

Motoriek kleuters
Wetenschappelijke verantwoording

Marieke op den Kamp, Jos Keuning en Herman van Boxtel



Motoriek kleuters

Wetenschappelijke verantwoording

Marieke op den Kamp
Jos Keuning
Herman van Boxtel

© Cito B.V. Arnhem (2016)

Niets uit dit werk mag zonder voorafgaande schriftelijke toestemming van Cito worden openbaar gemaakt en/of verveelvoudigd door middel van druk, fotokopie, scanning, computersoftware of andere elektronische verveelvoudiging of openbaarmaking, microfilm, geluidskopie, film- of videokopie of op welke wijze dan ook.

Inhoud

1	Inleiding	5
2	Uitgangspunten van de testconstructie	7
2.1	Meetpretentie	7
2.2	Doelgroep	7
2.3	Gebruiksdoel en functie	8
2.4	Theoretische inkadering	9
2.4.1	Het ontwikkelingspsychologisch perspectief	10
2.4.2	Het onderwijskundige perspectief	18
2.4.3	Uitgaan van normaal functioneren, niet van stoornissen	19
2.4.4	Continuïteit en stabiliteit in ontwikkeling: voorspelbaarheid over tijd?	19
2.4.5	Doorlopende leerlijnen	19
3	Beschrijving van de observatielijsten	21
3.1	Constructie van de observatielijst	21
3.2	Opbouw, afname en scoring van de observatielijst	28
3.3	Meetmodel en psychometrische eigenschappen	31
4	Normering	33
4.1	Steekproeftrekking en deelname	33
4.2	Representativiteit van de normeringssteekproef	33
4.3	Prestaties als functie van leeftijd	36
4.4	Normeringsmethodiek	39
4.5	Vaststelling van de normen	40
4.6	Nauwkeurigheid	42
5	Betrouwbaarheid	45
5.1	Generaliseerbaarheid van scores over versies	45
5.2	Generaliseerbaarheid van scores over leerkrachten	46
5.3	Accuraatheid van classificaties	46
6	Validiteit	49
6.1	Inhoudsvaliditeit	49
6.2	Begripsvaliditeit	49
7	Samenvatting	55
8	Literatuurverwijzingen	57

1 Inleiding

Deze wetenschappelijke verantwoording heeft betrekking op het digitale instrument Motoriek kleuters. Motoriek kleuters bestaat uit een observatielijst voor de leerkracht. Het instrument maakt deel uit van het Cito Volgstelsysteem primair en speciaal onderwijs.

Het doel van Motoriek kleuters is het in kaart brengen en volgen van het motorisch functioneren van leerlingen in groep 1 en 2.

Motoriek kleuters is gebaseerd op de gedachte dat leerkrachten die werkzaam zijn in de kleutergroepen praktijkdeskundigen zijn, die in staat mogen worden geacht om het gedrag en het functioneren van leerlingen te observeren en te beoordelen. Het instrument biedt hierbij steun in de vorm van een aantal uitspraken of stellingen (observatiepunten) waarbij de leerkracht aan kan geven in welke mate deze van toepassing zijn op de leerling. De observatielijst is gestandaardiseerd: alle leerlingen worden altijd beoordeeld aan de hand van dezelfde observatiepunten. Het grote voordeel daarvan is dat de gegevens van leerlingen op deze manier, via de normering, vergeleken kunnen worden met die van een grote en representatieve referentiegroep van leeftijdsgenoten.

In kaart brengen van het functioneren betekent, net als bij andere instrumenten in het Cito Volgstelsysteem primair en speciaal onderwijs, dat er uitspraken worden gedaan over het functioneren van de leerling, waarbij we veronderstellen dat aan dat functioneren (de ontwikkeling van) een bepaalde vaardigheid ten grondslag ligt.

De administratie- en afnameomgeving is gebaseerd op het Computerprogramma LOVS. Dit betekent dat de observatielijst digitaal wordt afgenomen (ingevuld) door gebruik te maken van het Computerprogramma LOVS. Ook de planning, scoring en rapportage gebeuren via dit programma, dat voor alle toetsen van het Cito Volgstelsysteem primair en speciaal onderwijs wordt gebruikt.

Samen met de inhoud van de observatielijst, de handleiding (Cito, 2015) en het Computerprogramma LOVS geeft deze verantwoording alle informatie die nodig is voor een snelle en efficiënte beoordeling van de kwaliteit van Motoriek kleuters. Het genoemde materiaal maakt een beoordeling van het instrument mogelijk op de volgende aspecten:

- Uitgangspunten van de testconstructie
- De kwaliteit van het testmateriaal
- De kwaliteit van de handleiding
- Normen
- Betrouwbaarheid
- Validiteit

Het laatstgenoemde aspect betreft alleen de begripsvaliditeit en geen criteriumvaliditeit. Omdat Motoriek kleuters niet bedoeld is voor 'voorspellend gebruik' is criteriumvaliditeit niet van toepassing.

Het voorliggende document heeft met name betrekking op de uitgangspunten van de constructie (hoofdstuk 2 en 3), de normen (hoofdstuk 4), de betrouwbaarheid en meetnauwkeurigheid (hoofdstuk 5) en de begripsvaliditeit (hoofdstuk 6) van Motoriek kleuters. De kwaliteit van het testmateriaal en de handleiding is te bepalen door kennis te nemen van de desbetreffende onderdelen en het Computerprogramma LOVS.

2 Uitgangspunten van de testconstructie

2.1 Meetpretentie

Motoriek kleuters geeft een globale indicatie van het niveau van motorisch functioneren. Daarbij worden twee dimensies onderscheiden.

- (1) *Grove motoriek*: het sequentieel en simultaan bewegen van grote delen van het lichaam op temporeel en ruimtelijk gecoördineerde wijze.
- (2) *Fijne motoriek*: het gecoördineerd bewegen van individuele lichaamssegmenten, in het bijzonder het gebruik van handen en vingers in het manipuleren van objecten.

De meetpretentie van Motoriek kleuters heeft geen betrekking op het cognitieve functioneren. Wel stelt Motoriek kleuters de leerkracht in staat om de samenhang tussen (aspecten van) motorisch functioneren en het verloop van onderwijs- en leerprocessen in kaart te brengen. De resultaten op de observatielijst kunnen gelegd worden naast de uitkomsten van toetsen die een indruk geven van de cognitieve mogelijkheden en leerprestaties van de leerlingen. Binnen het Cito Volgsysteem primair en speciaal onderwijs kan men voor de doelgroep kleuters hiervoor de toetsen Taal voor kleuters en Rekenen voor kleuters, en het screeningsinstrument beginnende geletterdheid gebruiken. Daarnaast kan men zicht krijgen op het sociaal-emotionele functioneren door VISEON voor kleuters af te nemen.

2.2 Doelgroep

Motoriek kleuters is genormeerd voor leerlingen van groep 1 en groep 2 in het reguliere basisonderwijs en is dan ook voor de leerkrachten van deze groepen bestemd.

Het invullen van de observatielijst kost in de regel niet meer dan tien minuten per leerling. Het instrument heeft een leeftijdsnormering. Dat betekent dat het motorisch functioneren van een leerling vergeleken kan worden met het motorisch functioneren van leeftijdsgenoten. Het instrument kan op ieder moment in het jaar ingevuld worden.

Normering

Het normeringsonderzoek heeft plaatsgevonden bij leerlingen van de groepen 1 en 2 van het basisonderwijs, inclusief de leerlingen die nog een jaar in het kleuteronderwijs verblijven terwijl ze op basis van hun leeftijd al hadden kunnen doorstromen naar groep 3.

Omdat het motorische functioneren sterk afhankelijk is van de kalenderleeftijd, hebben we gekozen voor een normering naar kalenderleeftijd. Daarbij maken we onderscheid tussen zes normgroepen (viermaandsgroepen):

- Normgroep 1: 4.0 t/m 4.3 jaar (4 jaar en 0 maanden tot en met 4 jaar en 3 maanden oud)
- Normgroep 2: 4.4 t/m 4.7 jaar
- Normgroep 3: 4.8 t/m 4.11 jaar
- Normgroep 4: 5.0 t/m 5.3 jaar
- Normgroep 5: 5.4 t/m 5.7 jaar
- Normgroep 6: 5.8 jaar of ouder

Dit betekent dat de normgroepen anders zijn dan bij andere instrumenten van het Cito Volgsysteem primair en speciaal onderwijs voor leerlingen van groep 1 en 2 (bijvoorbeeld VISEON voor kleuters), waarbij we de jaargroep als basis voor de normering hebben gebruikt.

Beperkingen

In de normering wordt de leerling vergeleken met andere leerlingen in de betreffende leeftijdsgroep in het reguliere basisonderwijs. Bij de normering zijn geen leerlingen uit het speciaal onderwijs of het speciaal basisonderwijs betrokken. Motoriek kleuters kan wel ingezet worden binnen het speciaal onderwijs of het speciaal basisonderwijs. Bij het interpreteren van de normering dienen leerkrachten zich wel te realiseren dat ze leerlingen vergelijken met leerlingen van dezelfde kalenderleeftijd in het reguliere basisonderwijs. Het is niet bekend of even oude kinderen binnen het reguliere basisonderwijs, het speciaal onderwijs en het speciaal basisonderwijs systematisch van elkaar verschillen in hun motorische functioneren zoals dit wordt beoordeeld met Motoriek kleuters.

Bij de ontwikkeling van Motoriek kleuters was het belangrijkste uitgangspunt dat leerkrachten goed in staat zijn het motorisch functioneren van hun leerlingen te beoordelen aan de hand van het concrete en direct waarneembare gedrag van de leerling op school.

Motoriek kleuters geeft een *globaal* beeld van het motorische functioneren. Wanneer men een nauwkeuriger en gedetailleerder beeld van het functioneren en de ontwikkeling van een kind wenselijk acht, is vervolgonderzoek noodzakelijk. Het kan nodig zijn dat ouders hiervoor met behulp van de huisarts een revalidatie- of kinderarts, kinderfysiotherapeut, ergotherapeut, motorisch remedial teacher of neuroloog inschakelen.

2.3 Gebruiksdoel en functie

Het doel van Motoriek kleuters is het in kaart brengen van het motorisch functioneren van leerlingen in groep 1 en 2 van het basisonderwijs. Het instrument biedt hierbij steun in de vorm van een aantal uitspraken of stellingen (observatiepunten) waarbij de leerkracht aangeeft in welke mate deze van toepassing zijn op de leerling. De lijst is gestandaardiseerd: alle leerlingen worden altijd beoordeeld aan de hand van dezelfde observatiepunten. Het grote voordeel daarvan is dat resultaten van leerlingen op deze manier, via de normering, vergeleken kunnen worden met de gegevens die verzameld zijn bij een grote en representatieve referentiegroep van leeftijdsgenoten. In kaart brengen betekent hier, net als bij andere instrumenten in het Cito Volgsysteem primair en speciaal onderwijs, dat er uitspraken worden gedaan over het functioneren van de leerling.

Door middel van een signaal geven we aan hoe een leerling functioneert vergeleken met leeftijdsgenoten. Daar kan dan vervolgens een passend begeleidingstraject bij gezocht worden. Voor leerlingen die zeer laag scoren in vergelijking met leeftijdsgenoten kan dit in de vorm van extra aandacht en stimulering (zie de handleiding voor de leerkracht). Hiervoor kan de leerkracht een plan van aanpak maken. Indien een school bij een zeer zwak functioneren denkt niet de hulp te kunnen bieden die het kind nodig heeft, kan men de leerling voor nader onderzoek doorverwijzen naar een expert. Dit betekent echter niet dat de leerling – alleen op grond van deze lage score – als risicoleerling moet worden aangemerkt. Deze observatielijst is nadrukkelijk niet bedoeld als screenend of diagnostisch instrument.

Niveaubepaling

De ruwe scores op de twee onderscheiden dimensies grove en fijne motoriek vormen de basis voor de niveaubepaling. Op zichzelf zeggen ruwe scores niets. Ze krijgen pas betekenis door ze te vergelijken met de ruwe scores in een representatieve referentiegroep (i.e. door relatieve normering). Bij dit instrument doen we dit door percentielscores toe te kennen aan de ruwe scores op basis van de cumulatieve verdeling van die ruwe scores in de referentiegroep. Vervolgens hebben we deze percentielscores in drie signaalgroepen verdeeld: niet afwijkend (N), afwijkend (A) en zeer afwijkend (Z) van de referentiegroep.

De signalen zijn op onderstaande wijze toegekend aan de percentielscores:

Signaal	Percentielscore	Interpretatie
niet afwijkend (N)	hoger dan 25	De score van de leerling is vergelijkbaar met de scores van de 75% hoogst scorende leeftijdgenoten uit de normgroep
afwijkend (A)	25 of lager, maar hoger dan 10	De score van de leerling is vergelijkbaar met de scores van de 10 tot 25% laagst scorende leeftijdgenoten uit de normgroep
zeer afwijkend (Z)	10 of lager	De score van de leerling is vergelijkbaar met de scores van de 10% laagst scorende leeftijdgenoten uit de normgroep

Om een overzicht te krijgen van de resultaten van alle leerlingen in een groep, kan de leerkracht in het Computerprogramma LOVS een *Categorieënanalyse groep* opvragen. Meer informatie hierover is te vinden in de handleiding voor de leerkracht.

