NCKO-schalen zijn mogelijk wat globaler in de beoordeling van de interacties.

De ITERS-R schetst een meer globaal beeld van de kwaliteit voor baby's en onderscheidt verschillende kwaliteitsaspecten. De dagelijkse individuele verzorging, de algehele kwaliteit van interacties waaronder interacties gericht op de Nederlandse taalstimulering van baby's en het programma van spel en activiteiten scoren hoog. De inrichting van de ruimte en aanwezige meubilering en vooral de aanwezigheid van materialen en het aanbod van activiteiten scoren beduidend lager. Voor de ITERS-R blijken er geen verschillen tussen horizontale en verticale groepen. Maar ook hier lijken iets andere patronen zichtbaar voor de kwaliteit van de inrichting van de ruimte, het taalgebruik met kinderen, de kwaliteit van het programma en het activiteitenaanbod. In horizontale groepen wordt over het algemeen vaker voldoende tot goed gescoord in vergelijking met matige scores voor verticale groepen, behalve voor het activiteitenaanbod. Het activiteitenaanbod blijkt in verticale groepen vaker

⁶ Slot et al. 2018

voldoende tot goed, maar ook vaker onvoldoende, wat duidt op meer spreiding dan in horizontale groepen die vaker matig scoren.

Het laatste aspect van de proceskwaliteit is gebaseerd op rapportages van de medewerkers. Er is in een interview gevraagd naar het aanbod van exploratie, spel en taalactiviteiten passend bij baby's en dreumesen. De ondersteuning van exploratie en spel blijkt voor baby's het hoogst in zowel horizontale als verticale groepen, terwijl voor dreumesen de mogelijkheden voor exploratie beter zijn in verticale groepen. De ondersteuning van taal voor baby's laat niet veel verschillen zien tussen horizontale en verticale groepen, maar voor dreumesen is er meer aanbod in verticale groepen.

De structurele kwaliteit is in kaart gebracht aan de hand van de vragenlijsten van pedagogisch medewerkers en laten over het algemeen weinig verschillen zien tussen medewerkers die werken in horizontale of verticale groepen. De medewerkers zijn over het algemeen tevreden met hun werk en de omstandigheden waaronder ze werken. Ze rapporteren relatief weinig stress in hun werk, waarbij medewerkers in babygroepen minder stress rapporteren dan collega's werkzaam in verticale groepen. Hoewel er aandacht is voor continue professionalisering, is dit aanbod nog niet heel frequent. Inhoudelijk teamoverleg, als onderdeel van teamprofessionalisering, komt het vaakst voor, maar gemiddeld minder dan één keer per maand. Overigens blijken medewerkers in verticale groepen vaker een vast contract te hebben dan medewerkers in horizontale babygroepen (57% en 78% respectievelijk).

Er blijken verschillen in groepsgrootte en kind-staf ratio die voortkomen uit de verschillen in leeftijd van de kinderen die opgevangen worden in horizontale en verticale groepen. Babygroepen zijn gemiddeld iets kleiner (11 kinderen vergeleken met 12 in verticale groepen) en gemiddeld zijn er vier kinderen per medewerker vergeleken met vijf kinderen voor verticale groepen. De grootste verschillen betreffen de groepssamenstelling van de groepen. Zo hebben verticale groepen een groter aandeel kinderen van laagopgeleide ouders, kinderen met een migratieachtergrond en kinderen met een zorgbehoefte. Dit laatste blijkt vooral uit een hoger aandeel kinderen met behoefte aan taalondersteuning, met gedragsproblemen en met lichamelijke of verstandelijke beperkingen. Dit zou erop kunnen duiden dat kinderen van laagopgeleide ouders, met een migratieachtergrond en kinderen met een zorgbehoefte (wellicht deels overlappend met opleidingsniveau en migratieachtergrond) later instromen in de kinderopvang en om die reden vaker in verticale groepen (en eventueel ook peutergroepen, maar die zijn nu buiten beschouwing gelaten) zitten dan in babygroepen. Hoewel het aandeel kinderen met een vluchtelingenstatus over het algemeen erg laag is, blijken deze kinderen wel iets vaker in babygroepen te zitten (en ook met een grotere variatie) dan in verticale groepen. Mogelijk duidt dit erop dat kinderen met een vluchtelingenstatus al vroeg geïndiceerd worden om in te stromen in de kinderopvang.

Verder blijken er geen verschillen te zijn in het pedagogisch groepsbeleid, zoals de stabiliteit van de groepen of het hanteren van een open-deurenbeleid. Ook de pedagogische omgang met kinderen, zoals de mate waarin affectie getoond wordt aan kinderen en de mate waarin kinderen gestimuleerd worden om lief te zijn voor elkaar en mee te doen aan groepsactiviteiten is vergelijkbaar tussen horizontale en verticale groepen. Het contact met ouders komt in beide type groepen overeen en is veelvuldig aanwezig.

2.3. Methodische aanpak en verantwoording

Dit rapport maakt gebruik van data verzameld tussen 2017 en 2019. Voor een uitgebreidere beschrijving van de verschillende meetinstrumenten en de werkwijze verwijzen we graag naar de rapportages die zijn verschenen in 2017, 2018 en 2019⁷. Dit hoofdstuk rapporteert vooral beschrijvende gegevens. Dat wil zeggen dat gemiddelde scores en de spreiding rondom de gemiddelde scores zijn opgenomen in de tabellen. Voor de leesbaarheid zijn er grafieken toegevoegd die in één oogopslag laten zien wat de resultaten zijn. Om te toetsen of er sprake is van verschillen in kwaliteit tussen horizontale en verticale groepen is middels t-toetsen getoetst of verschillen significant zijn en is er daarnaast bepaald hoe groot de effect grootte is. Hiervoor is gebruik gemaakt van de gangbare maat Cohen's *d* waarbij *d* tussen .20 en .50 een klein effect weergeeft; een *d* tussen .50 en .80 een middelmatig effect; en er bij een *d* van groter dan .80 sprake is van een groot effect.

2.4. Proceskwaliteit

In dit hoofdstuk wordt de proceskwaliteit gepresenteerd. Allereerst worden de gegevens van de observaties beschreven en vergeleken tussen horizontale en verticale groepen. Daarna volgt een sectie over het aanbod van exploratie en taal zoals gerapporteerd door de pedagogisch medewerkers.

2.4.1. CLASS Infant

De gegevens in Tabel 2.4.1 tonen de kwaliteit zoals geobserveerd met de CLASS Infant. Multivariate variantieanalyses laten zien dat er significante verschillen zijn tussen horizontale en verticale groepen op alle CLASS dimensies⁸. De kwaliteit is op alle aspecten hoger in horizontale babygroepen. Voor de emotionele kwaliteit is dit een middelmatig effect terwijl dit voor de educatieve kwaliteit een grote effect grootte laat zien. Op basis van de structurele kenmerken tijdens de observaties blijken de condities iets gunstiger in horizontale groepen

⁷ Slot et al. (2017); Slot et al. (2018); Slot et al. (2019)

⁸ Relatieel klimaat ($F(1,60) = 6.01, p = .02$), Sensitiviteit ($F(1,60) = 5.76, p = .02$), Faciliteren exploratie ($F(1,60) = 15.40, p < .001$) en Vroege taalondersteuning ($F(1,60) = 14.14, p < .001$)

waarbij er significant meer volwassenen aanwezig waren dan in verticale groepen (klein tot middelmatig effect).

Tabel 2.4.1 Geobserveerde kwaliteit van interacties op groepsniveau voor baby's in de kinderdagopvang. Eenheid: Groep (gegevens van 2017 - 2019). Duiding: scores < 3 = laag; scores ≥ 3 en < 5 = matig-voldoende; scores ≥ 5 = hoog.

	KDV baby horizontaal (N=26 groepen)			KDV baby verticaal (N=36 groepen)			Cohen's <i>d</i>
	Range	<i>M</i>	<i>SD</i>	Range	<i>M</i>	<i>SD</i>	
Structurele kenmerken tijdens de observaties							
Aantal kinderen	1.75-12.00	5.45	2.41	2.00-11.00	5.27	2.21	0.08
Aantal volwassenen	1.00-4.00	2.09	0.75	0.67-3.33	1.77	0.54	0.50
Kind-staf ratio	1.50-5.25	2.73	0.93	1.17-8.00	3.00	1.37	0.23
CLASS emotioneel							
Relationeel klimaat	4.50-6.50	5.88	0.62	4.33-6.33	5.51	0.56	0.64
Sensitiviteit	3.75-7.00	5.86	0.90	3.33-7.00	5.27	0.97	0.63
Gemiddelde emotionele kwaliteit							
CLASS educatief							
Faciliteren exploratie	2.00-6.00	4.45	1.16	1.33-5.67	3.31	1.09	1.01
Vroege taalondersteuning	1.75-5.00	3.80	0.89	1.33-4.67	2.91	0.94	0.97
Gemiddelde educatieve kwaliteit							

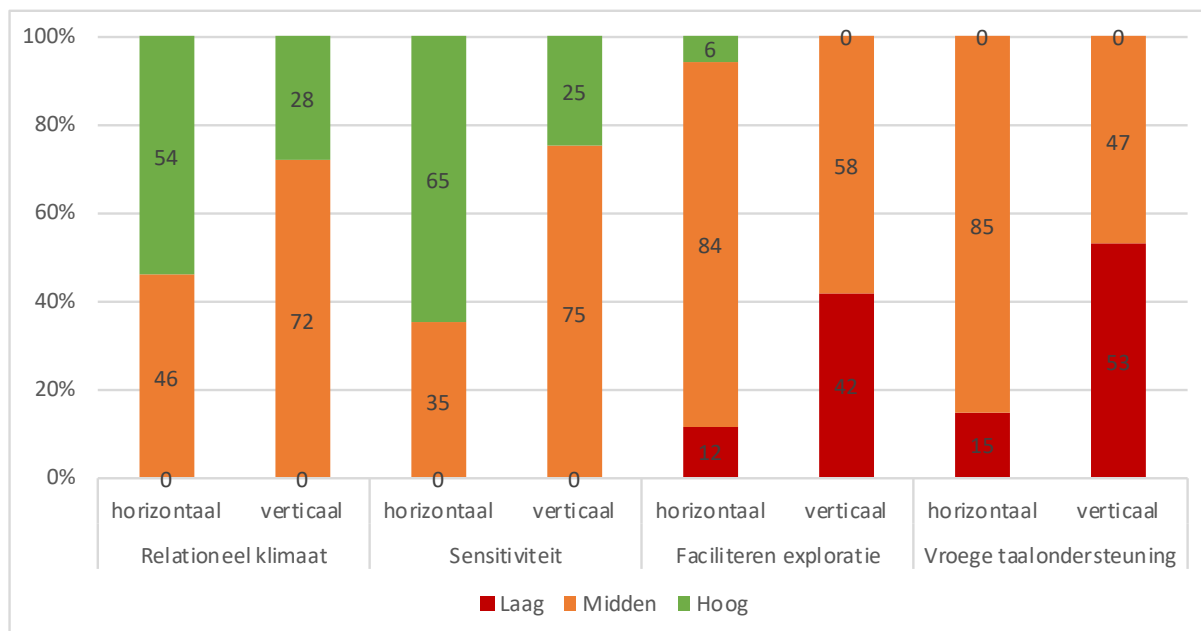
De CLASS Infant scores zijn gebaseerd op verschillende observatiemomenten per groep. Tabel 2.4.2 laat zien hoe de kwaliteit verschilt per type activiteit. Vrij spel is de activiteit die het vaakst geobserveerd werd, gevolgd door eten en drinken. Gerichte activiteiten kwamen in ruim 17% van de observaties voor en betroffen vooral voorlezen, puzzelen, een spelletje doen en in mindere mate knustel of muziekactiviteiten. Er bleken significante verschillen voor alle dimensies⁹. In alle gevallen was de kwaliteit het hoogst tijdens gerichte activiteiten en lager tijdens vrij spel en eten/drinken.