Het advies is om de observatielijst één keer per jaar in te vullen voor alle kinderen en twee keer voor kinderen die afwijkend of zeer afwijkend scores. Door de observatielijst één keer per jaar in te vullen, kan de leerkracht het motorisch functioneren van kinderen gedurende de kleuterperiode volgen. Dit maakt het mogelijk om bij een relatief zwak motorisch functioneren tijdig in te grijpen. Leerkrachten kunnen er ook voor kiezen om alleen over die kinderen de observatielijst in te vullen van wie ze uit hun dagelijkse observaties gemerkt hebben dat ze wat achterlopen op andere kinderen.

Bij de meeste toetsen in het Cito Volgsysteem primair en speciaal onderwijs kan men voor een bepaald kind of een groep kinderen de ontwikkeling van een vaardigheid rechtstreeks afleiden uit de zogeheten vaardigheidsscore. Dit komt omdat het gehanteerde meetmodel (in de regel het OPLM, het One Parameter Logistic Model) het mogelijk maakt om de scores van een leerling op verschillende toetsen, op verschillende momenten afgenomen, onderling te vergelijken. De ruwe scores op de toetsen worden daartoe getransformeerd in scores op één uni-dimensionele vaardigheidsschaal. Wanneer een leerling op een bepaald moment een duidelijk hogere vaardigheidsscore behaalt dan op een eerder afnamemoment, dan kan men aan de hand daarvan rechtstreeks concluderen dat de vaardigheid van de leerling is toegenomen. Bij Motoriek kleuters hebben we, vanwege het karakter van het instrument, zo'n vaardigheidsschaal niet geconstrueerd. We zullen dit verderop in deze verantwoording nog nader toelichten. De implicatie hiervan is dat de gebruiker het functioneren van een kind alleen kan volgen aan de hand van de signalen (N, A en Z). Heeft een leerling bijvoorbeeld een jaar geleden nog een ruwe score gekregen die resulteerde in een signaal A en krijgt dezelfde leerling nu op basis van zijn ruwe score een signaal Z, dan kan men in relatieve zin (vergeleken met leeftijdgenoten) constateren dat zijn niveau van functioneren erop achteruit is gegaan. Ook al is aan de hoogte van de ruwe scores wellicht te zien dat er sprake is van enige vooruitgang, doorslaggevend is de vraag of de leerling in zijn vooruitgang gelijke tred heeft gehouden met de 'normale' vaardigheidstoename van leeftijdgenoten. Kennelijk is dat in dit voorbeeld niet het geval. In het signaal Z komt naar voren dat het huidige niveau van functioneren in vergelijking met leeftijdgenoten nu als 'zeer afwijkend' moet worden geïnterpreteerd, terwijl het bij de vorige afname nog als 'afwijkend' kon worden gekarakteriseerd. Het relatief toch al zwakke functioneren is een jaar later relatief nog zwakker geworden. Mocht er sprake zijn van een ontwikkelingsachterstand, dan is deze in de achterliggende periode alleen maar groter geworden.

2.4 Theoretische inkadering

Het is van belang dat leerkrachten in het primair onderwijs aandacht besteden aan het motorisch functioneren van hun leerlingen. In de eerste plaats omdat de kerndoelen voor het primair onderwijs (Greven & Letschert, 2006) duidelijk maken dat het onderwijs ook gericht moet zijn op bewegingsonderwijs en beeldende vorming. In de uitwerking van de kerndoelen in tussendoelen en leerlijnen (SLO, 2008) staat

vermeld welke inhouden en activiteiten en welke materialen en technieken aan bod moeten komen in groep 1 en 2. Deze hebben ook betrekking op het grof- en fijnmotorisch functioneren van de leerlingen. In de tweede plaats is het van belang om aandacht te besteden aan het motorisch functioneren, omdat er sprake is van een duidelijke wisselwerking tussen de motorische ontwikkeling van een leerling en zijn ontwikkeling op andere terreinen. Problemen op het motorische vlak kunnen het leren op cognitief gebied negatief beïnvloeden. Maar ook is het mogelijk dat leerproblemen een storende invloed hebben op de ontwikkeling van een leerling op het motorische vlak.

Bij de constructie van Motoriek kleuters stond een aantal theoretische overwegingen centraal. Op de eerste plaats zijn we vertrokken vanuit *een ontwikkelingspsychologisch perspectief*. wat weten we over ontwikkeling, welke factoren en processen liggen eraan ten grondslag en wat zijn de belangrijkste ontwikkelingsmijlpalen? Vertrekkend vanuit algemene ontwikkelingsprincipes konden we ons baseren op de meer specifieke ontwikkelingspsychologische theorieën over de motoriek waarbij we de dynamische systeemtheorie centraal hebben gesteld. Daarover verderop meer. Op de tweede plaats hebben we ons laten leiden door *een onderwijskundig perspectief*. We hebben ons gebaseerd op de doelen die algemeen gehanteerd worden in het onderwijs aan leerlingen in groep 1 en 2. De leerdoelen sluiten aan bij de normale ontwikkelingsmogelijkheden van deze kinderen in termen van de ontwikkelingstaken waar zij voor staan. Op de derde plaats hebben we ons bij de ontwikkeling van de observatielijst laten leiden door de *normale ontwikkeling* en door functioneren *dat als normaal wordt beschouwd*, en dus niet door problematische ontwikkeling, respectievelijk problematisch gedrag of functioneren. Dat wil niet zeggen dat er geen relatie zou bestaan tussen de score op de observatielijst Motoriek kleuters en scores op lijsten die vooral bedoeld zijn om problematisch functioneren in kaart te brengen. De ontwikkelingspsychologische theorievorming met betrekking tot motoriek zoals we die in paragraaf 2.4.1 op hoofdlijnen zullen schetsen heeft immers niet alleen iets te zeggen over de manier waarop ontwikkeling zich normaliter voltrekt, maar ook over de processen en condities die een problematische ontwikkeling in de hand werken en versterken. Op de vierde plaats hebben we ons afgevraagd in hoeverre deze processen en factoren leiden tot een zekere *continuïteit en stabiliteit*, en daarmee ook tot een bepaalde mate van *voorspelbaarheid over tijd*. Hoewel we hier geen empirisch onderzoek naar hebben gedaan (en we dus voor dit instrument ook geen pretenties hebben met betrekking tot eventuele predictieve kwaliteiten), zullen we hier wel iets over zeggen. Op elk van deze overwegingen gaan we in de volgende paragrafen wat uitgebreider in.

2.4.1 Het ontwikkelingspsychologisch perspectief

Algemene ontwikkelingspsychologische principes

Een groot deel van de ontwikkelingspsychologische theorievorming houdt zich bezig met de vraag of en hoe ontwikkeling zich voltrekt onder invloed van genetisch-biologische factoren (*nature*) enerzijds en de sociale en culturele omgeving (*nurture*) anderzijds. Daarbij gaat het vooral om de vraag hoe *nature* en *nurture* samen, in hun onderlinge samenspel, de ontwikkeling bepalen. De ontwikkelingspsychologische principes en 'modellen' die hierbij centraal staan gelden als algemene theoretische achtergrond. Sameroff (2010) noemt de volgende:

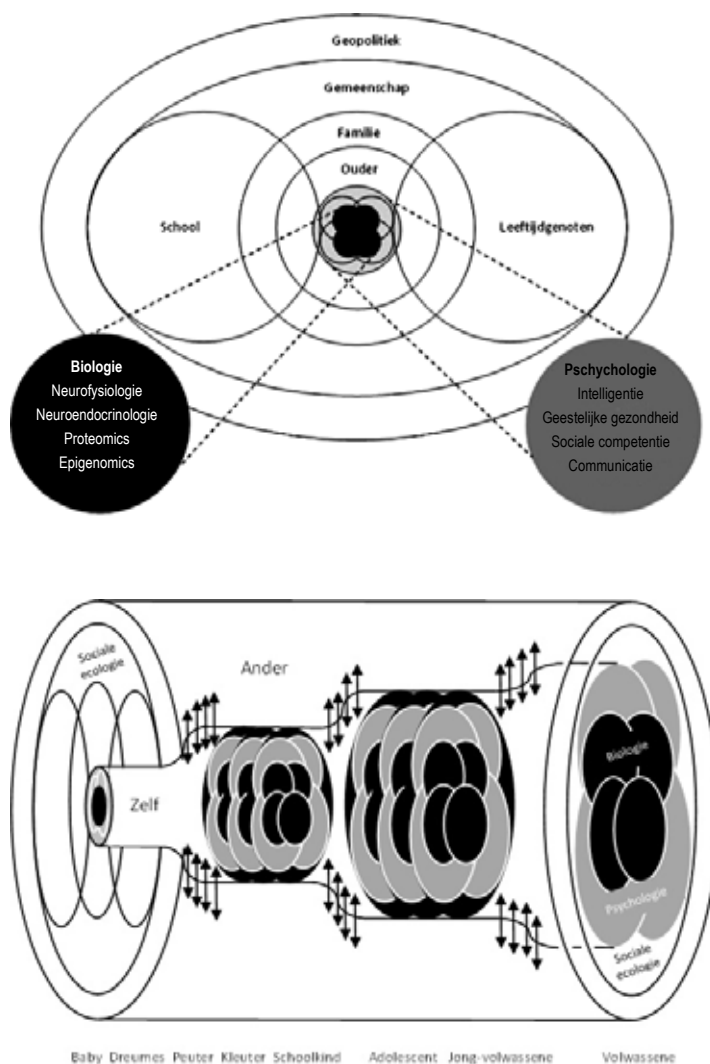
- *Het individueel persoonlijk veranderingsmodel (1)*. Het model dat men hier kiest, geeft aan hoe men aankijkt tegen persoonlijke verandering. Veronderstelt men de aanwezigheid van relatief stabiele kenmerken die nauwelijks aan verandering onderhevig zijn, is er sprake van (kwantificeerbare) groei of voltrekt de verandering zich in stadia die steeds resulteren in nieuw en kwalitatief verschillend gedrag?
- *Het contextueel model (2)*. Gedrag en ontwikkeling kunnen niet los worden gezien van de sociale context. In de klassieke analyse van Bronfenbrenner kan men in de sociale context verschillende lagen onderscheiden (micro-, meso- en macrosysteem, exo- en chronosysteem). Kenmerkend is dat het zich ontwikkelende individu met de leeftijd in steeds breder worden sociale systemen functioneert (ouders – gezin – school – peergroep et cetera).
- *Het regulatief model (3)*. Hierbij wordt het zich ontwikkelende kind opgevat vanuit een dynamische systeembenadering, waarin (in afnemende mate) door anderen gereguleerde en (in toenemende mate) door het kind zelf gereguleerde elementen zijn te onderscheiden. Sameroff bespreekt in dit verband het belang van de *transactionele relatie* tussen het kind en zijn groeiende zelfbewustzijn enerzijds en

betekenisvolle anderen (zoals ouders, leidsters in het kinderdagverblijf en op de peuterspeelzaal en leerkrachten op school) anderszijds.

- *Het representatiemodel (4)*. Ten slotte is het van belang aandacht te schenken aan de interne representatie van de buitenwereld en van de ervaringen die het kind opdoet. Afhankelijk van de specifieke theorie spreken we hier over cognitieve schema's over de buitenwereld en interne werkmodellen. Andere termen die in dit opzicht relevant zijn, zijn zelfbeeld, competentiebeleving en zelfwaardering.

Sameroff heeft bovenstaande vier invalshoeken samengevoegd tot één geïntegreerde ontwikkelingspsychologische theorie die in cross-sectioneel en longitudinaal perspectief is af te beelden zoals in figuur 2.2.

Figuur 2.2 Geïntegreerde ontwikkelingspsychologische theorie in cross-sectioneel (boven) en longitudinaal (onder) perspectief (afbeeldingen ontleend aan Sameroff, 2010)



Een geïntegreerde theorie als deze is vooral verhelderend als men tot doel heeft de belangrijkste factoren en processen in beeld te krijgen die bepalen *hoe* ontwikkeling zich voltrekt. Bij zeer jonge kinderen is voor deze ontwikkeling vooral het hoge tempo kenmerkend. Hun gedrag en functioneren wordt in hoge mate bepaald door de snelle en massieve veranderingen die zich voltrekken in het brein (voor een overzicht zie

bijvoorbeeld Johnson, 2011). Deze en andere genetisch gestuurde biologische veranderingen vormen de basis voor psychologische ontwikkelingen binnen verschillende domeinen (bijvoorbeeld intelligentie, sociale competentie, identiteit, motoriek), waarbij het kenmerkend is dat deze zich voltrekken in steeds breder wordende sociale contexten (gezin, kinderdagverblijf, peuterspeelzaal, school, buurt, maatschappij). De nadruk op processen en verklaringen kan weliswaar verhelderen waarom kinderen – gemiddeld gesproken – binnen een specifieke culturele en historisch bepaalde context in staat zijn tot bepaald gedrag, toch biedt een uitsluitend procesmatige ‘formatieve’ aanpak weinig aanknopingspunten als men zicht wil krijgen op de typische ontwikkelingsmijlpalen die men voor een bepaalde leeftijdsrange zou willen beschrijven. Het is daarom handig om nog een vijfde invalshoek toe te voegen, namelijk het ‘summatieve’ perspectief van de ontwikkelingsopgaven die de basis vormen voor deze mijlpalen. Het zijn juist deze ontwikkelingsopgaven en -mijlpalen die aan de basis liggen van leerdoelen (tussendoelen) zoals deze voor het motorische domein zijn te formuleren.

Motorische ontwikkeling

Na deze beschrijving van algemene ontwikkelingspsychologische principes, gaan we nu specifiek in op de theorievorming met betrekking tot het motorisch functioneren en de motorische ontwikkeling.