Tabel 2.4.2 Geobserveerde kwaliteit van interacties op groepsniveau in de kinderdagopvang met een uitsplitsing naar type activiteit (gegevens van 2017 - 2019 tezamen: N=200 observatiecycli). Duiding: scores < 3 = laag; scores ≥ 3 en < 5 = matig-voldoende; scores ≥ 5 = hoog.

Activiteit	Activiteit %	Relationeel klimaat		Sensitiviteit pm-er		Faciliteren exploratie		Vroege taalondersteuning	
		<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Eten/drinken	39.5	5.46 ^a	0.95	5.44 ^c	1.24	3.43 ^e	1.48	3.24 ^g	1.23
Vrij spel	43.0	5.66 ^b	0.82	5.49 ^d	1.18	4.03 ^{ef}	1.50	3.33	1.29
Gerichte activiteit	17.5	6.29 ^{ab}	0.67	6.23 ^{cd}	0.97	4.83 ^{ef}	1.47	3.89 ^g	1.23

⁹ Relationeel klimaat ($F(2,197) = 11.67, p < .001$), Sensitiviteit ($F(2,197) = 6.17, p = .003$) Faciliteren exploratie ($F(2,197) = 11.07, p < .001$) en Vroege taalondersteuning ($F(2,197) = 3.40, p = .035$)

Afbeelding 2.4.1 laat zien hoeveel horizontale en verticale groepen (in %) laag, midden en hoog scoren op de vier CLASS dimensies. Dit bevestigt de resultaten uit Tabel 2.4.1. In horizontale groepen wordt vaker hoog gescoord op de emotionele kwaliteit en minder vaak laag op de educatieve kwaliteit in vergelijking met verticale groepen.



Afbeelding 2.4.1 Percentage kinderopvanggroepen met lage, midden en hoge kwaliteit volgens de CLASS standaards uitgesplitst naar type groep (gegevens van 2017 - 2019 tezamen).

In Tabel 2.4.3 staat een overzicht van de kwaliteit zoals gemeten met de CLASS Infant in andere landen. Hieruit blijkt dat de kwaliteit in Nederland gemiddeld hoog is in vergelijking met andere landen, dit geldt vooral de emotionele kwaliteit.

Tabel 2.4.3: Overzicht van kwaliteit gemeten met de CLASS Infant in verschillende landen.

Land	Studie	Steekproef	N	Domein/dimensie	M	SD
Nederland	LKK 2017 - 2019	Gestratificeerde random steekproef (landelijk)	62	Emotionele kwaliteit	5.60	.80
				Educatieve kwaliteit	3.54	1.13
België	Hulpia et al. 2016	Gestratificeerde random steekproef (Vlaanderen)	167	Emotionele kwaliteit	4.99	.91
				Educatieve kwaliteit	3.15	1.18
Portugal	Barros et al., 2016	Stedelijk gebied rondom Porto	90	Emotionele kwaliteit	4.41	.82
				Educatieve kwaliteit	2.60	.76
V.S.	Jamison et al., 2014	Eén staat	30	Emotionele kwaliteit	5.10	.96
				Educatieve kwaliteit	3.96	1.05

2.4.2 NCKO-interactievaardigheden

De NCKO-schalen onderscheiden zes verschillende vaardigheden¹⁰: Sensitieve reponsiviteit, Respect voor autonomie, Structureren en grenzen stellen, Praten en uitleggen, Ontwikkelingsstimulering en Begeleiden van interacties (tussen kinderen). Sensitieve reponsiviteit, Respect voor autonomie en Structureren en grenzen stellen worden gezien als de basale vaardigheden¹¹ en vallen onder de noemer emotionele proceskwaliteit. De andere drie vaardigheden zijn meer educatieve vaardigheden¹² en behoren tot de educatieve proceskwaliteit. De zes vaardigheden worden beoordeeld op basis van videomateriaal dat verzameld is van meerdere pedagogisch medewerkers in de groep. De resultaten staan in Tabel 2.4.4 en laten zien dat de basale vaardigheden hoger scoren dan de educatieve vaardigheden. Er blijken daarbij geen significante verschillen tussen horizontale en verticale groepen.

Tabel 2.4.4: Gemiddelde scores op de NCKO-interactievaardigheden (op groepsniveau).

Duiding: scores < 3 = laag; scores ≥ 3 en < 5 = matig-voldoende; scores ≥ 5 = hoog.

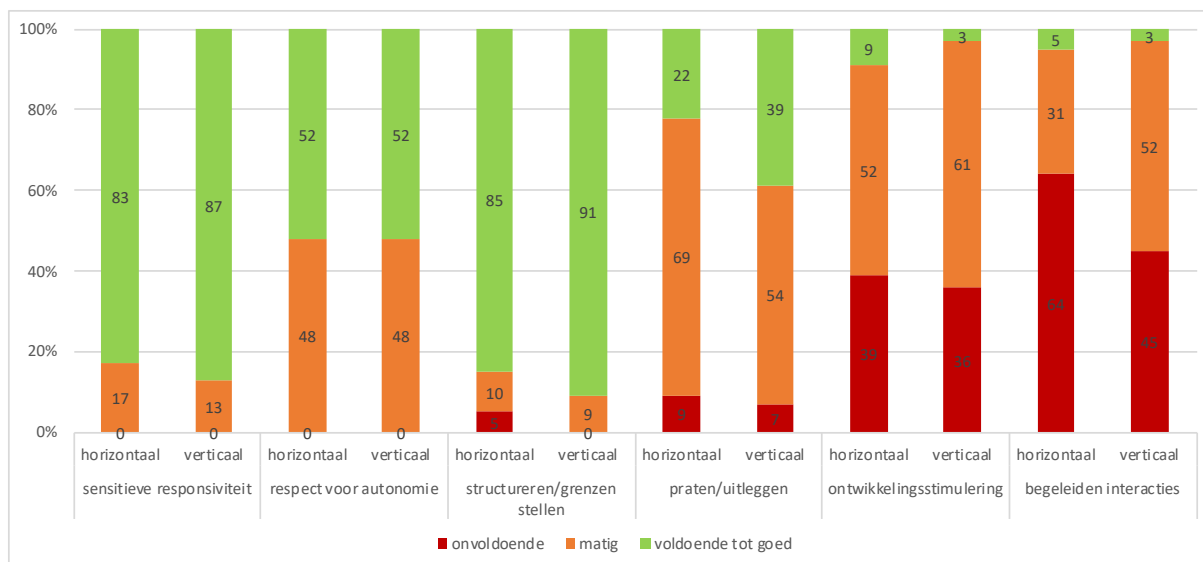
		KDV baby horizontaal (N=23 groepen)		KDV baby verticaal (N=31 groepen)	
	<i>Schaal</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Sensitieve reponsiviteit	1-7	5.49	0.99	5.43	0.76
Respect voor autonomie	1-7	4.93	0.71	5.09	0.84
Structureren en grenzen stellen	1-7	5.64	1.43	5.71	0.81
Praten en uitleggen	1-7	4.16	0.79	4.41	0.94
Ontwikkelingsstimulering	1-7	3.35	1.18	3.38	0.91
Begeleiding van interacties	1-7	2.62	1.14	2.78	1.19

In Afbeelding 2.4.2 staat de verdeling van onvoldoende, matige en voldoende tot goede scores voor horizontale en verticale groepen (in %). Voor de meeste vaardigheden lijken er weinig verschillen tussen horizontale en verticale groepen, behalve voor Structureren en grenzen stellen en Begeleiden van interacties. Een mogelijke verklaring kan zijn dat deze aspecten zoals ze omschreven worden in de schaal minder goed van toepassing zijn op jongere kinderen.

¹⁰ De Kruif et al. (2009).

¹¹ Fukkink et al. (2013).

¹² Fukkink et al. (2013).



Afbeelding 2.4.2 Verdeling van pedagogisch medewerkers naar onvoldoende, matige en voldoende tot goede kwaliteit in percentages (gegevens van 2017 - 2019 tezamen).

2.4.3. Omgevingschaal ITERS-R

De ITERS-R¹³ is een instrument speciaal gericht op baby's en meet op globale wijze verschillende aspecten van kwaliteit waaronder de inrichting van de ruimte en aanwezige meubilering, interacties die taal en redeneren bij kinderen aanmoedigen, een breed aanbod van activiteiten en materialen en het omvat een globale schaal voor de kwaliteit van interacties¹⁴. Tabel 2.4.5 rapporteert de uitkomsten en laat geen significante verschillen zien tussen horizontale en verticale groepen. De dagelijkse individuele verzorging, de schaal Luisteren en praten, de algemene interactieschaal en het Programma scoren hoog, terwijl de kwaliteit van de ruimte en vooral het activiteitenaanbod lager scoren.

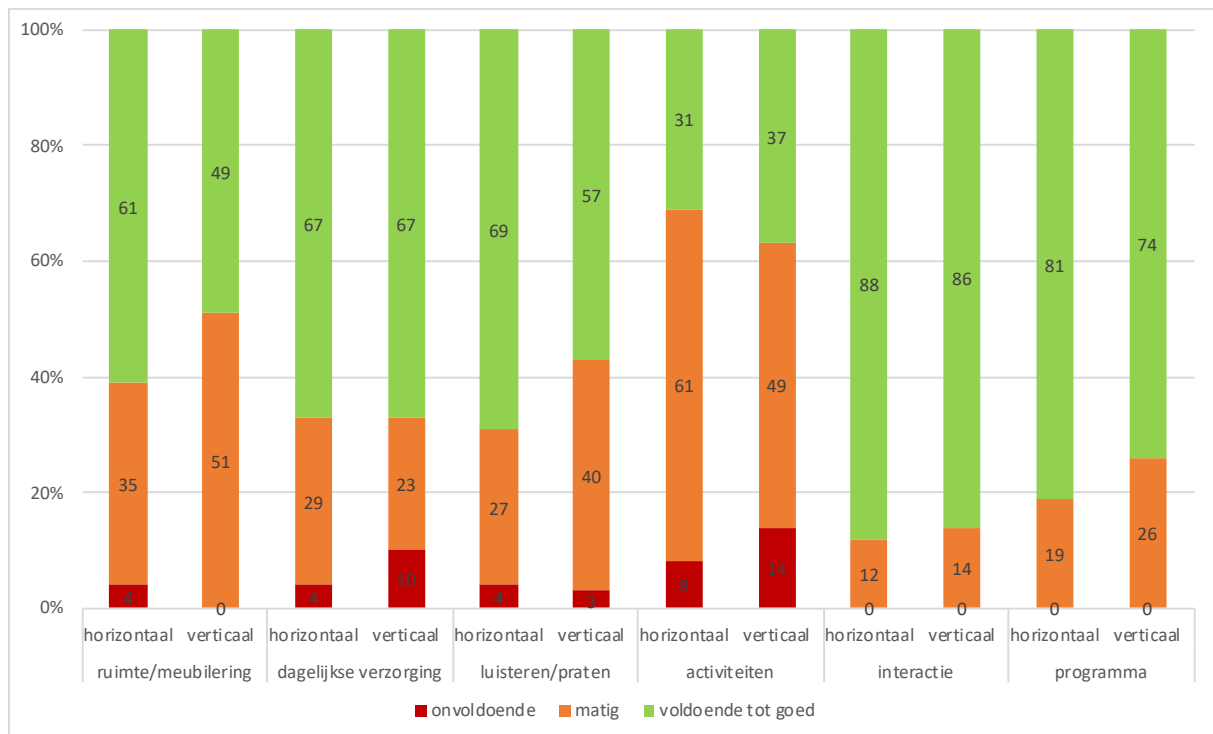
Tabel 2.4.5 Geobserveerde proceskwaliteit op basis van de omgevingschaal ITERS-R in de baby opvang. Eenheid: Groep (gegevens van 2017 - 2019 tezamen). Duiding: scores < 3 = laag; scores ≥ 3 en < 5 = matig-voldoende; scores ≥ 5 = hoog.

Subschalen	KDV babygroepen (N=26 groepen)			KDV verticale groepen (N=35 groepen)		
	Range	M	SD	Range	M	SD
Ruimte en meubilering	2.80-6.20	4.90	0.93	3.40-6.60	4.86	0.84
Dagelijkse individuele verzorging	2.00-7.00	5.79	1.67	1.00-7.00	5.60	1.89
Luisteren en praten	3.00-7.00	5.53	1.43	2.67-7.00	5.17	1.12
Activiteiten	2.00-6.44	4.34	1.13	2.56-6.22	4.50	1.07
Interactie	3.75-7.00	6.23	0.96	3.75-7.00	6.13	1.03
Programma	3.50-7.00	5.92	1.18	2.67-7.00	5.74	1.29
Totaal	3.40-6.36	5.14	0.88	3.32-6.44	5.10	0.83

¹³ Harms, Cryer, & Clifford (2006).

¹⁴ Zie Slot et al. (2018) voor een uitgebreidere beschrijving van de observatieschaal

Afbeelding 2.4.4 toont de verdeling van het percentage groepen dat onvoldoende, matig en voldoende tot goede score op de subschalen van de ITERS-R. Hoewel er geen verschillen zijn in gemiddelde scores, lijkt er voor sommige schalen wel een verschil te zijn in de verdeling van de scores. Zo scoren horizontale groepen vaker voldoende tot goed op de ruimte en meubilering (en ook enkele onvoldoendes), de mate waarin er geluisterd en met kinderen gepraat wordt en de kwaliteit van het programma. Voor het activiteitsaanbod daarentegen lijkt het erop dat verticale groepen iets vaker voldoende tot goed scoren evenals een onvoldoende, terwijl horizontale groepen vaker matig scoren.



Afbeelding 2.4.4 Percentage groepen met onvoldoende, matige en voldoende tot goede kwaliteit op de ITERS-R subschalen in 61 groepen (gegevens van 2017 - 2019 tezamen).

Tabel 2.4.6 laat een internationale vergelijking zien voor de ITERS-R. Deze bevindingen zijn grotendeels gebaseerd op verticale groepen, inclusief oudere kinderen. Desalniettemin laten de gegevens uit de LKK zien dat de kwaliteit in Nederland tot de hoogste behoort, vergelijkbaar met Australië en de Verenigde Staten. De studie uit de Verenigde Staten toonde dat de kwaliteit van de inrichting van de ruimte, de globale interacties en het programma hoger scoorde in horizontale babygroepen terwijl de verticale groepen hoger scoorden op de aanwezige materialen en het activiteitsaanbod.

Tabel 2.4.6 ITERS-R scores in verschillende landen op basis van recente studies. Duiding: scores < 3 = laag; scores \geq 3 en < 5 = matig-voldoende; scores \geq 5 = hoog.

Land	Studie	Steekproef	N	M	SD
Nederland	LKK 2017 -2019	Gestratificeerde random steekproef (landelijk)	61	5.12	0.85
Australië	Fenech et al., 2010	Horizontale en verticale baby-dreumesgroepen in één staat	48	5.17	0.72
Noorwegen	Bjørnstad & Os, 2018	Verticale groepen in representatieve steekproef obv vier regio's in Noorwegen	206	4.0	0.76
Portugal	Barros et al 2016	Stedelijk gebied rondom Porto	90	3.41	0.94
Verenigde Staten	King et al., 2016	Horizontale en verticale baby-dreumesgroepen in één staat	766	4.96	0.91

2.4.4 Activiteitsaanbod

Om te bepalen in hoeverre het aanbod aan regelmatig terugkerende activiteiten aansluit bij belangrijke leer-, ervarings- en ontwikkelingsdomeinen is een vragenlijst gecreëerd bestaande uit verschillende soorten activiteiten die relevant worden geacht voor baby's met betrekking tot exploratie-motorische ontwikkeling en taalontwikkeling. Onderzoek heeft laten zien dat de motorische ontwikkeling van kinderen samenhangt met de taalontwikkeling en dat exploratiegedrag van kinderen hier een belangrijke rol in speelt¹⁵. Om die reden wordt er hier ingezoomd op deze twee ontwikkelingsdomeinen, zie Tabel 2.4.7. De vragenlijst schalen zijn ontwikkeld om de hele leeftijdsrange af te dekken met leeftijdsspecifieke items van concrete activiteiten. Om die reden zijn er aparte schalen gemaakt van activiteiten die passend zijn voor baby's jonger dan één jaar en voor activiteiten beter passend bij de dreumesleeftijd.

¹⁵ Oudgenoeg-Paz, Leseman, & Volman (2015)

Tabel 2.4.7 Items van de vragenlijst schalen

	Baby	Dreumes
Domein		
Exploratie en spel	<p>Kinderen pakken verschillende voorwerpen, voelen eraan, sabbelen erop, laten het vallen (bijv. mobiel, bijtring, pop, ...)</p> <p>Kinderen spelen met voorwerpen (blokjes, bakjes, doos, bal) die ze kunnen stapelen , in elkaar doen, uit een bakje of doos halen, er in stoppen, op een rijtje leggen, weg rollen</p> <p>Kinderen spelen, alleen of naast elkaar, met voorwerpen, voelen en proeven eraan, herhalen vaak hun spelhandelingen zoals verplaatsen, laten vallen, draaien</p>	<p>Kinderen kruipen of lopen door de groepsruimte, verkennen actief de ruimte, verstoppem zich onder de tafel, in een andere ruimte, klimmen op een verhoging</p> <p>Kinderen rijden op loopfietsjes, stappen, klimmen op een glijbaan, bewegen op een vrije manier, spelen met een bal of springtouw</p> <p>Kinderen bouwen , alleen of naast elkaar, dingen met houten blokken, Duplo, Lego, Knexx, ze maken een Brio treinbaan</p>
Taal	<p>Kinderen doen een spelletje met klanken uitwisselen (brabbelen), aankijken en gebaartjes maken, alsof je een gesprekje voert</p> <p>Kinderen luisteren actief naar melodieën, kinderliedjes en ritmes, klappen ritmes, zingen en neuriën met de melodie mee</p>	<p>Kinderen doen een spelletje met voorwerpen aanwijzen, de vorm (bijv. de rondheid van een bal) uitbeelden en benoemen van voorwerpen of handelingen ('dit is een...')</p> <p>Kinderen lezen een plaatjesboek of bespreken een praatplaat, er wordt een verhaaltje verteld aan de hand van de plaatjes</p> <p>Kinderen luisteren actief naar melodieën, kinderliedjes en ritmes, klappen ritmes, zingen en neuriën met de melodie mee</p>

In Tabel 2.4.8 is een overzicht van gemiddelde scores gegeven op een frequentieschaal van 1 '(bijna) nooit' tot 7 'minstens twee keer per dag'. Bedacht moet worden dat een gemiddelde score voor de hele schaal verhult dat sommige activiteiten die in de schaal zijn samengevat vaker voorkomen, terwijl andere activiteiten minder vaak kunnen voorkomen. De medewerkers rapporteren relatief hoge gemiddelde scores voor exploratie en spel (gemiddeld tussen 'dagelijks' en 'minstens twee keer per dag'). Het aanbod van taal ligt iets lager, maar dit komt nog altijd wekelijks voor (gemiddeld tussen '2 tot 4 keer per week' en 'dagelijks'). Voor iets oudere kinderen (dreumesen van 1/1½ tot 2/2½ jaar) blijkt het aanbod van exploratie en taal iets lager. Opvallend is de grote spreiding in het aanbod van taal voor zowel baby's als peuters. Daarnaast blijkt er ook veel variatie in de exploratiemogelijkheden voor baby's in verticale groepen en omgekeerd voor dreumesen in babygroepen.