Motorische ontwikkeling kan gedefinieerd worden als ‘de verandering in bewegingsgedrag in de tijd én de processen die daaraan ten grondslag liggen’ (Clark & Whitall, 1989). In de nadruk op ontwikkelingsprocessen sluit deze definitie goed aan bij Sameroffs hoofdzakelijk procesmatige benadering van ontwikkeling. Niettemin kenmerken de vroegste theorieën over motorische ontwikkeling zich hoofdzakelijk door de nadruk die ze leggen op ontwikkelingsmijlpalen. Dit denken in termen van mijlpalen komt voort uit de neurale rijpingstheorie (Gesell & Amatruda, 1947). Hierbij wordt ervan uitgegaan dat de motorische ontwikkeling een weergave is van het uitrollen van genetische informatie (*nature*). In het algemeen geven de mijlpalen een goed beeld van de motorische ontwikkeling van jonge kinderen. Er wordt hierbij echter voorbijgegaan aan enkele belangrijke elementen:

- Er is een grote spreiding in het bereiken van mijlpalen. Sommige kinderen leren bijvoorbeeld veel eerder of later lopen dan andere kinderen. Dit hoeft geen probleem te zijn. Er kan dan nog steeds sprake zijn van een normale ontwikkeling. Zo geeft Poot (zoals geciteerd in Singer & Kleerekoper, 2009) aan dat acht maanden verschil in leren lopen normaal is.
- De ontwikkeling van motoriek is niet lineair, maar verloopt met sprongen.
- Er zijn ook kinderen die mijlpalen overslaan of in een andere volgorde bereiken. Er is dus geen sprake van een strikt sequentiële ontwikkeling (Largo, Kunda, Thum-Hohenstein, 1993).
- Er wordt geen rekening gehouden met bijvoorbeeld omgevingsfactoren, zoals de mate van training die het kind ervaart of het culturele belang van de taken waarvoor een bepaalde motorische vaardigheid nodig is.

Kortom, er is meer nodig om te komen tot een adequate beschrijving en verklaring van de motorische ontwikkeling. Volgens de dynamische systeemtheorie biedt een eenzijdig accent op (de rijping van) het centrale zenuwstelsel onvoldoende aanknopingspunten en komt motorische ontwikkeling tot stand door de dynamische interactie tussen individu, taak en omgeving (Piek, 2006; Thelen & Smith, 1994). Nieuwe vaardigheden als rollen en kruipen ontstaan door de samenwerking van de onderdelen in het lichaam (zoals spieren en gewrichten), maar een voorwaarde is wel dat de juiste omgeving- en taakomstandigheden aanwezig zijn, zoals bijvoorbeeld de juiste ondergrond. Nieuwe motorische vaardigheden ontstaan dus spontaan uit de interactie van het kind met de omgeving en zijn niet van tevoren in die concrete vorm in het systeem aanwezig: het systeem organiseert zich vanzelf (Empelen, Nijhuis-van der Sanden & Hartman, 2000; Thelen & Smith, 1994). De omgeving staat centraal bij de dynamische systeembenadering. Hierin onderscheidt deze benadering zich duidelijk van de rijpingsbenadering waarin motorische ontwikkeling wordt beschouwd als een neuraal proces dat losstaat van omgevingsinvloeden (Empelen, Nijhuis-van der Sanden & Hartman, 2000). De ontwikkeling van motorische vaardigheden wordt beïnvloed door de aanleg van een kind, de groei en ontwikkeling van het lichaam, de gezondheid, prikkels uit de omgeving en de mate waarin een kind ermee heeft geoefend. Het motorisch functioneren is verweven met andere ontwikkelingsgebieden, zoals de sociaal-emotionele en de cognitieve ontwikkeling. Voelt een kind zich bijvoorbeeld op zijn gemak in de groep? Kan het zich langere tijd op iets concentreren? Wordt het kind snel boos of angstig? Heeft een kind een sterke wil om iets te leren of geeft het snel op? Als het op de ene

manier niet lukt, kan het kind dan een andere manier bedenken? Begrijpt het wat de bedoeling is bij bewegingsactiviteiten, zoals dansen?

Wat verstaan we eigenlijk onder (adequaat) motorisch functioneren?

Om gecontroleerde bewegingen uit te kunnen voeren heeft het kind een goede motoriek nodig. Daarvan is sprake als het kind beschikt over de benodigde motorische vaardigheden. Bovendien moet het kind die vaardigheden kunnen toepassen in situaties die daarom vragen. Voorbeelden van dergelijke motorische vaardigheden zijn rennen, springen, vangen, rollen, bouwen, plakken en tekenen. Ter beheersing ervan moet het beschikken over voldoende rompbalans, krachtregulatie, timing, coördinatie van spiergroepen en ledematen, evenwicht, bewegingssnelheid en oog-hand/voetcoördinatie. Motorische vaardigheden, zoals springen moet het kind in diverse situaties kunnen toepassen. Een kind dat in de gymzaal kan 'slootje springen' door twee matten te overbruggen is daarmee nog niet in staat om bij het buitenspelen tussen vijftientig enthousiaste kleuters over een grote plas water te springen. Daarvoor zijn ook durf, een goede inschatting van de eigen verworvenheden en een adequate dosis zelfvertrouwen nodig. Kortom, beheersing van motorische vaardigheden is één ding, adequaat motorisch functioneren een ander. Voor het motorisch functioneren speelt het ook een grote rol hoe uitdagend bewegingssituaties zijn en of de plek om te bewegen en te spelen wel veilig genoeg is.

Er zijn grote verschillen in hoe kinderen in groep 1 en 2 motorisch functioneren. Sommige kinderen zijn motorisch buitengewoon getalenteerd en ontwikkelen zich heel snel, andere kinderen functioneren gemiddeld en er zijn ook kinderen waarbij de ontwikkeling vergeleken met leeftijdgenoten uitgesproken langzaam verloopt.

Bij sommige kinderen verloopt de ontwikkeling essentieel anders, bijvoorbeeld door een lichamelijke handicap of motorische stoornissen (zoals cerebrale parese). Voor deze kinderen is de ontwikkelde observatielijst Motoriek kleuters niet geschikt. Er zullen namelijk verschillende observatiepunten niet van toepassing zijn. Kinderen die heel onhandig zijn en grote problemen hebben in de coördinatie, kunnen DCD hebben. DCD staat voor *Developmental Coordination Disorder*, in het Nederlands vertaald als 'stoornis in de ontwikkeling van de coördinatie van bewegingen'. DCD is een verzamelnaam voor een aantal kenmerken van (licht) gestoorde motorische functies, zoals een lage spierspanning, een grote bewegings-onrust, coördinatieproblemen of problemen met fijnmotorische vaardigheden. Deze problemen kunnen apart voorkomen, maar veel vaker treden ze in combinatie op. DCD lijkt voor te komen bij 5 tot 10% van de schoolgaande kinderen. Vaak zijn het jongens. In diverse onderzoeken worden verhoudingen genoemd van 4 meisjes tegenover 10 jongens (Balansdigitaal.nl).

Voor diagnostiek op het terrein van DCD is het instrument Motoriek kleuters niet geschikt. Daarvoor zijn gespecialiseerde instrumenten nodig die worden ingezet door of onder supervisie van een revalidatie- of kinderarts, kindfysiotherapeut of neuroloog.

Motorische domeinen of taken

Er bestaan diverse visies over de onderverdeling van de motoriek in verschillende domeinen of taken, waarvan de indeling in grove of grote motoriek en fijne of kleine motoriek er één is. Deze indeling komt regelmatig terug in de literatuur, in observatie-instrumenten en verschillende VVE-methodes die de motorische ontwikkeling of het motorisch functioneren in kaart brengen. Maar er zijn ook domein-beschrijvingen waarbij geen onderscheid gemaakt wordt in grove of fijne motoriek, maar bijvoorbeeld wel in de onderliggende motorische capaciteiten, zoals rompbalans, krachtregulatie, timing, coördinatie van spiergroepen en ledematen, evenwicht, bewegingssnelheid en oog-hand/voetcoördinatie. En er zijn domeinbeschrijvingen waarbij grove en fijne motoriek beschreven worden naast enkele van bovenstaande motorische capaciteiten. Zo onderscheidt Noordstar (2009) de volgende domeinen of deelgebieden: rompbalans en spierkracht, evenwicht, oog-handcoördinatie en oog-voetcoördinatie, fijne motoriek en ruimtelijke oriëntatie. Noordstar (2009) geeft hierbij aan dat er geen overeenstemming bestaat over de vraag of ruimtelijke oriëntatie binnen de motoriek valt. Sommigen vinden dat de ruimtelijke oriëntatie niet bij de motoriek hoort, terwijl anderen deze juist beschouwen als een essentieel aspect van alle motorisch functioneren.

Op consultatiebureaus wordt het Van Wiechenschema uit het Van Wiechen ontwikkelingsonderzoek gebruikt. Het stapelen van blokjes of het staan op één been zijn bijvoorbeeld voor veel ouders herkenbare taken van respectievelijk fijne en grove motoriek uit het Van Wiechenschema. Naast het Van Wiechen ontwikkelingsonderzoek worden de domeinen grove en fijne motoriek ook onderscheiden in bijvoorbeeld, de Baecke-Fassaert Motoriektest voor kleuters en Kijk. Er zijn ook instrumenten die zich richten op één domein, zoals KOEK, waarmee fijnmotorische taken worden geobserveerd.

In veel VVE-methodes voor groep 1 en 2 worden de domeinen grove en of fijne motoriek onderscheiden, zoals in Peuter- en Kleuterplein. Soms wordt er in VVE-methodes in plaats van over grove motoriek gesproken over 'beweging' of 'bewegingsonderwijs', zoals bij Piramide, of worden met grove motoriek enkele grofmotorische vaardigheden bedoeld en worden de onderliggende motorische capaciteiten, zoals coördinatie of evenwicht apart benoemd.

In de tussendoelen voor groep 1 en 2 (SLO, 2008) zijn motorische domeinen voor groep 1 en 2 te herkennen in de leergebieden Bewegingsonderwijs en Kunstzinnige oriëntatie. Onder Bewegingsonderwijs vallen de onderdelen die volgens sommige definities gerekend kunnen worden tot de grove motoriek, waaronder: balanceren, stoeien, tikspelen, rijden, rollen, klimmen, springen, zwaaien, dansen en balspelen. Ook aspecten van het onderdeel 'Beweging', dat valt onder het leergebied Kunstzinnige oriëntatie, kunnen hiertoe gerekend worden. Het gaat hierbij over het verplaatsen in de ruimte en dansen op de plaats, met het hele lichaam of met afzonderlijke lichaamsdelen. Binnen het leergebied Kunstzinnige oriëntatie, onderdeel beeldende vorming, worden allerlei materialen en technieken genoemd die aangeboden moeten worden, zoals tekenen met potlood. Om deze materialen en technieken te hanteren zijn fijnmotorische vaardigheden nodig. De kwaliteit van deze vaardigheden wordt echter binnen het onderdeel 'Beeldende vorming' niet benoemd.

Bij de ontwikkeling van dit instrument hebben we ervoor gekozen om het motorisch functioneren van kinderen meetbaar te maken in de vorm van observaties van concrete, veel voorkomende dagelijkse motorische taken. Het bovenstaande overziende, reikt de literatuur ons voor de groepering van deze taken (observatiepunten) geen eenduidige indeling aan. De meest passende indeling lijkt daarom een indeling die goed aansluit bij de praktijk (de leerdoelen van het onderwijs, VVE-methoden, de gangbare instrumenten die consultatiebureaus hanteren). De indeling in grove en fijne motoriek, waarbij gericht gekeken wordt naar grof- en fijnmotorische taken, blijkt het beste bij dit pragmatische criterium te passen. Deze indeling heeft diverse voordelen. Door slechts twee domeinen of dimensies te onderscheiden, kan het aantal observatiepunten beperkt worden om tot een betrouwbare en valide meting te kunnen komen. Bovendien is de indeling herkenbaar voor leerkrachten en ook voor veel ouders. Er hoeven in de regel ook geen aparte activiteiten te worden uitgevoerd om de observatielijst in te kunnen vullen. Ten slotte, en dat is niet onbelangrijk in verband met de begripsvaliditeit (zie hoofdstuk 5), vonden we in de analyse van de verzamelde gegevens over de observatiepunten geen enkele aanleiding om een andere (bijvoorbeeld meer gedetailleerde) indeling aan te houden.

Grove motoriek

Over wat het domein 'grove motoriek' omvat, en hoe de ontwikkeling van de grove motoriek verloopt, bestaan diverse opvattingen. Zo vermelden Bilo & Voorhoeve (2008) dat de ontwikkeling van de grove motoriek het bewegingsapparaat betreft, waardoor het lichaam van een vrijwel passieve en onwillekeurige liggende houding tot een actieve en willekeurige rechtopstaande houding komt. Hiertoe behoren ontwikkeling van de hoofdbalans, het zitten, staan, kruipen, lopen en springen. Kurtz (2009) beschrijft hierop aansluitend dat alleen de meer fundamentele grofmotorische taken, zoals reiken, grijpen, zitten, kruipen en lopen tot de grove motoriek gerekend worden. Binnen deze opvatting worden de meer aangeleerde motorische vaardigheden, zoals het gooien met een bal 'culturele motorische vaardigheden' genoemd. Gedacht vanuit deze opvatting, wordt gesteld dat de grofmotorische ontwikkeling ongeveer na het tweede en in verfijnde vorm met het vierde levensjaar is voltooid (Kurtz, 2009).

Williams (zoals geciteerd in Netelenbos, 2000) definieert groot-motorisch gedrag (de grote of grove motoriek) als: 'het sequentieel en simultaan bewegen van grote delen van het lichaam op temporeel en ruimtelijk gecoördineerde wijze.' Volgens Williams behoren tot het grofmotorisch gedrag activiteiten als: lopen, rennen, vangen, balanceren en kopjebuitelen. Hierbij worden ook de activiteiten genoemd die hierboven de aangeleerde, culturele motorische vaardigheden werden genoemd. Netelenbos (1998) noemt deze vaardigheden in tegenstelling tot Bilo & Voorhoeve (2008) en Kurtz (2009) ook fundamentele motorische taken of vaardigheden. Hij doelt hiermee op bewegingsactiviteiten die een integraal en basaal onderdeel vormen van motorische spel- en sportvaardigheden. Hij zegt daarmee niet dat de grofmotorische ontwikkeling na het vierde jaar is voltooid. Hij benadrukt dat de periode van twee tot zes jaar een belangrijke periode is voor het verwerven van allerlei zogenoemde fundamenteel motorische taken of vaardigheden. Netelenbos geeft hierbij aan dat er geen uitputtende lijst is van deze taken. Elke onderzoeker op dit terrein komt tot zijn eigen inventarisatie.

Om een theoretische discussie zonder duidelijk eindpunt te ontlopen, hanteren we hier een pragmatische afbakening in termen van een praktische domeinbeschrijving die is ontleend aan Williams waarbij de grove motoriek wordt omschreven als het opeenvolgend en gelijktijdig bewegen van grote delen van het lichaam op temporeel en ruimtelijk gecoördineerde wijze. Voor de beschrijving van fundamentele grofmotorische vaardigheden gebruiken we de indeling zoals door SLO (2008) geformuleerd in tussendoelen en leerlijnen (TULE) voor het Bewegingsonderwijs: balanceren, stoeien, tikspelen, rijden, rollen, klimmen, springen, zwaaien, dansen en balspelen. Daarnaast hebben we de vaardigheid 'zitten' opgenomen als basisvoorwaarde voor het schrijven en tekenen. Dit lichten we hieronder toe.

De grove en fijne motoriek zijn nauw met elkaar verbonden. Kinderen moeten vaak bepaalde basisvoorwaarden binnen de grove motoriek beheersen, voordat zij toe zijn aan bepaalde taken binnen de fijne motoriek. Zo moet een kind zijn houding kunnen aanpassen aan de fijnmotorische taak die het wil gaan uitvoeren. En stabiel kunnen zitten of staan, zonder dat het de hele arm als stabiliserende factor (of evenwichtsfactor) hoeft te gebruiken. Dan kan het kind namelijk de onderdelen van de arm meer los van elkaar gebruiken. Dit zien we bijvoorbeeld terugkomen in de voorwaarden om te kunnen (leren) schrijven. Het kind moet daarvoor recht en stabiel op zijn stoel kunnen blijven zitten, moet zijn evenwicht kunnen bewaren en zijn houding kunnen aanpassen aan wat het wil gaan doen. Bovendien moet het zijn arm in een stabiele positie kunnen houden om met de vingers de fijne, nauwkeurige bewegingen te kunnen maken die voor het duidelijk schrijven essentieel zijn (Kurtz, 2009; Nouws, 1985). Als het kind nog te veel met de hele arm werkt, gaat dit ten koste van de nauwkeurigheid (Hettinga, 2011). De ontwikkeling vindt plaats via de bewegingen in de arm (achtereenvolgens van ellebogen en polsen) naar bewegingen met de afzonderlijke vingers. Kinderen die problemen hebben met deze onafhankelijke vingerbewegingen hebben dikwijls ook problemen met het efficiënt vasthouden van een pen. Vaak gaat dit samen met het reguleren van de spierspanning in de vingers (het kind drukt bijvoorbeeld te hard of te zacht op het potlood bij het schrijven en tekenen). Het kind moet beide handen onafhankelijk van elkaar kunnen bewegen en daarbij zelf de spierspanning kunnen regelen (Noordstar, 2009). Kortom, er is een duidelijke, deels voorwaardelijke, relatie tussen de grove en de fijne motoriek in relatie tot het schrijven en tekenen waarbij het goed kunnen 'zitten' een belangrijke rol speelt.

Fijne motoriek

Ook over de definitie van 'fijne motoriek' bestaan allerlei verschillende visies. Zo definieert Netelenbos (2000) fijnmotorisch gedrag (de fijne motoriek) naar voorbeeld van Williams als het gecoördineerd bewegen van individuele lichaamssegmenten, in het bijzonder het gebruik van handen en vingers in het manipuleren van objecten. Dit wordt volgens hem ook van oog-hand-coördinatie genoemd. Voorbeelden die tot het fijnmotorische gedrag behoren zijn volgens hem activiteiten als knippen, schrijven en kleien. In zijn definitie benadrukt Netelenbos weliswaar het gebruik van handen en vingers, maar de definitie sluit niet uit dat het ook om het coördineren van andere individuele lichaamssegmenten gaat. Dit zouden bijvoorbeeld de individuele lichaamssegmenten kunnen zijn die nodig zijn om te komen tot spraak of zindelijkheid.

Bilo en Voorhoeve (2008) beperken zich in hun definitie van de fijnmotorische ontwikkeling tot het gebruik van de handen en vingers en wel in de fase voor het grijpen, bij het grijpen zelf en bij het manipuleren van een voorwerp.

Ook voor de fijne motoriek kiezen we voor een pragmatische omschrijving, waarbij we uitgaan van de definitie van Netelenbos, naar voorbeeld van Williams. We beperken ons in de observatielijst Motoriek kleuters tot de gecoördineerde kleinere bewegingen van handen, vingers en pols in het manipuleren van objecten. Voor de beschrijving van de fijnmotorische taken/vaardigheden zijn de tussendoelen 'Beeldende vorming' (SLO, 2008) gebruikt. Hierin zijn de materialen en technieken die beschreven zijn in aan bod moeten komen in groep 1 en 2 van het basisonderwijs. Het gaat hierbij om:

- Tekenen met potlood, kleurpotlood, viltstift, waskrijt en bordkrijt
- Schilderen met vingerverf, plakkaatverf en gekleurde inkt
- Stempelen met aardappels, kurken en werken met sjablonen
- Knippen, scheuren en plakken met verschillende soorten papier, waaronder zijdevloei en sitspapier
- Repen knippen van textiel
- Rijgen met naald en draad (rietjes en kralen)
- Werken met kosteloos materiaal
- Bouwen met blokken
- Werken met constructiemateriaal
- Boetseren met plastische materialen, plasticine, natuursklei en brooddeeg
- Spelen met zand en water
- Digitale foto's maken
- Op de computer werken met eenvoudige tekenprogramma's

Naast de materialen en technieken, die gebruikt worden in het kader van beeldende vorming, zien we in groep 1 en 2 ook andere taken en vaardigheden terug die een beroep doen op de gecoördineerde bewegingen van handen, vingers en pols in het manipuleren van objecten. Deze zien we in het kader van zelfredzaamheid en het spelen met materialen. Na raadpleging van methodes die gebruikt worden in groep 1 en 2, diverse observatie-instrumenten en na overleg met verschillende deskundigen werden onderstaande taken en vaardigheden aan de domeinbeschrijving toegevoegd.

- Zelfredzaamheid:
 - aan- en uitkleden (open- en dichtmaken van knopen en ritsen, veters strikken);
 - eten en drinken (hanteren van een beker, openmaken van verpakkingen van eten en drinken);
 - toiletbezoek (handen wassen en drogen).
- Spel met materialen, zoals:
 - spelen met klein spel materiaal (zoals poppenhuis, auto's, dieren of natuurlijke materialen, zoals stenen, schelpen, kastanjes, ...);
 - manipuleren met kleine voorwerpen (zoals mozaïek, puzzels, kralenplank);
 - schrijfsels maken (met potlood, krijt, penseel of met vingers).

De ontwikkeling van de fijne motoriek is belangrijk, omdat een goede fijne motoriek noodzakelijk is om op latere leeftijd te leren schrijven.

Omdat het leren schrijven een van de belangrijkste activiteiten is waar kinderen in de basisschool mee te maken krijgen (Alkema, Van Dam, Kuipers, Lindhout en Tjerkstra, 2006) gaan we hier uitgebreid in op de relatie tussen fijnmotorische ontwikkeling en het schrijfonderwijs. Daarbij zullen we ook het tekenen betrekken omdat voor schrijven en tekenen deels dezelfde fijnmotorische vaardigheden nodig zijn¹.

¹ We ontleen deze passage aan een interne rapportage met betrekking tot de legitimering van de observatiepunten over fijnmotorische voorwaarden voor het (leren) schrijven en tekenen in de observatielijst op basis van literatuuronderzoek van de hand van Maaïke Foppen.

Om zich een goed handschrift te verwerven is het belangrijk dat zowel de grove als de fijne motoriek veel gestimuleerd is. Kinderen die zich voorafgaand aan de kleuterperiode optimaal motorisch hebben kunnen ontwikkelen hebben een betere startpositie dan kinderen die een achterstand hebben opgelopen (Kwast, 2006). In het onderstaande gaan we vooral in op de fijne motoriek.

De fijne motoriek, de oog-hand coördinatie en ruimtelijke oriëntatie zijn belangrijk bij het schrijven (Noordstar, 2009). Het kind stuurt met zijn hand de pen over het papier op basis van de informatie die het via de ogen binnenkrijgt. Het gaat om de afstemming tussen de handelingen die het kind met de hand uitvoert en de waarneming daarvan met de ogen, en omgekeerd. De ogen volgen het schrift. Hiervoor is de oogbeweging belangrijk (Keulen & Eerd-Smetsers, 2007). Schweitzer (2000) vat dit alles samen door te stellen dat schrijven een vorm van bewegingsonderwijs is. Het heeft volgens hem dan ook pas zin om aan leren schrijven te beginnen als het kind daar motorisch gezien echt aan toe is. Soms is het beter om eerst de fijne motoriek nog verder te stimuleren in de vorm van spelletjes en oefeningen. Het schrijven ontwikkelt zich vervolgens vanuit het 'schrijfftekenen' van met name letters en vanuit voorbereidende schrijfoefeningen (Louwe, 2012).

Methoden voor schrijf- en bewegingsonderwijs verwijzen soms naar het fasemodel van Mesker (beschreven in Kwast, 2006) met betrekking tot de ontwikkeling van de motoriek. We beperken ons hier tot de zogenoemde symmetriefase en lateralisatiefase die voor de ontwikkeling van de kleuter (en het leren schrijven) essentieel zijn. In de symmetriefase (2-6 jaar) beweegt het kind zich symmetrisch: de twee lichaamshelften doen hetzelfde in spiegelbeeld; als de ene hand knijpt, doet de andere dat ook. In de lateralisatiefase (5-9 jaar) kan het kind de lichaamsdelen steeds beter onafhankelijk van elkaar bewegen. Uiteindelijk leidt dit tot de ontwikkeling van dominantie en de keuze voor een voorkeurshand (bij ongeveer 80% van de kinderen). Juist de fijnmotorische vaardigheidsontwikkeling vanuit de symmetriefase naar de lateralisatiefase maakt het leren schrijven mogelijk. In de fase van het voorbereidend schrijven waarin de kleuter verkeert, overlappen de symmetrie- en de lateralisatiefase elkaar.

Uit het bovenstaande blijkt dat er een onlosmakelijke relatie bestaat tussen de algemene fijnmotorische ontwikkeling enerzijds en het (kunnen leren) schrijven anderzijds. Beide zijn niet los van elkaar te denken. Dit komt ook naar voren als we specifiek de ontwikkeling van de penguip beschrijven (een observatiepunt in de observatielijst heeft betrekking op de penguip). Veel leerkrachten gaan ervanuit dat er slechts één goede penguip bestaat, namelijk de dynamische driepuntsgreep. In de paramedische wereld wordt hierover vaak anders gedacht (Noordstar, 2009). De Korte Observatie Ergotherapie Kleuters (KOEK) en de Standaard Observatie Ergotherapie Schrijven en de Sensomotorische Schrijfvoorwaarden (SOESS) maken onderscheid tussen drie soorten penguipen: de onrijpe penguip, de overgangsgreep en de rijpe penguip. Bij de onrijpe grepen houdt het kind de pen met veel kracht in de palm van de hand vast. De onderarm zweeft hierbij vaak boven het tafelblad. Het kind maakt dan nog geen goede duim- en vingerbewegingen. Een voorbeeld van een onrijpe greep is de vuistgreep. Onrijpe grepen komen normaal gesproken voor bij kinderen tot vier jaar (Noordstar, 2009). Overgangsgrepen zijn vooral te zien bij kinderen van drie tot zes jaar. Deze worden gebruikt door kinderen die van een onrijpe naar een rijpe penguip gaan. Een voorbeeld is de statische driepuntsgreep. De penguip is hetzelfde als bij de dynamische driepuntsgreep, maar de beweging wordt ingezet vanuit de pols (Corstens-Mignot et al., 2006; Noordstar, 2009). Bij de rijpe penguipen is er sprake van een goede polscontrole en een gecoördineerde vingercontrole. Bij deze penguipen wordt adequaat gebruikgemaakt van de spieren in de hand. Daardoor kan op een goede manier worden geschreven. De meest bekende en meest gebruikte rijpe penguip is de dynamische driepuntsgreep. Dit is de greep waarbij de duim, de wijsvinger en de middelvinger kunnen buigen en strekken en waarbij de hand, zonder buiging of draaiing naar links of rechts in het verlengde van de onderarm ligt. Het is de greep die elke schrijfmethode aanbeveelt (Hettinga, 2011). Deze greep wordt daardoor vaak als meest ideale gezien, maar er zijn meerdere rijpe penguipen waarmee kinderen een goed handschrift kunnen ontwikkelen, zoals de 'dynamische vierpuntsgreep' en de 'à la ronde-greep'.