Voor dreumesen blijkt het aanbod van exploratie en spel in verticale groepen significant hoger dan in babygroepen¹⁶. Een paarsgewijze vergelijking laat zien dat in babygroepen het aanbod

¹⁶ Exploratie dreumesen ($F(1,91) = 6.47, p = .013$) en een trend voor Taalaanbod voor dreumesen ($F(1,91) = 3.11, p = .081$)

van exploratie, spel en taal hoger is voor baby's dan dreumesen¹⁷. Voor verticale groepen geldt dat alleen het aanbod van taal significant hoger is voor baby's in vergelijking met dreumesen¹⁸.

Tabel 2.4.8 Activiteitenaanbod zoals gerapporteerd door pedagogisch medewerkers (gegevens van 2017 - 2019 tezamen).

Activiteitenaanbod	KDV baby (N=26 groepen)				KDV verticaal (N=36 groepen)			
	N	Range	M	SD	N	Range	M	SD
Exploratie voor baby's (1-7)	37	5.67-7.00	6.58	0.48	56	1.00-7.00	6.34	0.89
Taal voor baby's (1-7)	36	1.00-7.00	5.71	1.46	55	1.00-7.00	5.54	1.35
Exploratie voor dreumesen (1-7)	37	2.67-7.00	5.77	1.14	56	4.00-7.00	6.24	0.65
Taal voor dreumesen (1-7)	36	1.00-6.75	4.24	1.46	55	1.67-6.50	4.70	1.03

2.4.5. Samenvattend beeld van de proceskwaliteit

Eerdere LKK-rapportages in 2017 en 2018 hebben aangetoond dat er kwaliteitsverschillen zijn tussen horizontale babygroepen en verticale kinderopvanggroepen. In het huidige rapport gaan we hier dieper op in. Daarvoor gebruiken we gegevens die zijn verzameld tussen 2017 en 2019 bij in totaal 62 groepen. Het beeld op basis van de observaties laat op verschillende aspecten verschillen zien in de gemiddelde kwaliteit en de patronen van lage, matige en voldoende tot goede kwaliteit. Met name de CLASS Infant laat grote verschillen zien op zowel de emotionele als de educatieve kwaliteit, waarbij de laatste grote effecten laat zien. De kwaliteit blijkt in horizontale babygroepen gemiddeld hoger te zijn en er blijken minder groepen die laag scoren in vergelijking met verticale groepen. Vooral voor de educatieve kwaliteit zijn de gevonden verschillen groot. Voor horizontale groepen geldt dat meer dan 80% in de middenrange of een enkele groep zelfs hoog scoort op de educatieve aspecten terwijl voor verticale groepen (bijna) de helft laag scoort.

De resultaten met de NCKO-schalen laten geen verschillen zien in gemiddelde scores tussen horizontale en verticale groepen. Een mogelijke verklaring die voor de hand ligt, is dat de NCKO gescoord is voor de gehele groep, dus inclusief de peuters als die er waren. De CLASS Infant heeft expliciet de interacties met baby's in ogenschouw genomen en de eventuele interacties met peuters buiten beschouwing gelaten en geeft daardoor een scherper beeld van de ervaringen van baby's. Overigens bleken er voor twee NCKO-schalen wel iets andere scorepatronen te zijn voor horizontale en verticale groepen. Zo scoorden horizontale babygroepen vaker onvoldoende op het structureren en grenzen stellen en het begeleiden van interacties tussen kinderen wat er mogelijk op duidt dat deze vaardigheden niet voldoende benut werden of mogelijk minder goed van toepassing zijn voor groepen met alleen baby's.

¹⁷ Exploratie ($t = 4,58$, $p < .001$) en taalaanbod ($t = 6,48$, $p < .001$)

¹⁸ Taalaanbod $t = 5,37$, $p < .001$)

Voor de omgevingschaal ITERS-R geldt dat er gemiddeld genomen geen verschillen zijn tussen horizontale en verticale groepen. Maar ook hier lijken iets andere patronen zichtbaar voor de kwaliteit van de inrichting van de ruimte, het luisteren en praten met kinderen, de kwaliteit van het programma en het activiteitenaanbod. In horizontale groepen wordt over het algemeen vaker voldoende tot goed gescoord in vergelijking met matige scores voor verticale groepen, behalve voor het activiteitenaanbod. Het activiteitenaanbod blijkt in verticale groepen vaker voldoende tot goed, maar ook vaker onvoldoende, wat duidt op meer spreiding dan in horizontale groepen die vaker matig scoren.

Op basis van het zelf gerapporteerde activiteitenaanbod blijkt dat er dagelijks aandacht is voor exploratie en spel voor baby's in zowel horizontale als verticale groepen. Ruimte voor exploratie voor oudere kinderen lijkt beter tot zijn recht te komen in verticale groepen dan in horizontale babygroepen. Het aanbod van taal komt iets minder frequent voor dan aanbod van exploratie en spel, maar nog altijd meerdere keren per week. Er blijken verschillen tussen horizontale en verticale groepen, maar ook binnen deze groepen blijkt het aanbod voor baby's en dreumesen anders. Zowel horizontale als verticale groepen kunnen baby's (0 tot 1/1 ½ jaar) en dreumesen (1/1 ½ tot 2/2 ½ jaar) opvangen en daarom is het relevant om naar het aanbod voor beide leeftijdsgroepen te kijken. Dreumesen lijken beter af in verticale groepen waar meer ruimte is voor exploratie en de verschillen in taalaanbod voor baby's en peuters minder groot zijn. Baby's daarentegen lijken het meest te profiteren van het aanbod in horizontale babygroepen.

2.5 Structurele kwaliteit

In totaal zijn tijdens de metingen van 2017, 2018 en 2019 met 105 pedagogisch medewerkers, werkzaam op 62 kinderopvanggroepen, interviews gehouden. Daarvan werkten er 44 in groepen met alleen baby's en dreumesen, 37 in peutergroepen en 61 in verticale groepen met baby's, dreumesen en peuters. Hieronder worden de bevindingen uit de interviews gerapporteerd voor verschillende aspecten: structurele kwaliteit en groepssamenstelling, pedagogisch beleid en pedagogisch klimaat, en activiteitenaanbod.

2.5.1 Kenmerken van de medewerkers en de groep

Tabel 2.5.1 geeft een overzicht van de achtergrond van de geïnterviewde medewerkers van horizontale en verticale groepen. Voor de meeste aspecten zijn er weinig verschillen in achtergrond tussen medewerkers die op horizontale of verticale groepen werken. De gemiddelde leeftijd ligt tussen de 32 en 34 jaar. De spreiding in leeftijd is groot. De werktijd per week ligt tussen de 26 en 28 uur en ook hier is er grote spreiding. Het overgrote deel van de medewerkers is opgeleid op MBO3 of MBO4 niveau (het gemiddelde voor- en vervolgopleidingsniveau ligt rond de 5). Er zijn wat betreft vooropleiding zowel lager opgeleiden (VMBO of MBO2 niveau) als hoger opgeleiden (HBO). Door de combinatie van

vooropleiding en vervolgopleiding voldoen alle medewerkers aan de wettelijke eis op dit vlak. Er zijn wel verschillen in de gemiddelde werkervaring van de geïnterviewden waarbij medewerkers werkzaam in verticale groepen beduidend meer ervaring hebben, maar de variatie is groot voor beide groepen. Ook voor de mate waarin medewerkers een vast contract hebben, blijkt er een verschil tussen medewerkers van horizontale en verticale groepen. Zo heeft bijna 57% van de medewerkers van horizontale groepen een vast contract tegenover 78% van de medewerkers in verticale groepen.

Tabel 2.5.1 Achtergrond van de pedagogisch medewerkers (gegevens van 2017 - 2019 tezamen).

Achtergrond medewerker	KDV baby (N= 26 groepen)				KDV verticaal (N= 36 groepen)			
	N	Range	M	SD	N	Range	M	SD
Leeftijd	44	19.00-65.00	32.39	10.08	60	19.00-54.00	34.33	9.25
Werkdagen	44	1.00-5.00	3.34	1.08	60	1.00-5.00	2.93	0.88
Werktijd in uren ¹⁹	43	11.75-55.00	28.49	8.25	60	12.00-45.00	26.36	6.76
Vooropleiding (1-8)	42	1.00-8.00	4.86	1.68	60	1.00-7.00	4.30	1.76
Niveau vervolgopleiding (1-8)	42	2.00-8.00	5.40	0.86	60	2.00-7.00	5.25	0.84
Werkervaring in kinderopvang in jaren	43	0.00-30.00	8.56	8.56	60	0.00-30.00	12.08	7.79
Vast contract (%)	25		56.80		46		78.00	

De werkomgeving wordt in het algemeen als positief ervaren en hierbij zijn weinig verschillen tussen medewerkers van horizontale en verticale groepen, zoals Tabel 2.5.2 laat zien. Op vragen naar de onderlinge afstemming en samenwerking in het team, het werkklimaat met aspecten als zich gerespecteerd en gewaardeerd voelen, en het gevoel betrokken te zijn bij beleidsontwikkeling binnen de organisatie, wordt gemiddeld rond de 5 ('mee eens') gescoord op een schaal van 1 tot en met 6. Het ervaren stressniveau is laag (positief gehercodeerd; het antwoord op stellingen als "ik ervaar regelmatig stress op werk" of "ik ben meer uitgeput dan normaal na een werkdag" is gemiddeld 'een beetje mee oneens' tot 'mee oneens'). Wel geven medewerkers van babygroepen iets minder werkstress aan dan medewerkers op verticale groepen.

¹⁹ Het maximum aantal uur van 55 betreft hier een ureninschatting van een leidinggevende/eigenaar die tevens werkzaam is als pedagogisch medewerker op de groep.

Tabel 2.5.2 Werkbeleving zoals gerapporteerd door pedagogisch medewerkers (gegevens van 2017 - 2019 tezamen).

Organisatieklimaat en werkbeleving	KDV baby (N= 26 groepen)				KDV verticaal (N= 36 groepen)			
	N	Range	M	SD	N	Range	M	SD
Continue professionalisering (1-5)	39	1.00-4.00	2.01	0.64	57	1.00-3.50	1.95	0.53
Teamgerichte professionalisering (1-7)	39	1.56-5.56	2.78	0.97	57	1.00-5.33	2.50	0.85
Frequentie teamoverleg (0-8)	39	0.00-8.00	3.90	1.93	57	0.00-7.00	4.02	2.02
Teamcohesie (1-6)	39	2.60-6.00	4.81	0.72	57	3.20-6.00	4.92	0.54
Zelfvertrouwen (1-6)	39	4.33-5.78	4.89	0.39	57	4.11-5.67	4.96	0.41
Werktevredenheid (1-6)	39	3.67-5.88	5.06	0.50	57	3.42-5.83	4.95	0.53
Beleving van inclusie op de werkvloer (1-6)	39	4.40-6.00	5.37	0.52	57	3.40-6.00	5.42	0.54
Werkstress (1-6) (R)	39	2.40-6.00	4.61	0.80	57	2.60-6.00	4.29	0.84

Er is aandacht voor continue professionalisering door middel van cursusaanbod, coaching, teamoverleg en conferentiedeelnemers, hoewel de gemiddelde score rond de 2 ('zelden') ligt, wat suggereert dat het aanbod niet overdadig is. De meest voorkomende vorm van continue professionalisering is pedagogisch-inhoudelijk teamoverleg ('regelmatig'), oftewel één keer per twee maanden, zoals ook meer gedetailleerd is nagevraagd. De wijze waarop het teamoverleg vorm wordt gegeven is apart nagevraagd. De schaal teamgerichte professionalisering gaat in op overleg, samenwerking, feedback en reflectie in het team rond onder andere het activiteitenaanbod, individuele kinderen die speciale aandacht behoeven, nieuwe vakinhoudelijke kennis en de pedagogische doelen van het werk. De gemiddelde score tussen 2.50 en 2.78 op een schaal van 1 ('nooit') tot 7 ('elke dag') komt overeen met minder dan één keer per maand.

Gegevens over de groep kinderen waarmee de geïnterviewden werken staan vermeld in Tabel 2.5.3. De gemiddelde groepsgrootte op de drie drukste dagen van de week (maandag, dinsdag en donderdag) is ongeveer 11 kinderen en iets groter voor verticale groepen met ruim 12 kinderen, passend bij de oudere leeftijd van kinderen. De hoogste waarde voor de groepsgrootte is 18 (samengevoegde groepen, vermoedelijk aan het begin- of het eind van de dag). De kind-staf ratio op de drukste dagen is 4.06 voor babygroepen en 5.05 voor verticale groepen wat samenhangt met de verschillen in leeftijd van de kinderen. De samenstelling van de groepen naar leeftijdsfase komt overeen met de aard van de groepen: deels babygroepen met daarin ook dreumesen van 1½ tot 2 à 2½ jaar, deels verticale groepen met baby's, dreumesen en peuters. De pedagogisch medewerkers hebben ingeschat hoeveel van de kinderen in hun groep ouders hebben die laag zijn opgeleid (maximaal VMBO of MBO1/2) of een migratieachtergrond hebben. Op basis van hun antwoorden schatten wij het aandeel kinderen van laagopgeleide ouders op gemiddeld 13.55% en, daarmee deels overlappend, het aandeel kinderen met ouders met een migratieachtergrond, Westers zowel als niet-Westers, op 25.69% voor babygroepen en 22.74% respectievelijk 31.32% voor verticale groepen. Er blijkt dus meer diversiteit in verticale groepen wat mogelijk verklaard zou kunnen worden

door instroom op oudere leeftijd van deze groep kinderen. Er zijn kinderopvanggroepen die volledig uit kinderen bestaan van laagopgeleide of migrantenouders. Ook zijn er groepen met vluchtelingenkinderen of kinderen met een lichamelijke of verstandelijke beperking, hoewel het aantal kinderen met deze achtergronden beperkt is. Ten dele corresponderend hiermee, rapporteren de pedagogisch medewerkers een aandeel kinderen die extra zorg of (educatieve) steun nodig hebben in babygroepen van gemiddeld 11.44% en 21.54% voor verticale groepen, waarvan een deel met behoefte aan taalondersteuning, en een deel met behoefte aan gedragsbegeleiding. Het aandeel kinderen met een zorgbehoefte is aanzienlijk groter in verticale groepen, wat mogelijk ook duidt op instroom van deze kinderen op oudere leeftijd.

Tabel 2.5.3 Samenstelling van de groep gerapporteerd door pedagogisch medewerkers (gegevens van 2017 - 2019 tezamen).

		KDV baby (N=26 groepen)				KDV verticaal (N=36 groepen)			
Structurele kenmerken en groepssamenstelling		N	Range	M	SD	N	Range	M	SD
Groepsgrootte op de drie drukste dagen		41	5.00-17.00	11.29	2.80	59	7.00-18.00	12.37	2.34
Kind-staf ratio op de drie drukste dagen		41	2.17-7.00	4.06	0.84	58	3.81-7.67	5.05	0.77
Baby's (%)		38	7.00-100.00	50.32	22.92	51	0.00-73.00	26.31	14.14
Dreumesen (%)		37	0.00-89.00	43.20	20.27	52	0.00-69.00	37.81	14.31
Peuters (%)		37	0.00-64.00	7.82	16.56	50	0.00-100.00	39.84	21.25
Kinderen van laagopgeleide ouders (%)		24	0.00-100.00	13.55	21.91	41	0.00-100.00	22.74	33.20
Kinderen met een migratieachtergrond (%)		25	0.00-100.00	25.69	32.63	41	0.00-100.00	31.32	31.50
Kinderen met zorgbehoefte (%)		38	0.00-50.00	11.44	13.79	54	0.00-100.00	21.54	22.24
Waarvan taalondersteuning (%)		38	0.00-75.00	8.17	16.85	53	0.00-100.00	13.31	19.28
Waarvan gedragsproblemen (%)		37	0.00-38.00	2.34	6.94	52	0.00-50.00	6.03	10.12
Waarvan beperkingen (%)		37	0.00-25.00	1.14	4.49	54	0.00-18.00	2.13	4.22
Waarvan vluchtelingenstatus (%)		37	0.00-45.00	2.92	9.02	54	0.00-19.00	1.09	3.98

2.5.2 Pedagogisch beleid en pedagogisch klimaat

Het pedagogisch beleid van de deelnemende instellingen is op verschillende manieren in kaart gebracht; zie Tabel 2.5.4. Op vragen naar aspecten van stabiliteit van de groep (onder andere of er gedurende de dag groepen worden samengevoegd of gesplitst, of er gemakkelijk tijdelijk van opvangdag gewisseld kan worden) wordt een gemiddelde score van ruim een 3 gehaald ('de genoemde stabiliteitsaspecten zijn min of meer van toepassing'). Vooral de wijde scorerange wijst er op dat er locaties en groepen zijn met een sterke oriëntatie op stabiliteit

van de groep. Er kan in deze groepen niet makkelijk van dag gewisseld worden, de groepen worden niet of zelden samengevoegd of gesplitst. Er zijn ook groepen waarin het wisselen van dagen en veranderen van groepen juist veel gebeurt. In samenhang hiermee is er sterke variatie tussen de groepen in de mate waarin een opendeurenbeleid wordt gehanteerd, afgemeten aan aspecten zoals kinderen die zelf kunnen kiezen in welke groep ze bepaalde activiteiten willen doen. De gemiddelde score suggereert dat het opendeurenbeleid nog geen overwegend pedagogisch model is in de kinderopvang voor 0- tot 4-jarigen, althans niet in deze steekproef, maar er zijn groepen waarin een dergelijk beleid wel geïmplementeerd wordt.

Tabel 2.5.4 Pedagogisch beleid en klimaat zoals gerapporteerd door pedagogisch medewerkers (gegevens van 2017 - 2019 tezamen).

Pedagogisch beleid en pedagogisch klimaat	KDV baby (N=26 groepen)				KDV verticaal (N=36 groepen)			
	N	Range	M	SD	N	Range	M	SD
Stabiliteit groepen (1-5)	42	2.50-4.50	3.51	0.73	59	2.00-4.83	3.42	0.59
Opendeurenbeleid (1-5)	42	1.00-5.00	3.76	0.95	59	1.00-5.00	3.69	0.87
Affectief-inclusief gedrag (1-7)	39	4.25-7.00	6.04	0.60	58	4.25-7.00	5.99	0.69
Contact met ouders (% van 14 opties)	39	0.07-0.93	0.59	0.19	58	0.00-0.86	0.57	0.14

Het pedagogisch klimaat en de algemene strategie van omgaan met kinderen is in verschillende onderdelen nagevraagd tijdens de interviews. Affectief-inclusief gedrag verwijst naar een benadering waarin de pedagogisch medewerker kinderen regelmatig knuffelt, persoonlijk troost als dat nodig is, een aai over de bol geeft en bevordert dat alle kinderen in de groep mee doen met groepsactiviteiten en lief zijn voor elkaar. De gemiddelde score wijst op een sterke en dominante (want geringe spreiding van scores) affectief-inclusieve pedagogische houding van de medewerkers in de kinderopvang (6.04 en 5.99 op een schaal van 1 tot 7: zulke handelingen worden 'vaak' tot 'zeer vaak' uitgevoerd).

Oudercontacten verlopen op verschillende manieren: via haal- en brengcontacten, ouderavonden, digitale platforms en app-groepen, tentoonstellingen en feesten, deelname aan een oudercommissie, enzovoort. Uit een lijst van 14 vormen van oudercontact vinkten de geïnterviewde medewerkers er gemiddeld acht aan (57-59%). Gemiddeld is er veelvuldig en veelvormig contact met ouders. De variatie is echter heel groot, met medewerkers die zeggen maar op één of twee manieren contact hebben met ouders en medewerkers die op bijna alle 14 manieren contact hebben. Vrijwel alle medewerkers melden dat er een oudercommissie is.

2.5.4 Samenvattend beeld van de interviewgegevens

Er zijn interviews gehouden met in totaal 105 pedagogisch medewerkers die werkzaam waren in 62 groepen. Over het algemeen blijken er weinig verschillen in structurele groepskenmerken, werkbeleving en pedagogisch aanbod in de groep tussen horizontale en verticale groepen. Er zijn een aantal aspecten die wel verschillen laten zien en die lichten we eruit. Zo hebben medewerkers die werken in horizontale groepen minder vaak een vast contract in vergelijking met medewerkers werkzaam in verticale groepen (57% en 78% respectievelijk). Een ander verschil betreft de ervaren werkstress, waarbij medewerkers van babygroepen minder werkstress rapporteren. Verder wordt de werkomgeving als positief ervaren en lijken die ervaringen niet wezenlijk te verschillen tussen medewerkers in horizontale en verticale groepen.