2.4.2 Het onderwijskundige perspectief

Op 1 augustus 2010 is de zogeheten Wet OKE (Ontwikkelingskansen door Kwaliteit en Educatie) van kracht geworden. Daarmee is voor- en vroegschoolse educatie (VVE) nadrukkelijk op de kaart gezet, onder meer in de vorm van kwaliteitseisen met betrekking tot het onderwijs aan jonge kinderen (binnen de kaders van kinderopvang, peuterspeelzaalwerk en basisonderwijs) en het toezicht daarop. Het doel van de wet is om de ontwikkelingskansen voor kinderen die al vroeg een achterstand (met name taalachterstand) hebben opgelopen te vergroten. In het kader van de Wet OKE dienen kindercentra en peuterspeelzalen te werken met een erkend VVE-programma. In de kleutergroepen in het basisonderwijs geeft men hieraan een vervolg, niet alleen binnen de domeinen taal en rekenen, maar ook op het gebied van motoriek en motorische ontwikkeling.

Men kan zich afvragen in hoeverre in de toepassing van programma's onderwijsleerprocessen verschillen van ontwikkelingsprocessen. Voor zover er sprake is van een erkend VVE-programma zal er in ieder geval sprake zijn van gerichte en doelbewuste processen, die in gang worden gezet en onderhouden op basis van welomschreven didactische principes waarin een belangrijke rol is toebedacht aan deskundige en professionele volwassenen. Of de vorderingen die kinderen in deze setting met toepassing van zo'n VVE-programma maken, nu te duiden zijn als *leren* of als *zich ontwikkelen*, lijkt een grotendeels academische vraag. Temeer omdat de didactische principes die men hanteert het beste werken wanneer ze aansluiten bij de eerder beschreven ontwikkelingsprocessen. In het VVE-programma Piramide bijvoorbeeld, dat bedoeld is voor jonge kinderen van 0 tot 7 jaar (in kinderopvang, peuterspeelzaal, gastouderopvang of groep 1 en 2 van het basisonderwijs) wordt aangehaakt bij ontwikkelingspsychologische theorieën met betrekking tot hechting, de zogeheten 'distancing'-theorie en de dynamische systeemtheorie (vergelijk Van Kuyk, Breebaart & Op den Kamp (2012)). Daarin is veel te herkennen van de eerder beschreven ontwikkelingspsychologische theorie volgens Sameroff (2010). Specifiek voor het domein van de motoriek en de motorische ontwikkeling sluit de dynamische systeemtheorie zoals Esther Thelen die omschrijft hier naadloos bij aan (vergelijk bijvoorbeeld Thelen & Smith, 1994). Het is te verwachten dat dit in meerdere of mindere mate geldt voor elk zichzelf respecterend VVE-programma, wil het kunnen verhelderen waarom het – in theorie – leidt tot bepaalde opbrengsten. Gewoonlijk worden deze opbrengsten van educatieve programma's en methoden echter niet omschreven in termen van *ontwikkelingsopgaven*, maar in termen van *leerdoelen* (kerndoelen, tussendoelen). *Kerndoelen* zijn op te vatten als *eindtermen*, maar dan wel als eindtermen die niet door ieder kind kunnen worden gehaald. Omdat er geen sprake is van een afsluitend examen waarin eindtermen worden geoperationaliseerd, worden de kerndoelen in de praktijk geïnterpreteerd als *streefdoelen*. Overigens pleit de Onderwijsraad (2010) in *Naar een nieuwe kleuterperiode in de basisschool* voor het gebruik van de term *ontwikkeldoelen* (zoals dat in Vlaanderen voor het kleuteronderwijs gebeurt). In deze ontwikkeldoelen moet dan worden verwoord, wat men in het onderwijsprogramma zou moeten aanbieden, rekening houdend met het individuele karakter van ontwikkeling (als "een groeiproces dat elk kind op een eigen manier doorloopt in een eigen tempo" (p. 57)). Tegelijkertijd verwijst de raad in dit kader echter weer naar "basiscompetenties waarover kleuters zouden moeten beschikken" (p. 57) als waren het eindtermen.

De SLO biedt een overzicht van de kern- en tussendoelen zoals deze kunnen worden onderscheiden in het basisonderwijs (zie voor het SLO-project Tussendoelen & Leerlijnen, TULE, de website <http://tule.slo.nl>).

Ook voor groep 1 en 2 staat vermeld wat aan bod moet komen, welke stappen kinderen zetten en er worden aanwijzingen voor de leerkracht gegeven. Ten aanzien van het motorisch functioneren betreft het:

- Bewegingsonderwijs: balanceren, klimmen, zwaaien, over de kop gaan, springen, hardlopen, mikken, jongleren, doelspelen, tikspelen, stoeispelen en bewegen op muziek (zie uitwerking kerndoel 57)
- Aanbieden van materialen en technieken bij beeldende vorming (zie uitwerking kerndoel 54, Kunstzinnige oriëntatie, beeldende vorming)

SLO ontwikkelt (beheersings)doelen Bewegingsonderwijs en (beheersings)doelen op het gebied van fijne motoriek (uitgave najaar 2015). Deze geven aan welke onderdelen belangrijk zijn in het aanbod en naar welke doelen men kan toewerken.

2.4.3 Uitgaan van normaal functioneren, niet van stoornissen

Bij de constructie van de observatielijst zijn we uitgegaan van normaal motorisch functioneren en niet van motorische stoornissen, zoals dat bijvoorbeeld het geval is bij diagnostische instrumenten waarmee bijvoorbeeld cerebrale parese vastgesteld kan worden. In dit type instrumenten zijn voornamelijk items te vinden die verwijzen naar aspecten van een stoornis. Motoriek kleuters heeft als belangrijkste doel leerkrachten ondersteuning te bieden bij het in kaart brengen van het motorische functioneren van kinderen en de daaraan ten grondslag liggende ontwikkeling. Het instrument is niet bedoeld om kinderen op te sporen met een stoornis in het motorische functioneren of met een verhoogd risico op een stoornis zoals DCD. Dergelijke (predictieve) ambities heeft Cito met deze instrumenten niet. In de keuze van de observatiepunten en de formulering daarvan is daarom normaal motorisch functioneren het uitgangspunt geweest.

2.4.4 Continuïteit en stabiliteit in ontwikkeling: voorspelbaarheid over tijd?

De observatielijsten voor jonge kinderen zijn ontwikkeld vanuit de gedachte dat er in de ontwikkeling sprake is van een zekere mate van stabiliteit. Verschillende elementen in de ontwikkelingspsychologische theorievorming, zoals Sameroff (2010) deze samenvat, dragen bij aan deze stabiliteit. Op de eerste plaats is er sprake van al dan niet genetisch bepaalde biologische eigenschappen die het kind als betrekkelijk stabiele factor meebrengt. Binnen het domein van de motorische ontwikkeling is er onmiskenbaar sprake van een genetische factor, zoals de eerder besproken neurale rijpingstheorie (Gesell & Amatruda, 1947) veronderstelt. Er zijn echter ook redenen om aan te nemen dat deze veronderstelde stabiliteit lager uitvalt dan men zou verwachten, bijvoorbeeld door ingrijpende gebeurtenissen die lichamelijk letsel met zich meebrengen (zoals een beenbreuk) of het ontstaan van angst voor specifieke bewegingssituaties naar aanleiding van bepaalde ervaringen. En zelfs al is de genetische invloed groot, motorische ontwikkeling is méér dan het simpelweg 'afspelen' van een genetisch programma, zoals de dynamische systeemtheorie laat zien. Biologische processen waarin taak- en omgevingsfactoren invloed uitoefenen laten substantieel ruimte voor het effect van deze factoren. Stabiliteit en voorspelbaarheid over tijd zijn bovendien geringer naarmate er meer tijd is verstreken tussen de metingen en het om jongere kinderen gaat.

We beschikken niet over longitudinaal onderzoek met betrekking tot stabiliteit en voorspellende waarde van de scores op Motoriek kleuters. Daarom reikt de functie van Motoriek kleuters niet verder dan het beschrijven van gedrag zoals het kind dat op dit moment laat zien. Dit is op zich echter al belangrijk genoeg omdat leerkrachten hiermee informatie in handen krijgen die hen aanzet tot het zoeken naar verandermogelijkheden in de huidige situatie. Daarmee kan een negatieve spiraal worden omgebogen (vergelijk Hermanns, Nijnatten, Verheij & Reuling, 2008). Voorwaarde is wel dat zij voldoende kennis hebben van en gevoelig zijn voor de aard van (motorische) ontwikkelingsprocessen om goed te kunnen inschatten hoe deze processen kunnen worden bijgestuurd en geredimeerd.

2.4.5 Doorlopende leerlijnen

In het Cito Volgsysteem primair en speciaal onderwijs voor de groepen 1 tot en met 8, waarvan Motoriek kleuters deel uitmaakt, wordt zoveel mogelijk gewerkt met doorlopende leerlijnen als basis voor de instrumenten. In de regel gaat het om vaardigheden zoals rekenen-wiskunde, technisch lezen, spelling, begrijpend lezen en beheersing van de woordenschat. Deze vaardigheden laten over de leerjaren heen een stelselmatige groei zien, waardoor de leervorderingen over een groot deel van de basisschoolperiode voor een beperkt aantal leerlijnen in de vorm van vaardigheidsscores te volgen zijn. Bij de constructie van toetsen voor deze leerlijnen worden kerndoelen en tussendoelen, en de uitwerkingen daarvan in methodes, gehanteerd die houvast bieden bij het operationaliseren van de vaardigheidsgroei.

Ook ten aanzien van het motorisch functioneren is getracht uit te gaan van een doorlopende leerlijn. Voor kleuters en peuters hebben we dezelfde indeling naar grove en fijne motoriek als uitgangspunt genomen. Voor nóg jongere kinderen (zes maanden tot twee jaar) is afgezien van deze tweedeling. In de Observatielijst voor baby's en dreumesen is een reeks concrete en hoofdzakelijk grofmotorische

ontwikkelingsmijlpalen opgenomen die het mogelijk maken de motorische ontwikkeling efficiënt te karakteriseren. Een onderscheid naar grove en fijne motoriek is bij deze doelgroep nauwelijks zinvol. Hoewel het onderscheid naar grove en fijne motoriek van begin af aan het uitgangspunt vormde, is de uiteindelijk gerealiseerde schaalstructuur toch vooral langs empirische weg tot stand gekomen. Bij de constructie van observatiepunten gingen we namelijk niet primair uit van dit globale onderscheid, maar van clusters van onderling vrij sterk samenhangende observatiepunten, zoals 'kleine voorwerpen manipuleren', 'tekenen en schrijven', 'mikken en jongleren', 'klimmen' en 'dans' die tezamen het volledige domein van motorisch functioneren op kleuterleeftijd leken te dekken. Achteraf, dat wil zeggen op basis van empirische bevindingen, bleken deze clusters en de daarin ondergebrachte observatiepunten prima onder te brengen langs de genoemde hoofdlijnen van grove en fijne motoriek.

Tot slot van deze paragraaf willen we graag een kanttekening maken bij de specifieke aard van observatieschalen als instrument. Anders dan bij de leervorderingstoetsen in het Cito Volgstelsel primair en speciaal onderwijs kunnen leerlingen niet rechtstreeks *zelf* laten zien waartoe zij in staat zijn. Bij de genoemde leervorderingstoetsen is de ruwe toetscore op basis van de itemkenmerken en door toepassing van IRT te transformeren tot de al genoemde vaardigheidsscore. Voor een bepaald vaardigheidsdomein, zoals bijvoorbeeld rekenen-wiskunde, kunnen de individuele scores van een leerling op *verschillende* toetsen voor verschillende leerjaren en afnamemomenten op deze manier worden afgebeeld op één en dezelfde vaardigheidsschaal. De leervorderingen van een leerling over tijd kunnen op deze manier dus direct in beeld worden gebracht.

Bij observatielijsten kan de leerling *niet zelf* laten zien wat hij kan: zijn competentie wordt *beoordeeld* door de leerkracht aan de hand van een aantal uitspraken en stellingen. De leerkracht geeft in Motoriek kleuters met behulp van een vierpuntsschaal aan in welke mate de uitspraken over het motorisch functioneren van toepassing zijn. De ervaring leert dat de beoordelende leerkracht hierbij een betrekkelijk subjectief en relatief referentiekader hanteert, ook al is er geprobeerd om de observatiepunten zo concreet mogelijk te formuleren. Het gevolg is dat afname van een lijst met stellingen als "Valt niet vaak" bij een grote en representatieve groep kleuters ongeveer dezelfde gemiddelde ruwe score zou kunnen opleveren als afname bij een grote, representatieve groep peuters. Het betrekkelijk geringe verschil in gemiddelde betekent niet dat kleuters ongeveer even vaak vallen als peuters; het duidt op een verschuiving in het referentiekader van de beoordelaar.