Er zijn verschillen in de gemiddelde groeps grootte en kind-staf ratio die gerelateerd zijn aan de verschillende leeftijden van de kinderen in horizontale en verticale groepen. Babygroepen zijn gemiddeld iets kleiner (11 kinderen vergeleken met 12 in verticale groepen) en gemiddeld zijn er vier kinderen per medewerker vergeleken met vijf kinderen voor verticale groepen. De grootste verschillen betreffen de groeps samenstelling van de groepen. Zo hebben verticale groepen een groter aandeel kinderen van laagopgeleide ouders en kinderen met een migratieachtergrond. Het aandeel kinderen met een zorgbehoefte is ook groter in verticale groepen (ruim 21% vergeleken met bijna 14%) en dit is vooral terug te zien in een hoger aandeel kinderen met behoefte aan taalondersteuning, met gedragsproblemen en met lichamelijke of verstandelijke beperkingen. Dit zou erop kunnen duiden dat kinderen van laagopgeleide ouders, met een migratieachtergrond en kinderen met een zorgbehoefte (wellicht deels overlappend met opleidingsniveau en migratieachtergrond) later instromen in de kinderopvang²⁰ en om die reden vaker in verticale groepen (en eventueel ook peutergroepen, maar die zijn nu buiten beschouwing gelaten) zitten dan in babygroepen. Hoewel het aandeel kinderen met een vluchtelingenstatus over het algemeen erg laag is, blijken deze kinderen wel iets vaker in babygroepen te zitten en ook met een grotere variatie) dan in verticale groepen. Mogelijk duidt dit erop dat kinderen met een vluchtelingenstatus al vroeg geïndiceerd worden om in te stromen in de kinderopvang.

Met betrekking tot het pedagogisch groepsbeleid blijken er weinig verschillen tussen horizontale en verticale groepen. De stabiliteit van de groepen en het werken met opendeurenbeleid zijn erg vergelijkbaar. Ook zijn er weinig verschillen in de pedagogische omgang met kinderen in de groep. Er is een hoge mate van affectief-inclusief gedrag wat inhoudt dat kinderen regelmatig geknuffeld en getroost worden, dat kinderen worden gestimuleerd om mee te doen aan groepsactiviteiten en om lief te zijn voor elkaar. Ook het contact met ouders is vergelijkbaar tussen horizontale en verticale groepen.

²⁰ Uit recent onderzoek van SEO blijkt dat laagopgeleide moeders minder gebruik maken van kinderopvang en dat het gebruik sterk stijgt na de tweede verjaardag van het kind (Kok, Plantenga, Schwartz, en ter Weel, 2020).

3. Verklarende analyses babyopvang

Een belangrijke vraag voor beleid en praktijk is welke kenmerken van de opvang de geobserveerde kwaliteit en verschillen daarin tussen opvanggroepen kunnen verklaren. In dit hoofdstuk geven we een eerste voorzichtig antwoord op basis van de gecombineerde gegevens van 2017 - 2019.

3.1 Introductie en leeswijzer

Alle kenmerken die in dit hoofdstuk aan de orde komen, zijn steeds op hetzelfde moment gemeten. Het onderzoeksdesign is daarom als *cross-sectioneel* te karakteriseren, wat inhoudt dat alleen *samenhangen* kunnen worden onderzocht. Voor causale interpretaties zijn sterkere onderzoekdesigns nodig. Hoewel we spreken van *verklarende analyses*, gaat het feitelijk om analyses van *samenhangen* en beschouwen we de kenmerken die we aan de geobserveerde kwaliteit relateren als 'voorspellers' van kwaliteit, niet meteen als 'oorzaken'. Voor sommige van deze voorspellers ligt door de aard ervan een oorzakelijke interpretatie van een gevonden verband voor de hand. Dat geldt bijvoorbeeld voor structurele randvoorwaarden als de groepsgrootte en kind-staf ratio. Het is aannemelijker dat relatief stabiele structurele kenmerken 'oorzaken' zijn van meer variabele proceskenmerken dan omgekeerd. Structurele kenmerken staan bovendien centraal in wettelijke kwaliteitsregelingen en worden van buiten opgelegd. Voor andere voorspellers is over de causale interpretatie meer discussie mogelijk. Hoewel er theorievorming is in het kinderopvangonderzoek, en sommige aspecten al vaker in verklarend onderzoek zijn meegenomen, is de dagelijkse werkelijkheid in de kinderopvang complex, wat het lastig maakt heel specifieke verwachtingen op te stellen die vervolgens eenduidig getoetst kunnen worden om te zien of ze uitkomen of niet. Ook als bepaalde kenmerken vaker zijn onderzocht, zijn de relaties met de geobserveerde kwaliteit niet altijd consistent. Om die reden dragen de analyses waarover in dit hoofdstuk wordt gerapporteerd vooralsnog een exploratief karakter, bedoeld om ideeën voor vervolgonderzoek te verkrijgen.

Voor de verklarende analyses die in dit hoofdstuk worden gerapporteerd zijn verschillende gegevens gebruikt. Naast de observatiegegevens verzameld met de CLASS, ITERS-R en de NCKO-schalen en het aanbod van exploratie, spel en taal gebaseerd op de rapportages door medewerkers (zie Hoofdstuk 2) zijn er gegevens uit de interviews met de pedagogisch medewerkers over kenmerken van de groep en de achtergrond van de medewerkers gebruikt. Er is gestart met het bekijken van afzonderlijke relaties tussen structurele kenmerken van de groep en van de medewerkers met de proceskwaliteit aan de hand van Pearson correlaties ($r = .10$ kleine effect grootte; $r = .30$ medium effect grootte; $r = .50$ grote effect grootte volgens Cohen, 1988). Maar de relaties tussen de geobserveerde proceskwaliteit en de door middel van interviews gemeten kenmerken van de groep en de medewerkers zijn complex. De relatie van één kenmerk met de proceskwaliteit wordt mede beïnvloed door de relaties van andere

kenmerken met de proceskwaliteit. De voorspellende kenmerken zelf hebben onderlinge verbanden, waardoor de relatie van één kenmerk met de kwaliteit deels de relatie van een ander kenmerk met de kwaliteit weerspiegelt. De groepsgrootte en kind-staf ratio, bijvoorbeeld, hangen samen met de leeftijdssamenstelling van de groep. Dit vraagt om een zogenaamde *multivariate* aanpak, waarbij kenmerken die onderling samenhangen gelijktijdig als voorspellers van de kwaliteit worden geanalyseerd om zo de unieke bijdrage van elk kenmerk afzonderlijk aan de verklaring van kwaliteit vast te kunnen stellen. Voor dit doel maken we gebruik van de techniek multivariate regressieanalyse, onderdeel van het statistische analyseprogramma SPSS (versie 24) waarbij het percentage verklaarde variantie een maat voor effect grootte is ($R^2=.02$ kleine effect grootte; $R^2=.13$ medium effect grootte; $R^2=.26$ grote effect grootte).

Het hoofdstuk is als volgt opgebouwd. Na deze inleiding volgt eerst een samenvatting op hoofdlijnen van de belangrijkste uitkomsten en een beschouwing daarover. Voor de geïnteresseerde lezer wordt vervolgens op een meer technische manier de aanpak van de analyses beschreven. Daarna worden de gedetailleerde resultaten per opvangsoort beschreven aan de hand van samenvattende tabellen. In de bijlagen bij dit hoofdstuk zijn uitgebreide tabellen met de gedetailleerde analyseresultaten opgenomen. Het meer technische gedeeltelijk beschrijft de uitkomsten, maar geeft weinig interpretatie of verklaring. Dat gebeurt wel in het eerste, beschouwende deel.

3.2 Samenvatting en beschouwing van de resultaten

In het voorgaande hoofdstuk is uitgebreid ingegaan op de geobserveerde en gerapporteerde kwaliteit. Daaruit blijkt dat er verschillen zijn tussen groepen en ook tussen opvangvormen (in horizontale en verticale groepen respectievelijk) voor baby's. Het doel van dit hoofdstuk is om na te gaan in hoeverre die verschillen in kwaliteit van babyopvang samenhangen met, en mogelijk verklaard kunnen worden door, kenmerken van de pedagogisch medewerkers en de groepen waar zij mee werken. We hebben daarvoor informatie uit de interviews met pedagogisch medewerkers gebruikt. Aangezien er nog weinig bekend is over de relaties tussen structurele kenmerken en proceskwaliteit voor baby's worden alle gangbare kwaliteitskenmerken geïnccludeerd in de analyses. Een voorbeeld betreft de 'ijzeren driehoek' van de structurele kwaliteitskenmerken groepsgrootte, kind-staf ratio en opleidingsniveau van de medewerker, kenmerken die een belangrijke rol spelen in de wettelijke kwaliteitsregulering. Andere kenmerken zijn afgeleid uit het voorgaande hoofdstuk en uit eerdere onderzoeken in de baby-opvang en uit de verklarende analyses in de kinder- en peuteropvang. Zo bleken werkervaring en de werktijd in uren twee aspecten te zijn met veel spreiding rondom de gemiddelde scores. Ook om die reden zijn beide aspecten meegenomen in de analyses. Structurele aspecten die een rol spelen in het verklaren van verschillen in proceskwaliteit bleken afhankelijk van de uitkomstmaat (CLASS Infant, ITERS-R, NCKO-schalen of gerapporteerd aanbod van exploratie, spel en taal).

Kenmerken van de pedagogisch medewerker

Een hoger opleidingsniveau van de pedagogisch medewerkers blijkt samen te gaan met een hogere kwaliteit van interacties zoals gemeten met de CLASS en toont een medium effect grootte. Voor de overige kwaliteitsaspecten bleken geen significante verbanden. De ITERS-R en de NCKO-schalen zijn mogelijk wat globaler en minder specifiek in vergelijking met de CLASS Infant wat een mogelijke verklaring kan zijn voor het uitblijven van relaties met deze instrumenten. De CLASS Infant meet mogelijk specifiekere vaardigheden van medewerkers, vooral met het oog op het ondersteunen van exploratiegedrag en de vroege taalontwikkeling van kinderen, dat sterker samengaat met een hoger niveau van de vooropleiding.

Werkervaring hangt positief samen met de algemene kwaliteit zoals gemeten met de ITERS-R. Medewerkers met meer ervaring in de kinderopvang scoren hoger op een algemene maat voor kwaliteit, waaronder de kwaliteit van de ruimte en meubilering, de aanwezige materialen en het activiteitsaanbod, brede interactiekwaliteit en aandacht voor gesprekken en taal in het bijzonder en een programma. Mogelijk hebben meer ervaren krachten hun groepen beter ingericht op de behoeften van jonge kinderen en laten zij op basis van een globale beoordeling na een hele ochtend een hogere interactiekwaliteit zien. Voor dit patroon vinden we echter geen bevestiging wanneer gekeken wordt naar specifiekere aspecten van deze interacties zoals gemeten met de CLASS Infant.

Pedagogisch medewerkers die meer uren werken per week tonen een hogere kwaliteit van interacties zoals gemeten met de CLASS Infant en een meer aandacht voor taal op basis van zelfrapportages. Er kunnen verschillende verklaringen zijn voor deze bevinding. Medewerkers die meer uren werken, hebben de baby's in de groep wellicht beter leren kennen en kunnen op die manier wellicht beter afstemmen op hun behoeften. Daarnaast kan het zo zijn dat medewerkers die meer uren werken per week ook meer ervaring hebben opgedaan in het ondersteunen van de ontwikkeling van kinderen in zowel hun interacties als het aanbod van spel en activiteiten.

Tot slot: continue professionalisering blijkt een sterke voorspeller van de algemene kwaliteit. Het gaat hier om verschillende professionaliseringsvormen zoals cursussen, coaching, teamoverleg, conferentiebezoek en frequent pedagogisch-inhoudelijk overleg. Deze activiteiten dragen blijkbaar bij aan een hogere kwaliteit, zoals de inrichting van de groep, aanwezige materialen en activiteiten en een globale beoordeling van de interacties. Echter, voor de meer specifieke aspecten van de interactiekwaliteit zijn misschien andere vormen van professionalisering nodig. Met de Wet IKK moeten alle medewerkers specifieke scholing volgen gericht op het werken met baby's. Toekomstige metingen kunnen inzicht geven in de effecten van deze scholing op de kwaliteit.

Kenmerken van de groep

De groeps grootte en kind-staf ratio hangen deels met elkaar samen, maar de ratio is meer onderscheidend naar leeftijd van kinderen en is om die reden opgenomen in de analyses. De kind-staf ratio hangt niet samen met proceskwaliteit. Dit komt niet omdat het niet belangrijk is, maar omdat er relatief weinig variatie is in de ratio vanwege de sterke regulering. Het aantal kinderen in de groep of per medewerker lijkt minder belangrijk dan de leeftijd van de kinderen in de groep.

Het aandeel baby's (onder de 1 jaar) op de groep laat zowel positieve als negatieve relaties zien met proceskwaliteit. Voor de kwaliteit van de interacties blijkt het aandeel baby's een positieve uitwerking te zijn. Dit kan erop duiden dat in groepen met meer baby's er ook meer expliciete aandacht is voor deze baby's. Het eerder gevonden verschil tussen horizontale en verticale groepen wijst ook in die richting. Daarnaast kan het zo zijn dat pedagogisch medewerkers die werkzaam zijn in groepen met meer baby's, meer ervaren zijn in de omgang met baby's en beter in staat zijn om tegemoet te komen aan de behoeften van die baby's. Het kan ook zijn dat dit een expliciete voorkeur en affiniteit en dus betrokkenheid van de medewerker weerspiegelt. Het aanbod van exploratie, spel en taal voor dreumesen lijkt eerder te lijden onder de aanwezigheid van veel baby's. Een verklaring kan zijn dat in groepen met zowel baby's als dreumesen de aandacht meer uitgaat naar de jongere kinderen en dat er relatief weinig aanbod is gericht op de dreumesen en hun ontwikkelingsbehoeften met betrekking tot het exploreren van de ruimte en expliciete aandacht voor taal. Het kan zijn dat de zorg voor de jongere baby's belemmerend werkt om aandacht te besteden aan voorlezen of liedjes zingen en klappen of het uitgebreid benoemen van de objecten in de omgeving of handelingen. Ook de inrichting van de ruimte is mogelijk niet optimaal voor dreumesen om hun omgeving actief te verkennen.

Verdere verkenning van de kwaliteitsconcepten

Er zijn verschillende kwaliteitsinstrumenten gebruikt om de kwaliteit voor baby's te beoordelen. Om meer inzicht te krijgen in de kwaliteit voor baby's is daarnaast gekeken naar de samenhang tussen die verschillende aspecten, maar dan op het laagste niveau van een instrument (dus niet voor de gehele schaal, maar voor een specifiek onderdeel ervan). De resultaten laten een aantal interessante patronen zien. Zo blijkt de kwaliteit van de inrichting van de groep zoals gemeten met de ITERS-R samen te hangen met het zelf-gerapporteerde aanbod van exploratie, spel en taal. Het bevorderen van exploratie en spel bij baby's en dreumesen blijkt dus samen te hangen met fysieke mogelijkheden die de inrichting van de groep te bieden heeft. Dat heeft onder meer te maken met de oppervlakte van de binnenruimte (dus vrije beweegruimte voor kinderen), maar ook de aanwezigheid van meerdere, aparte hoekjes met bijvoorbeeld zachte materialen (kussens, zitzak of bank op kindniveau) en afgescheiden plekken voor rustig en meer actief spel, en een kindgerichte aankleding van de groep (met bijvoorbeeld platen, posters of foto's op kindniveau). Voor

dreumesen blijkt bovendien de kwaliteit van de materialen en het activiteitenaanbod zoals gemeten met de ITERS-R samen te gaan met meer aanbod van exploratie, spel en taal.

3.3 Voorspellers van de kwaliteit in de baby-opvang

De wijze waarop de kwaliteit wordt gemeten door verschillende observatieinstrumenten vertoont een kleine tot matige samenhang met elkaar ($r = .30 - .38$), zie Tabel 3.1. Ook de schalen met betrekking tot het aanbod van spel en taalactiviteiten voor baby's en dreumesen zoals gerapporteerd door de pedagogisch medewerkers vertoont een matige tot sterke onderlinge samenhang ($r = .24 - .54$). De kwaliteit zoals gemeten met de ITERS-R hangt bovendien samen met het gerapporteerde aanbod van exploratie en spel voor dreumesen ($r = .37$).

Gezien de exploratieve aard van de studie is gestart met een verkenning van de verbanden tussen proceskwaliteit en verschillende structurele kenmerken van de groep en pedagogisch medewerkers, zie eveneens Tabel 3.1. Er blijken verschillende patronen afhankelijk van het type instrument.

Zo geldt voor de CLASS Infant dat er een hogere kwaliteit is in groepen met een gunstigere kind-staf ratio (minder kinderen per medewerker) waarin medewerkers gemiddeld minder werkervaring hebben. Deze verbanden duiden op een medium effect grootte. Naarmate medewerkers meer uren werken per week en meer baby's opvangen in de groep, blijkt de kwaliteit eveneens hoger, met medium tot grote effecten. Dit kan er mogelijk op duiden dat medewerkers die meer baby's opvangen hier meer kennis in kunnen opdoen en ervaring mee hebben en beter in staat zijn om hun interacties af te stemmen op hun behoeften. Een soortgelijke verklaring kan opgaan voor het werken van meer uren per week. Medewerkers leren kinderen dan beter kennen en kunnen meer ervaring opdoen in het werken met baby's wat er mogelijk toe leidt dat zij een hogere kwaliteit van interacties kunnen leveren.

Voor de ITERS-R blijkt alleen het professionaliseringsaanbod positief samen te hangen met de kwaliteit met een matig effect. Dus medewerkers die meer mogelijkheden hebben voor cursussen, coaching, teamoverleg, conferentiebezoek en frequent pedagogisch-inhoudelijk overleg blijken een hogere globale kwaliteit te laten zien. Er blijken geen verbanden tussen de kwaliteit zoals gemeten met de NCKO-schalen en structurele kenmerken van de groep en de medewerkers.

Tabel 3.1 Correlaties tussen de proceskwaliteit en structurele kenmerken van de groep en medewerkers (N=48-62 groepen)

		2	3	4	5	6	7	8	9	10	11	12	13	14
1	CLASS Totaal	.30**	.38**	.12	.00	.02	-.04	-.08	-.26*	.42**	.16	-.23 ⁺	.40**	.20
2	ITERS-R Totaal		.37**	.19	.16	.37**	.22	.11	-.06	-.07	-.01	.18	.12	.32*
3	NCKO Totaal			.12	.07	.13	.13	.07	-.08	.07	.09	-.00	-.02	.08
4	Spel baby				.51**	.44**	.24 ⁺	.14	.04	.01	.09	-.03	.11	-.01
5	Taal baby					.34**	.54**	.08	.18	-.22	.01	.16	.12	-.08
6	Spel dreumes						.69**	-.01	.19	-.53***	.01	.27*	.02	.14
7	Taal dreumes							-.07	.36**	-.58**	.10	.27*	-.02	.09
8	Groeps grootte								.37**	-.02	-.32*	.14	-.26 ⁺	-.25 ⁺
9	Kind-staf ratio									-.49**	-.17	.38**	-.33*	-.17
10	Aandeel baby's										-.11	-.32*	.23 ⁺	.10
11	Vooropleiding											-.06	-.10	.10
12	Werkervaring												-.32*	-.20
13	Dienstverband in uren													.46**
14	Professionalisering													

Noot. + $p < .100$; * $p < .050$; ** $p < .010$; *** $p < .001$

In de volgende stap zijn multivariate regressie-analyses gedaan om te bekijken in welke mate de genoemde structurele kenmerken tezamen de verschillen in proceskwaliteit tussen horizontale en verticale groepen kunnen verklaren. In Tabel 3.2 staan de resultaten voor de drie observatieinstrumenten. Voor de CLASS Infant versie wordt 33% van de variantie verklaard door de set van kenmerken wat een substantieel en significant effect is. De belangrijkste voorspellers zijn het aandeel baby's in de groep, het aantal uren dat een medewerker werkt per week en het opleidingsniveau. In groepen met meer baby's blijkt de kwaliteit van de interacties hoger met een medium-sterk effect. Naarmate medewerkers meer uren werken, blijkt de geboden kwaliteit van interacties eveneens hoger, met een medium-sterk effect. Deze verbanden tezamen zouden er op kunnen wijzen dat medewerkers meer kennis, ervaring en ook vaardigheden op kunnen doen in het werken met baby's als er meer baby's in de groep zijn al dan niet gecombineerd met het werken van meer uren. Dit geeft medewerkers de gelegenheid om baby's beter te leren kennen en hun interacties beter af te stemmen op hun behoeften. Een hogere vooropleiding blijkt een kleine tot medium samenhang te vertonen met de interactiekwaliteit.