Het bovenstaande impliceert dat we ons bij de interpretatie van de ruwe scores dienen te realiseren dat het lastig is om deze te beschouwen als een *rechtstreekse* maat voor groei of ontwikkeling. Door relatieve normering kunnen we – door de percentielscores te interpreteren – wel vaststellen hoe goed of hoe slecht een leerling het doet ten opzichte van andere leerlingen. Pas in tweede instantie mogen we aannemen dat een leerling met een hoge percentielscore beter functioneert dan een leerling met een lage percentielscore en dus waarschijnlijk verder is in zijn of haar ontwikkeling. Deze conclusie is niet uniek voor Motoriek kleuters, maar geldt voor de meeste observatie-instrumenten die op soortgelijke wijze zijn geconstrueerd, al wordt dit zelden expliciet aangegeven. Te vaak wordt ten onrechte impliciet aangenomen dat scores op dit soort schalen een rechtstreekse operationalisatie van groei en ontwikkeling representeren. Kortom, de gebruiker dient zich bij het gebruik van Motoriek kleuters en vergelijkbare instrumenten steeds goed te realiseren dat interpretaties van de resultaten in termen van ontwikkeling of ontwikkelingsachterstand slechts indirect mogelijk zijn.

Wel is het zo dat ten aanzien van de ruwe scores duidelijke leeftijdsverschillen blijken op te treden, die wijzen op verschillen in motorisch functioneren die samenhangen met de leeftijd (zie hoofdstuk 4). Het ligt voor de hand om deze verschillen te interpreteren in termen van vaardigheidsgroei en ontwikkeling. Niettemin pleiten we ervoor om verschillen in scores tussen twee of meer meetmomenten voor een individuele leerling niet rechtstreeks te interpreteren in termen van ontwikkeling. Er kan weliswaar sprake zijn van toename in de motorische vaardigheid zoals deze is geobserveerd door een leerkracht, maar men kan uit het verschil tussen de ruwe scores niet aflezen of deze toename gelijke tred houdt met de 'normale' vaardigheidsgroei van leeftijdgenoten. Daarom hebben we de ruwe scores relatief genormeerd in viermaandelijks normgroepen. We adviseren om altijd de resulterende normscores (en niet de ruwe scores) in termen van 'signalen' te gebruiken om eventuele verschuivingen in het motorische functioneren van leerlingen vast te stellen.

3 Beschrijving van de observatielijst

3.1 Constructie van de observatielijst

Bij de ontwikkeling van Motoriek kleuters werden voorafgaand aan en tijdens het ontwikkeltraject de bij Cito gebruikelijke procedures gehanteerd, die het waarborgen van een optimale kwaliteit beogen. In onderstaande beschrijving van het constructieproces gaan we kort in op deze procedures.

Theoretische inkadering en inhoudsverantwoording

Voordat de daadwerkelijke constructie startte vond conceptontwikkeling plaats waarbij de meetpretentie werd geformuleerd en de theoretische inkadering op conceptueel niveau werd vastgelegd en verantwoord. Daarbij werden onderstaande bronnen geraadpleegd:

- Kerndoelen Bewegingsonderwijs en Kunstzinnige oriëntatie.
- Tussendoelen en leerlijnen Bewegingsonderwijs en Kunstzinnige oriëntatie (beeldende vorming): Op de website <http://tule.slo.nl> zijn alle kerndoelen van 2006 uitgewerkt in inhouden en activiteiten.
- Concept SLO-doelen Bewegingsonderwijs: SLO ontwikkelt (beheersings)doelen voor jonge kinderen voorafgaand aan groep 1 en eind groep 2 op het gebied van motoriek (uitgave najaar 2015). De subdomeinen binnen deze nieuwe SLO-doelen voor het bewegingsonderwijs komen grotendeels overeen met de eerder genoemde tussendoelen Bewegingsonderwijs die in 2006 geformuleerd zijn door het SLO. Hierbij is één uitzondering en dat is 'hardlopen'. Dit subdomein wordt in de nieuwe uitwerking van doelen niet genoemd.
- Beschrijvingen van methodes die op kinderdagverblijven, in peutergroepen en in groep 1 en 2 gebruikt worden, waaronder Kleuterplein, Schatkist, Kaleidoscoop, Startblokken, Doe meer met Bas, Speelplezier, Sporen en Piramide. We hebben ons vooral gericht op de beschrijving van de doelen bewegingsonderwijs, sensomotoriek en schrijfvoorwaarden, en kunstzinnige ontwikkeling, waaronder beeldende ontwikkeling en dans, in deze methodes (Cito, 2012).
- Bestaande instrumenten (zoals een eerdere ontwikkelde, nooit uitgebrachte versie Motoriek peuters en Motoriek kleuters en de Cito Peuterobservatielijst POL en Kleuterobservatielijst KOL).
- Onderzoeksrapporten en achtergrondliteratuur.

Vervolgens werd op operationeel niveau een inhoudsverantwoording geschreven, met daarin een weergave van het beoogde ontwikkelproces, een blauwdruk van het instrument (omvang, inhoud, vorm van de observatiepunten), de beschrijving van de antwoordcategorieën, de kwaliteitscriteria die met betrekking tot de observatiepunten werden gehanteerd en een beschrijving van de methodologie die nodig werd geacht om uiteindelijk de betrouwbaarheid en validiteit van het instrument te kunnen aantonen. Deze conceptbeschrijving is voorgelegd aan verschillende experts: toetsdeskundigen binnen Cito, twee wetenschappers op het gebied van motorische ontwikkeling verbonden aan twee universiteiten, een fysiotherapeut en een peuterleidster. Op basis van hun feedback is de theoretische inkadering en de inhoudsverantwoording enigszins bijgesteld.

Constructie van items

Zoals aangegeven in het vorige hoofdstuk zijn we voor Motoriek kleuters bij de constructie van de observatielijst uitgegaan van twee (clusters van) ontwikkelingsopgaven, die we vooral hebben opgevat in de betekenis van specifieke domeinen 'in ontwikkeling' ('ontwikkeldoelen'), te weten:

- Grove motoriek
- Fijne motoriek

Uitgaande van de bijgestelde theoretische inkadering en inhoudsverantwoording (voor een beschrijving zie hoofdstuk 2) zijn voor elk domein in een constructiegroep van vier toetsdeskundigen, gespecialiseerd in de ontwikkeling van het jonge kind, domeinomschrijvingen gemaakt, waarbij elk domein zo breed mogelijk operationeel werd gedefinieerd. Deze domeinomschrijvingen vormden de basis voor een toetsmatrijs met

een onderverdeling in hoofd- en subdomeinen inclusief de specificatie van de aantallen gewenste observatiepunten per subdomein. Het aanvankelijke streven was om 33 observatiepunten te ontwikkelen. Op grond van eerdere ervaringen werd dit aantal nodig geacht om een voldoende betrouwbare observatielijst te construeren. Gezien het doel en rekening houdend met de praktische bruikbaarheid werd dit aantal tevens beschouwd als het maximum. Uiteindelijk bleken 30 observatiepunten per lijst ook te voldoen.

Omdat de ontwikkeling van Motoriek kleuters, inclusief de proeftoetsing, parallel werd uitgevoerd (met een gedeeltelijke overlap qua inhoud) met een soortgelijke observatielijst voor peuters, bespreken we hier beide observatielijsten tezamen. Bij de motorische ontwikkeling van peuters ligt het zwaartepunt op de grove motoriek en minder op de fijne motoriek. Daarom is ervoor gekozen om in de peuterlijst meer observatiepunten te formuleren die betrekking hebben op het grofmotorisch functioneren. In groep 1 en 2 van het basisonderwijs ligt het accent wat sterker op het fijnmotorische functioneren (onder andere in verband met het voorbereidend schrijven).

Op basis van advies van experts op het gebied van motorische ontwikkeling en eerdere ervaring met het ontwikkelen van een soortgelijk observatie-instrument, werd besloten om meer observatiepunten te ontwikkelen dan er nodig zouden zijn en deze uit te testen in een proefonderzoek. Uitgaande van een geschatte mortaliteit van 60% werden 75 observatiepunten per observatielijst ontwikkeld en opgenomen in het proefonderzoek. Voor een verdeling van de observatiepunten over de hoofddomeinen zie onderstaand schema.

Globale toetsmatrix Motoriek peuters

	Beoogd	In proeftoets
Grove motoriek	19	47
Fijne motoriek	14	28
Totaal	33	75

Globale toetsmatrix Motoriek kleuters

	Beoogd	In proeftoets
Grove motoriek	15	37
Fijne motoriek	18	38
Totaal	33	75

In de eerste constructiefase werden zoveel mogelijk potentiële indicatoren voor adequaat functioneren binnen elk domein benoemd. Dit gebeurde voor alle leeftijdsgroepen tegelijk (peuters, kleuters). Steeds werd de vraag gesteld of specifiek functioneren kon worden opgevat als een goede indicator van adequaat functioneren voor een van de twee leeftijdscategorieën. Vervolgens werden deze indicatoren omgezet in twee gepaarde stellingen die per item als eindpunten van een bipolaire dimensie kunnen worden opgevat.

Voorafgaand aan de constructie van de observatiepunten werd een aantal specifieke aandachtspunten geformuleerd:

- Aansluitend bij de dynamische systeemtheorie dienen de observatiepunten niet zozeer gericht te zijn op het beheersen van specifieke motorische vaardigheden, maar meer op het motorisch functioneren van kinderen tijdens bezigheden (spelen/bewegen) in de dagelijkse context van de peuterspeelzaal, het kinderdagverblijf en groep 1 en 2 van het basisonderwijs.
- De bezigheden moeten in de dagelijkse praktijk regelmatig waarneembaar zijn, zodat leidsters/-leerkrachten er een beeld van hebben of zich er een beeld van kunnen vormen door gericht te observeren. Weven met papierstroken of stof komt bijvoorbeeld niet vaak voor. Verven, tekenen, kleien zijn daarentegen bezigheden die vaak en in diverse situaties aan de orde zijn. Deze bezigheden zijn op

te vatten als (ontwikkelings)taken die het kind leert beheersen, maar deze 'taken' moeten in deze context niet verward worden met (al dan niet gestructureerde) opdrachten. Alle taken dienen betrekking te hebben op zelfgekozen of uitgelokte bewegingssituaties, bijvoorbeeld vrij rennen op de speelplaats, alleen of samen met andere kinderen, in de speelzaal of tijdens een bewegingsspel.

- Aangezien naast de resultaten op de observatielijst ook de formulering van de observatiepunten de onderwijspraktijk kan beïnvloeden, is het belangrijk dat er alleen zinvolle taken benoemd worden, die aansluiten bij de doelen waaraan leidsters/leerkrachten met hun aanbod naartoe werken.
- Aansluitend bij het opleidingsniveau van de gebruikersgroep is het belangrijk om geen vaktermen op het gebied van de motoriek en motorische ontwikkeling te bezigen, maar eenvoudige woorden te gebruiken die verwijzen naar de alledaagse situatie in de groep en de observatiepunten simpel en kort te formuleren.

Op basis van de genoemde aandachtspunten en inhoudelijke bronnen, is een verfijndere toetsmatrijs ontwikkeld met daarin herkenbare clusters van taken, waarbij observatiepunten ontwikkeld zijn. Deze vormen een weerspiegeling van veel voorkomende taken in de praktijk, in de doelen, methodes en in andere observatie-instrumenten. We beschrijven hier de gedetailleerde toetsmatrijs voor Motoriek kleuters, met de gewenste en gerealiseerde aantallen observatiepunten per 'taak'. Er is steeds naar gestreefd om van de verzameling observatiepunten die meegenomen werd in de proeftoetsing en in tweede instantie de definitieve versie van de observatielijst vormt, een evenwichtige en representatieve steekproef te maken uit de totale itemdomeinen met betrekking tot de grove en fijne motoriek. Dit valt na te gaan door de observatiepunten te vergelijken met de door de SLO geformuleerde tussendoelen en leerlijnen met betrekking tot bewegingsonderwijs en beeldende vorming.

Toetsmatrijs Motoriek kleuters

Domein	Taken	Wens	In proeftoets	Definitieve lijst
Fijne motoriek	aan- en uitkleden	1	1	1
Fijne motoriek	bouwen	1	3	1
Fijne motoriek	eten en drinken	1	1	1
Fijne motoriek	in elkaar zetten	1	3	1
Fijne motoriek	kleien	1	3	1
Fijne motoriek	kleine voorwerpen manipuleren	1	3	1
Fijne motoriek	knippen	1	3	1
Fijne motoriek	plakken	1	3	1
Fijne motoriek	prikken	1	2	1
Fijne motoriek	tekenen en schrijven	4	9	4
Fijne motoriek	verven	1	3	0
Fijne motoriek	vouwen	1	3	2
Grove motoriek	aan- en uitkleden	1	1	0
Grove motoriek	algemene coördinatie	3	5	2
Grove motoriek	balanceren	2	4	2
Grove motoriek	groepsspel	2	4	1
Grove motoriek	klimmen	1	4	1
Grove motoriek	dans	2	4	2
Grove motoriek	lopen en rennen	1	3	1
Grove motoriek	mikken en jongleren	2	4	2
Grove motoriek	over de kop gaan	1	2	1
Grove motoriek	springen	2	4	2
Grove motoriek	zitten	1	2	1
Aantal items		33	75	30

Operationalisatie doelen Tule Bewegingsonderwijs

Op de website <http://tule.slo.nl> zijn de kerndoelen 'Bewegingsonderwijs' van 2006 concreet uitgewerkt in inhouden en activiteiten. Ook voor groep 1 en 2 staat vermeld wat aan bod moet komen, welke stappen kinderen zetten en er worden aanwijzingen voor de leerkracht gegeven. De doelen 'Bewegingsonderwijs' hebben betrekking op balanceren, klimmen, zwaaien, over de kop gaan, springen, hardlopen², mikken, jongleren, doelspelen, tikspelen, stoeispelen en bewegen op muziek. In de taken geformuleerd in de toetsmatrijs zijn de doelen grotendeels te herkennen. Hierop zijn enkele uitzonderingen:

- De taak 'zwaaien' hebben we uit de matrijs verwijderd. Hiervoor hadden we in eerste instantie wel observatiepunten geconstrueerd, maar tijdens een eerste screening door de constructiegroep bleek dat in de inrichting van de binnen- en buitenruimte op veel locaties geen gelegenheid is voor kinderen om te zwaaien (op een schommel of aan ringen of touwen). Alhoewel dit een doel is waaraan in het onderwijs gewerkt wordt, is het vanwege de ontbrekende randvoorwaarden niet zinvol en psychometrisch gezien niet correct om hierover een observatiepunt op te nemen.
- Stoeispelen en tikspelen zijn geoperationaliseerd in de observatiepunten onder 'groepsspel'.
- De observatiepunten onder de taak 'Dans' hebben deels te maken met de doelen 'bewegen op muziek', maar ook met coördinatie in het algemeen en met lichaamsbesef. Deze observatiepunten richten zich op het soepel bewegen in de ruimte en het imiteren van grote bewegingen.