Voor de ITERS-R blijkt de verklaarde variantie medium tot groot, maar niet significant²¹. Hier blijken echter wel andere aspecten een rol te spelen. Zo blijkt de kwaliteit hoger wanneer medewerkers meer werkervaring hebben, met een medium effect grootte. De belangrijkste en sterkste voorspeller blijkt echter het professionaliseringsaanbod te zijn. Voor de NCKO-schalen blijkt de verklaarde variantie voor alle kenmerken tezamen klein en er blijken geen relaties van afzonderlijke aspecten.

Tabel 3.2 Resultaten van de regressieanalyses met de gegevens van de pedagogisch medewerkers.

Voorspellers van kwaliteit (N _{groepen} = 51)			
	CLASS Infant	ITERS-R	NCKO
Kind-staf ratio	.11	-.16	-.05
Percentage baby's	.41**	-.10	.08
Opleidingsniveau pm-er	.26 ⁺	-.06	.08
Werkervaring pm-er	-.01	.28 ⁺	.04
Dienstverband in uren	.38*	.00	-.08
Professionalisering	-.03	.36*	.10
<i>R</i> ²	.33	.18	.03
	F(6,45)=3.713, <i>p</i> = .004	F(6,45)=1.658, <i>p</i> = .153	F(6,41)=.319, <i>p</i> = .980

Noot. + *p* < .100; * *p* < .050; ** *p* < .010; *** *p* < .001

Voor het gerapporteerde aanbod van spel en mogelijkheden voor exploratie evenals het aanbod van taalstimulerende activiteiten worden de resultaten in Tabel 3.3 weergegeven. Voor het aanbod specifiek gericht op jonge baby's (onder 1 jaar) is de verklaarde variantie

²¹ Dit kan waarschijnlijk verklaard kan worden door de relatief lagere power vanwege de kleine steekproef

klein en niet significant. Er blijken dus voor de ondersteuning van spel en exploratie geen afzonderlijke significante voorspellers. Voor het aanbod van taalontwikkelingstimulerende activiteiten blijken medewerkers die meer uren werken per week meer aandacht te besteden aan taalstimulering in de vorm van activiteiten zoals brabbelen of actief luisteren naar muziek, ritme en klanken of zingen. Wellicht dat medewerkers die meer werken, de kinderen beter kennen en ook meer ervaring hebben opgedaan in het aanbieden van taalontwikkelingstimulerende activiteiten. Voor het aanbod voor iets oudere kinderen (dreumesen tussen 1 en 2/2,5 jaar oud) blijkt de verklaarde variantie groter (tussen de 36-39%). De belangrijkste voorspeller blijkt hier het aandeel baby's in de groep. Naarmate er meer baby's in de groep zijn, blijken de medewerkers minder mogelijkheden te bieden voor exploratie en spel voor oudere kinderen.

Tabel 3.3 Resultaten van de regressieanalyses met de gegevens van de pedagogisch medewerkers.

Pedagogisch medewerkers ($N_{\text{groepen}} = 51$)				
	Exploratie spel baby's	Taalaanbod baby's	Exploratie spel dreumesen	Taalaanbod dreumesen
Kind-staf ratio	.14	.15	-.12	.14
Percentage baby's	.06	-.16	-.57***	-.51**
Opleidingsniveau pm-er	.15	.08	-.07	.08
Werkervaring pm-er	-.01	.13	.19	.12
Dienstverband in uren	.20	.34*	.08	.13
Professionalisering	-.10	-.17	.19	.12
R^2	.04	.14	.36**	.39**
	F(6,45)=.334, $p = .915$	F(6,45)=1.205, $p = .321$	F(6,45)=4.170, $p = .002$	F(6,45)=4.839, $p = .001$

Noot. + $p < .100$; * $p < .050$; ** $p < .010$; *** $p < .001$

Om meer inzicht te krijgen in het aanbod van exploratie, spel en taal op de groep is er gekeken naar de samenhang met de geobserveerde proceskwaliteit, zie Tabel 3.4. Hieruit komt naar voren dat de kwaliteit van de inrichting van de ruimte en aanwezige meubilering samenhangt met het aanbod van exploratie en spel voor zowel baby's als dreumesen. Ook is er een positieve relatie met het taalaanbod voor baby's. Mogelijk reflecteert dit groepen die expliciet aandacht schenken aan een uitdagende en kindgerichte inrichting die mogelijkheden biedt voor exploratiegedrag van kinderen. Er blijkt eveneens een positieve relatie van het exploratie en taalaanbod voor dreumesen met het activiteitenaanbod, maar alleen voor dreumesen. Mogelijk duidt dit erop dat het activiteitenaanbod beter past bij dreumesen dan bij jonge baby's. Interessant in dit opzicht is de relatie tussen de door de medewerker gerapporteerde exploratie met het faciliteren van exploratie. Deze correlatie lijkt erop te duiden dat beide schalen op soortgelijke wijze het stimuleren van exploratie bij baby's meten. Er blijkt eveneens een verband tussen de responsieve interacties van de medewerker en de ondersteuning van exploratie en spel bij dreumesen.

Tabel 3.4 Correlaties tussen de zelf-gerapporteerde activiteitenschalen en de observatieinstrumenten

		Exploratie baby	Taalonder steuning baby	Exploratie dreumes	Taalonder steuning dreumes
ITERS-R	Ruimte en meubilering	.23 ⁺	.27*	.30*	.22
	Luisteren en praten	.11	.07	.19	.07
	Activiteiten	.11	.09	.37**	.26 ⁺
	Interactie	.11	.04	.18	-.01
CLASS	Relationeel klimaat	-.08	-.19	-.06	-.04
	Sensitiviteit	.15	-.06	.03	-.12
	Faciliteren van exploratie	.22 ⁺	.11	.05	-.02
	Vroege taalondersteuning	.04	.04	.03	.02
NCKO	Sensitieve responsiviteit	.17	.11	.24 ⁺	.22
	Respect voor autonomie	.06	.04	.10	.14
	Structureren en grenzen stellen	.00	.17	.06	.07
	Praten en uitleggen	.09	-.06	.23	.22
	Ontwikkelingsstimulering	.11	.17	.07	.10
	Begeleiden van interacties	-.01	-.01	.04	.10

Referenties

- Barros, S., Cadima, J., Bryant, D. M., Coelho, V., Pinto, A. I., Pessanha, M., & Peixoto, C. (2016). Infant child care in Portugal: Associations with structural characteristics. *Early Childhood Research Quarterly*, *37*, 118-130. doi:10.1016/j.ecresq.2016.05.003.
- Bjørnstad E. & Os, E. (2018). Quality in Norwegian childcare for toddlers using ITERS-R. *European Early Childhood Education Research Journal*, *26*, 111-127
- Fenech, M., Sweller, N., & Harrison, L. (2010). Identifying high-quality centre-based childcare using quantitative data-sets: what the numbers do and don't tell us. *International Journal of Early Years Education*, *18*(4), 283-296.
- Harms, T., Cryer, D. & Clifford, R.M. (2006). *Beoordelingsschaal Pedagogische Kwaliteit. Leer- en Leefomgeving voor baby's en dreumesen [Infant/Toddler Environment Rating Scale-Revised]* (R.E.L. de Kruif & L.W.C. Tavecchio, Vert.), Amsterdam: Uitgeverij SWP.
- Helmerhorst, K. O. W., Riksen-Walraven, J. M. A., Gevers Deynoot-Schaub, M. J. J. M., Tavecchio, L. W. C., & Fukkink, R. G. (2015). Child care quality in the Netherlands over the years: A closer look. *Early Education and Development*, *26*, 89-105. doi:10.1080/10409289.2014.948784
- Jamison, K. R., Cabell, S. Q., LoCasalel-Crouch, J., Hamre, B. K., & Pianta, R. C. (2014). CLASS-Infant: An observational measure for assessing teacher-infant interactions in center-based child care. *Early Education and Development*, *25*, 553-572. doi:10.1080/10409289.2013.822239.
- King, E. K., Pierro, R. C., Li, J., Porterfield, M. L., & Rucker, L. (2016). Classroom quality in infant and toddler classrooms: impact of age and programme type. *Early Child Development and Care*, doi: 10.1080/03004430.2015.1134521
- Kruif, R. E. K. de, Vermeer, H. J., Fukkink, R. G., Riksen-Walraven, J. M. A., Tavecchio, L. W. C., IJzendoorn, M. H., van, & Zeijl, J. van (2007). *De nationale studie pedagogische kwaliteit kinderopvang eindrapport project 0 en 1*. Amsterdam: Nederlands Consortium Kinderopvang Onderzoek (NCKO).
- Oudgenoeg-Paz, O., Leseman, P. P. M., & Volman, M. (C.) J. M. (2015). Exploration as a mediator of the relation between the attainment of motor milestones and the development of spatial cognition and spatial language. *Developmental Psychology*, *51*(9), 1241–1253. doi: [10.1037/a0039572](https://doi.org/10.1037/a0039572)
- Slot, P. L., Jepma, IJ., Muller, P., Romijn, B. R. & Leseman, P. P. M. (2017). *Kwaliteit van de Nederlandse kinderdagopvang, peuteropvang, buitenschoolse opvang en gastouderopvang - Landelijke Kwaliteitsmonitor Kinderopvang Meting 2017*. (119 p.). Sardes.
- Slot, P. L., Jepma, IJ., Muller, P., Romijn, B.R. & Leseman, P.P.M. (2018). *Kwaliteit van de Nederlandse kinderdagopvang, peuteropvang, buitenschoolse opvang en gastouderopvang - Gecombineerde metingen 2017-2018*. (176 p.). Sardes.

Colofon

Titel	Kwaliteit van de Nederlandse kinderdagopvang voor baby's op basis van de gecombineerde metingen 2017 - 2019
Auteur	Pauline Slot, IJsbrand Jepma, Paulien Muller, Bodine Romijn, Celeste Bekkering, Paul Leseman
Datum	Juli 2020
Project	LKK 2020

Landelijke Kwaliteitsmonitor Kinderopvang

 (030) 232 62 00

 info@monitorlkk.nl

 www.monitorlkk.nl



Universiteit Utrecht

Sardes 