Er zijn ook enkele taken toegevoegd die niet te herkennen zijn in de doelen Bewegingsonderwijs:

- Observatiepunten over 'zitten' zijn toegevoegd, aangezien een goede zithouding een belangrijke voorwaarde vormt voor het schrijven en andere 'schoolse taken' die kinderen in groep 3 zittend aan een tafeltje uitvoeren.
- Alhoewel aan- en uitkleden geen onderdeel van de doelen rondom bewegingsonderwijs vormen, hebben we ook hierover een observatiepunt uitgeprobeerd in het proefonderzoek. Voor de zelfredzaamheid en zelfstandigheid van kinderen is het zelfstandig aan- en uitkleden een belangrijke taak binnen het grof- en fijnmotorisch functioneren.
- Ook enkele observatiepunten over de algemene coördinatie zijn opgenomen. Dit zijn observatiepunten die iets over het motorisch functioneren zeggen maar 'taakoverstijgend' zijn, zoals 'Valt niet vaak'.

Operationalisatie doelen Tule Beeldende vorming

In Tule is het kerndoel 'Kunstzinnige oriëntatie' uitgewerkt in concrete inhouden en activiteiten met betrekking tot 'beeldende vorming'. Alhoewel de doelen 'beeldende vorming' geformuleerd zijn om richting te geven aan het onderwijsaanbod waarin de kunstzinnige oriëntatie van kinderen wordt gestimuleerd, geven deze doelen ook richting aan het aanbod van fijnmotorische taken. Er wordt namelijk geformuleerd welke materialen en technieken in groep 1 en 2 aan bod moeten komen. Het gaat hierbij om:

- Tekenen met potlood, kleurpotlood, viltstift, waskrijt en bordkrijt
- Schilderen met vingerverf, plakkaatverf en gekleurde inkt
- Stempelen met aardappels, kurken en werken met sjablonen
- Knippen, scheuren en plakken met verschillende soorten papier, waaronder zijdevloei en sitspapier
- Repen knippen van textiel
- Rijgen met naald en draad (rietjes en kralen)
- Werken met kosteloos materiaal
- Bouwen met blokken
- Werken met constructiemateriaal
- Boetseren met plastische materialen, plasticine, natuurklei en brooddeeg
- Spelen met zand en water
- Digitale foto's maken
- Op de computer werken met eenvoudige tekenprogramma's

In de formulering van de doelen met betrekking tot 'beeldende vorming' worden echter geen aanwijzingen gegeven voor de kwaliteit waarmee deze 'taken' moeten worden uitgevoerd.

² In de nieuwste versie van de SLO-doelen van 2015 is dit doel niet meer opgenomen.

Op basis van bovenstaande opsomming van materialen en technieken is een aantal taken toegevoegd aan de toetsmatrijs, namelijk tekenen, schilderen (verven), knippen, plakken, rijgen met naald en draad (bij manipuleren met kleine voorwerpen), bouwen, werken met constructiemateriaal (in elkaar zetten), boetseren (kleien).

Deze taken die ontleend zijn uit de doelen 'beeldende ontwikkeling', zijn aangevuld na methode-onderzoek en het bestuderen van bestaande instrumenten met taken die te maken hebben met zelfredzaamheid (aankleden; eten en drinken), het spelen met materialen (kleine voorwerpen manipuleren zoals mozaïek, puzzels, kralenplank), prikken, vouwen en schrijven. Een taak die het fijnmotorisch functioneren goed bleek te meten was vouwen. Daarom zijn hier ook twee observatiepunten over opgenomen.

Feedback expertgroep

De ontwikkelde observatiepunten zijn vervolgens voorgelegd aan een groep experts en aangepast naar aanleiding van hun feedback. De expertgroep bestond uit:

- twee kinderfysiotherapeuten;
- een pabo-docent beeldende vorming/schrijven;
- een ervaren peuterleidster;
- een startende pedagogisch medewerkster;
- drie (ex-)leerkrachten, van wie twee toetsdeskundigen en een educatief adviseur.

Hoe vaak items ook besproken en bijgesteld zijn, het blijft moeilijk om te voorspellen of de leerkrachten de items eenduidig begrijpen en of de items betrouwbaar en valide zijn. Om een meer verantwoorde selectie van items te kunnen maken, is in het proefonderzoek statistische, kwantitatieve en kwalitatieve informatie verzameld.

Proefonderzoek

In de periode maart/april 2013 zijn de observatiepunten onderzocht in een proefonderzoek dat ons gegevens opleverde over de kwaliteit ervan. De betreffende gegevens geven zicht op het inhoudelijk en psychometrisch functioneren en vormden de basis voor de definitieve selectie van observatiepunten. De onderzochte groep leerlingen en leerkrachten is weliswaar niet volledig representatief voor de populatie, maar gaf ons door de omvang ervan voldoende informatie over de vraag hoe de items zich in de populatie zullen gedragen (Eggen & Sanders, 1993).

Het proefonderzoek had, zoals aangegeven, zowel betrekking op Motoriek kleuters als op Motoriek peuters en vond plaats in de periode maart – april 2013. Er werden zowel (kleutergroepen van) basisscholen als peuterspeelzalen en kinderdagverblijven in betrokken.

Met het oog op de lijst voor kleuters werden scholen benaderd in twee afzonderlijke rondes: er is een aanvullende steekproef getrokken en aangeschreven toen de deelnamebereidheid niet groot genoeg bleek om de gewenste aantallen scholen en leerlingen te realiseren. Uiteindelijk meldden zich 26 scholen aan met in totaal 759 leerlingen. Daarvan hebben uiteindelijk 386 kleuters aan het proefonderzoek deelgenomen die alle in de analyses konden worden betrokken. Daarnaast konden resultaten voor 676 peuters, afkomstig van 30 peuterspeelzalen en kinderdagverblijven, in de analyses worden meegenomen.

Omdat het totaal aantal uit te testen observatiepunten (75 voor peuters en 75 voor kleuters) veel te groot was om in één vragenlijst af te nemen en er bovendien sprake was van een aantal voor leeftijd specifieke observatiepunten, werd besloten de observatiepunten te verdelen over tien verschillende lijsten in een zogenaamd 'verbonden (onvolledig) design'. Dit design zorgt ervoor dat bij IRT-schaling alle items (voor een bepaalde leeftijdsgroep) op dezelfde schaal kunnen worden gebracht³. In dit verbonden design komt elk observatiepunt in twee lijstversies voor. Dit betekent dat er in totaal 300 observatiepunten over 10 observatielijstversies werden verdeeld, dus 30 observatiepunten per lijst. De lengte van de uitgeteste lijstversies kwam hiermee ongeveer overeen met die van de uiteindelijk te construeren observatielijst.

³ Er is voor de schaling van de items geen gebruikgemaakt van IRT. Voor enkele analyses is wel gebruikgemaakt van op IRT gebaseerde technieken.

Op deze manier werden 75 verschillende observatiepunten aan peuters voorgelegd (in vijf versies van de lijst) en 75 aan kleuters (eveneens in vijf versies). Er waren 28 observatiepunten die zowel in de peuter- als in de kleuterlijsten zijn opgenomen. Vanwege relatief hoge percentages ontbrekende waarden werden drie van deze observatiepunten in de peuterlijsten niet in de analyses meegenomen. Het totaal aantal *verschillende* observatiepunten waarover analyses konden worden uitgevoerd kwam daarmee op 119 en uiteindelijk (door het uitvallen van één observatiepunt op grond van hardop-denk-onderzoek en evaluatieformulieren) 118.

De tien verschillende versies werden zo gelijk mogelijk over leerkrachten en leidsters, respectievelijk de bijbehorende groepen kinderen verdeeld. Gemiddeld waren er 154 observaties per observatiepunt over de kleuterversies van de lijst en 270 observaties over de peuterversies. Voor de peuterversies was dit ruim voldoende, voor de kleuterversies – met het oog op de scheve verdeling van de antwoordcategorieën – enigszins aan de krappe kant.

Evaluatieformulieren

In het kader van het proefonderzoek is gebruik gemaakt van een evaluatieformulier, waarbij de leidsters, pedagogisch medewerkers en leerkrachten opmerkingen over de observatiepunten en de observatielijst in zijn geheel konden maken. Hier is door de leerkrachten van groep 1 en 2 maar zeer beperkt gebruik gemaakt. Van slechts vier respondenten werden ingevulde evaluatieformulieren terug ontvangen. De reden hiervan is niet bevestigd bij de respondenten.

Hardop-denk-onderzoek

Om een bijdrage te leveren aan het valideren van het instrument, is er parallel aan het proefonderzoek een kleinschalig kwalitatief vooronderzoek gehouden om het gebruik van (delen van) de observatielijst in de praktijk te testen. Het doel van het kwalitatief vooronderzoek was om na te gaan of pedagogisch medewerkers, leidsters en leerkrachten de observatiepunten eenduidig begrijpen. Er is gebruikgemaakt van de hardop-denk-methode (Someren, Barnard & Sandberg, 1994) met achteraf een semigestructureerd interview naar hun mening over de observatiepunten⁴. Hierdoor konden we enerzijds een indruk krijgen van de (koers van de) denkprocessen die zich afspelen tijdens het invullen van de observatielijst, anderzijds van de kwaliteit van de observatiepunten zoals deze wordt geëvalueerd door praktijkdeskundigen. Een leidster van een peuterspeelzaal, een pedagogisch medewerkster op een kinderdagverblijf en een leerkracht op een school werkten mee aan dit onderzoek. De data uit hardop-denk-protocollen zijn geanalyseerd. Dit is gedaan aan de hand van de methode 'systematische analyse van verbale data', waarbij gebruikgemaakt is van het MEPA-analyseprogramma (Multiple Episode Protocol Analysis). Op basis van de analyses werden inhoudelijke adviezen opgesteld met betrekking tot de kwaliteit van de observatiepunten en de selectie ervan.

Bijstellen en selecteren van observatiepunten

In latere sessies van de constructiegroep werd alle verzamelde inhoudelijk-kwalitatieve informatie over de afzonderlijke observatiepunten en over de observatielijst in zijn geheel geïnterpreteerd. De uiteindelijke selectie van items is gebaseerd op de omschrijving van het inhoudsdomain (zie de eerdere beschrijving van de toetsmatrijs Motoriek kleuters), psychometrische gegevens, aansluiting bij de gebruikerscontext en praktische bruikbaarheid.

De volgende informatie is hierbij gebruikt: de empirische gegevens van het hierboven beschreven proefonderzoek en kwalitatieve gegevens voorkomend uit de reacties via het evaluatieformulier van de bij het proefonderzoek betrokken pedagogisch medewerkers, leidsters en leerkrachten en uit het hardop denk-onderzoek. We beperken ons in het navolgende tot de observatielijst Motoriek kleuters.

Voor elk van de twee onderscheiden domeinen werd een somscore berekend. Vervolgens werden intercorrelaties en de correlatie van de score op elk observatiepunt met de twee somscores berekend (vanwege het onvolledige karakter van het design voor de proeftoetsing was het niet mogelijk om in deze

⁴ Voor de opzet, uitvoering en rapportage van dit onderzoek zijn we Maaïke Foppen veel dank verschuldigd.

fase een factoranalyse uit te voeren; voor factor-analytische informatie op basis van de gegevens van het normeringsonderzoek verwijzen we naar hoofdstuk 6 over validiteit).

De selectie van observatiepunten werd bepaald door (een combinatie van) de volgende overwegingen.

Aantal observatiepunten en betrouwbaarheid

De omvang van de observatielijst mocht niet te groot zijn. Het instrument moest in de praktijk snel in te vullen zijn en een niet al te grote belasting voor de leerkracht zijn (gegeven het aantal kinderen in een doorsnee kleutergroep). Op basis van de gegevens van het proefonderzoek werd ingeschat dat *vijftien observatiepunten* per domein respectievelijk subschaal zouden volstaan om een goede betrouwbaarheid te garanderen. Bij een aantal van $2 * 15 = 30$ observatiepunten blijft de omvang van de lijst binnen de perken.

Hoge correlatie met andere items en met somscore binnen hetzelfde domein

Er werd in principe enerzijds gestreefd naar grote interne consistentie, waarbij anderzijds de iteminhouden elkaar zo min mogelijk zouden mogen overlappen. Op basis hiervan werden items verwijderd die inhoudelijk gezien overlap hadden met de andere items en items geselecteerd met een hoge item-resttotaalcorrelatie en een zo hoog mogelijke correlatie met de rest van de items in de observatielijst binnen hetzelfde domein (minimaal .40).

Spreiding van de antwoorden over de antwoordcategorieën

Eerder in deze verantwoording hebben we aangegeven vooral te hebben gekozen voor observatiepunten die betrekking hebben op de normale ontwikkeling als basis voor de itemconstructie, en niet voor abnormaal functioneren zoals bij motorische stoornissen. Daarmee hoopten we een zeer scheve verdeling over de antwoordcategorieën te voorkomen. De proefvoetsgegevens lieten zien dat dit voor een aantal observatiepunten goed gelukt is. Sommige items lieten niettemin een verdeling zien waarbij de minder wenselijke antwoordalternatieven erg weinig werden gebruikt. Daarnaast vormde de informatiewaarde een belangrijk criterium: observatiepunten met een zo hoog mogelijke informatiewaarde hadden de voorkeur.

Inhoudelijke overwegingen

Soms bleken meerdere items op basis van de eerder genoemde overwegingen in aanmerking te komen voor selectie, maar bleek de overlap met een of meer andere items groot. In dit soort gevallen werden dubbelgangers vermeden.

Daarnaast werd de dekking van elk domein breed gehouden. Soms werd een observatiepunt met wat minder uitmuntende empirische eigenschappen toch gekozen omdat een bepaald aspect van het domein in onze ogen niet in de subschaal mocht ontbreken.

Het hardop-denken-onderzoek en de evaluatieformulieren gaven aanvullende informatie over de aansluiting bij de gebruikerscontext. Ons streven was met de observatielijst zoveel mogelijk aan te sluiten bij bewegingssituaties die in de praktijk veel voorkomen, zodat leerkrachten een duidelijk beeld hebben van het functioneren van de kinderen op bepaalde observatiepunten. Als bepaalde taken uit de observatiepunten veel aan bod komen, dan is de kans groter dat de leerkracht hier een betrouwbaar beeld van heeft, dan wanneer een taak zelden wordt uitgevoerd. Daarnaast kost het gebruik van de observatielijst minder tijd wanneer leerkrachten niet eerst allerlei activiteiten moeten organiseren en observeren. Toch gaven leerkrachten soms aan dat ze bij bepaalde observatiepunten speciale activiteiten moesten organiseren om een beeld te krijgen van het functioneren van een leerling. Indien we binnen de dekking van het inhoudsdomein een keuze konden maken tussen observatiepunten, hebben we dit argument laten meespelen. Alhoewel 'stoeien' wel in de doelen van het Bewegingsonderwijs genoemd wordt, hebben we het observatiepunt dat hier betrekking op had verwijderd. Meerdere leerkrachten gaven namelijk aan "dit nooit te doen". Het observatiepunt kon in dat geval niet gescoord worden. Wanneer de formulering van een observatiepunt door de gebruikers als complex, ambigu of onduidelijk werd ervaren, is dit verwijderd.

Op basis van het proefonderzoek konden we het aantal observatiepunten reduceren tot een voor de praktijk aanvaardbaar aantal van 30 en de ordening in twee domeinen handhaven. Door psychometrisch optimale observatiepunten te selecteren wisten we twee subschalen te construeren met uitstekende eigenschappen en een hoge betrouwbaarheid (zie hoofdstuk 5). Daarbij is overbodige inhoudelijke overlap zoveel mogelijk

uit de weg gegaan. Essentieel is dat alle domeinen en aspecten binnen domeinen inhoudelijk zijn gedekt, waarbij de verdeling van de observatiepunten naar type bij benadering overeenkomt bij de vooraf geformuleerde wensen (zie daarvoor de toetsmatrijs die het uitgangspunt vormde voor de constructie van de observatiepunten).

3.2 Opbouw, afname en scoring van de observatielijst

Opbouw

Zoals in paragraaf 3.1 is beschreven heeft Motoriek kleuters een opbouw gekregen waarin twee dimensies (domeinen of subschalen) worden onderscheiden. Deze hebben de vorm van afzonderlijke subschalen voor grove en fijne motoriek.

1. Grove motoriek

De dimensie *Grove motoriek* betreft het functioneren waarbij leerlingen gecoördineerde grote, grove bewegingen met hun lichaam maken. De betreffende subschaal bestaat uit de volgende 15 observatiepunten.

1	Beweegt zich soepel (niet houterig) door de ruimte.	Beweegt zich niet soepel (maar houterig) door de ruimte.
2	Valt niet vaak.	Valt vaak.
3	Rent zonder overdreven meebewegen van de armen.	Rent met overdreven meebewegen van de armen.
4	Kan als tikker enkele kinderen tikken.	Kan als tikker niet enkele kinderen tikken.
5	Kan grote bewegingen met het lichaam imiteren (nadoen).	Kan grote bewegingen met het lichaam niet imiteren (nadoen).
6	Leert nieuwe bewegingen in de speelzaal vlot aan (bijvoorbeeld tijdens een bewegingscircuit).	Leert nieuwe bewegingen in de speelzaal niet vlot aan (bijvoorbeeld tijdens een bewegingscircuit).
7	Kan op een bank stappen zonder handen of hulp te gebruiken.	Kan niet op een bank stappen zonder handen of hulp te gebruiken.
8	Heeft geen moeite om in een klimparcours van een lager op een hoger vlak te stappen (van de bank op de kast).	Heeft moeite om in een klimparcours van een lager op een hoger vlak te stappen (van de bank op de kast).
9	Kan over een bank lopen (van minimaal 20 cm breed).	Kan niet over een bank lopen (van minimaal 20 cm breed).
10	Kan ergens overheen springen (blokken of touw).	Kan niet ergens overheen springen (blokken of touw).
11	Kan ergens vanaf springen (bank of traprede) en landen op twee voeten tegelijk zonder om te vallen.	Kan niet ergens vanaf springen (bank of traprede) en landen op twee voeten tegelijk zonder om te vallen.
12	Kan liggend van een schuin vlak naar beneden rollen (van een mat of helling buiten).	Kan niet liggend van een schuin vlak naar beneden rollen (van een mat of helling buiten).
13	Kan een grote bal vangen met twee handen voor het lichaam.	Kan een grote bal niet vangen met twee handen voor het lichaam.
14	Kan een kleine bal van dichtbij ergens ingooien (bijvoorbeeld in een bak).	Kan een kleine bal niet van dichtbij ergens ingooien (bijvoorbeeld in een bak).
15	Kan recht en stabiel op een stoel zitten tijdens het tekenen/schrijven.	Kan niet recht en stabiel op een stoel zitten tijdens het tekenen/schrijven.

2. Fijne motoriek

De dimensie *Fijne motoriek* heeft betrekking op de kleinere gecoördineerde bewegingen van handen, vingers en pols. De betreffende subschaal bestaat eveneens uit 15 observatiepunten.

16	Maakt de ritssluiting van een jas zelfstandig dicht.	Maakt de ritssluiting van een jas niet zelfstandig dicht .
17	Heeft geen moeite met openen van sluitingen van verpakkingen van eten en drinken (broodtrommel, rietje, dop beker).	Heeft moeite met openen van sluitingen van verpakkingen van eten en drinken (broodtrommel, rietje, dop beker).
18	Kan blokjes van klein constructiemateriaal (zoals lego) aan elkaar bevestigen.	Kan blokjes van klein constructiemateriaal (zoals lego) niet aan elkaar bevestigen.
19	Is in staat om houten blokjes nauwkeurig tegen elkaar te schuiven (zodat ze precies aansluiten).	Is niet in staat om houten blokjes nauwkeurig tegen elkaar te schuiven (zodat ze precies aansluiten).
20	Is in staat om kleine kraaltjes beheerst en precies op een kralenplank te plaatsen.	Is niet in staat om kleine kraaltjes beheerst en precies op een kralenplank te plaatsen.
21	Boetseert klei door soepel te duwen, drukken en kneden met duim en vingers.	Boetseert klei niet door soepel te duwen, drukken en kneden met duim en vingers.
22	Slaagt erin om een vouwblaadje precies dubbel te vouwen met de punten op elkaar.	Slaagt er niet in om een vouwblaadje precies dubbel te vouwen met de punten op elkaar.
23	Is in staat om een vouwblaadje dubbel te vouwen, zonder dat het papier erg kreukt.	Is niet in staat om een vouwblaadje dubbel te vouwen, zonder dat het papier erg kreukt.
24	Brengt de juiste hoeveelheid lijm aan (en niet te veel of te weinig).	Brengt niet de juiste hoeveelheid lijm aan (te veel of te weinig).
25	Is in staat om een eenvoudig figuur uit te knippen.	Is niet in staat om een eenvoudig figuur uit te knippen.
26	Kan met een prikpen vlak naast elkaar gaatjes prikken.	Kan met een prikpen niet vlak naast elkaar gaatjes prikken.
27	Kan eenvoudige vormen tekenen (cirkel, driehoek, vierkant).	Kan geen eenvoudige vormen tekenen (cirkel, driehoek, vierkant).
28	Kan kleine, eenvoudige vormen egaal inkleuren.	Kan kleine, eenvoudige vormen niet egaal inkleuren.
29	Drukt niet te hard of te zacht op het potlood tijdens het tekenen/schrijven.	Drukt te hard of te zacht op het potlood tijdens het tekenen/schrijven.
30	Kan met een rijpe drie- of vierpuntsgreep een potlood vasthouden.	Kan niet met een rijpe drie- of vierpuntsgreep een potlood vasthouden.

Aard van het instrument: afname-instructies

Motoriek kleuters is een volledig computergestuurd observatie-instrument. Om het te kunnen gebruiken dienen het Computerprogramma LOVS en het instrument Motoriek kleuters geïnstalleerd te zijn (zie gebruikershandleiding). De observatielijst is uitsluitend digitaal af te nemen. Via het computerprogramma kunnen leerlingen worden ingepland, vragenlijsten worden ingevuld en rapportages opgevraagd. De observatielijst is genormeerd voor de groepen 1 en 2 van het reguliere basisonderwijs. Bij afname in het speciaal (basis)onderwijs dient men zich te realiseren dat men de resultaten van een kind vergelijkt met de scores van kinderen in het reguliere onderwijs. Motoriek kleuters kan gedurende het hele jaar worden ingevuld. Ons advies is om de lijst na een aanlooperperiode in het begin van het schooljaar in te vullen, bijvoorbeeld in oktober. Zo kan de leerkracht zijn aanbod snel aanpassen aan de behoeften van de kinderen in zijn groep. Uiteraard moet de leerkracht wel voldoende tijd hebben gehad om zich een beeld te kunnen vormen van het motorisch functioneren van de leerling. Het wordt aanbevolen om de observatielijst

in principe één keer per schooljaar in te zetten. Bij kinderen die ‘afwijkend’ of ‘zeer afwijkend’ scoren ten opzichte van de gemiddelde score van leeftijdsgenoten uit de normgroep, adviseren we om het instrument na een half jaar nog eens in te vullen.

Leerkrachten krijgen bij het invullen het advies om niet uit te gaan van een algemene, gunstige of ongunstige indruk van een leerling. Ieder paar tegengestelde uitspraken moet gezien worden als een op zichzelf staande beschrijving van een specifiek aspect van het motorisch functioneren van een leerling. Daarnaast wordt afgeraden om veel leerlingen achter elkaar of in een specifieke volgorde te beoordelen. Dit kan tot gevul hebben dat zich een bepaalde tendens tot beoordelen ontwikkelt. Duobaan-leerkrachten vullen de observatielijst bij voorkeur in overleg in; bij gezamenlijk invullen ontstaat een objectiever beeld van het motorisch functioneren. Bij duidelijke verschillen van mening kunnen zij besluiten om het gedrag van de leerling (nogmaals) te observeren om op grond daarvan tot een definitieve score te komen.

De afname (het invullen) van de observatielijst vindt digitaal plaats. Leerkrachten dienen zich vooraf op de hoogte te stellen van de afname-instructie en deze minimaal één keer te hebben doorgenomen. De uitleg voor het invullen van de observatielijst voor de kleuters is in het afnameprogramma opgenomen.

De observatiepunten van de observatielijst voor kleuters worden op een vierpuntsschaal afgebeeld. Elk observatiepunt bestaat uit twee tegengestelde stellingen.

Voorbeeld van een observatiepunt op een vierpuntsschaal (niet in de observatielijst opgenomen)



De leerkracht bepaalt eerst welke van de twee stellingen het *meest* van toepassing is op de leerling waarvoor de lijst wordt ingevuld. Vervolgens bepaalt deze of die stelling *helemaal* van toepassing is of slechts *gedeeltelijk*. Als de stelling helemaal van toepassing is, dient het eerste of het vierde rondje te worden aangeklikt. Als de stelling gedeeltelijk van toepassing is, dient het tweede of het derde rondje te worden aangeklikt.

Omdat het invullen van een observatielijst intensief is, wordt aanbevolen niet meer dan vijf observatielijsten achter elkaar in te vullen. Het invullen van een observatielijst vraagt ongeveer 10 minuten per leerling.

Het is niet mogelijk om een observatiepunt over te slaan. De knop *verder* verschijnt pas als een antwoord gegeven is. Het is ook mogelijk om naar het volgende observatiepunt te navigeren door te dubbelklikken op het gekozen antwoord.

Scoring

De scoring van het instrument is volledig geautomatiseerd. Per observatiepunt worden de waarden 3, 2, 1 en 0 toegekend aan de gekozen antwoordalternatieven, waarbij de waarde 3 staat voor functioneren dat als positief wordt beoordeeld en 0 voor functioneren dat als (extreem) negatief wordt gezien. De resulterende waarden worden per dimensie (subschaal) opgeteld tot ruwe scores, die op hun beurt weer worden omgezet in percentielscores en niveauscores (zie hoofdstuk 4 voor de normering).

Daarnaast is het computerprogramma OPCAT gebruikt om de toegepaste scores voor de antwoordalternatieven te evalueren. Nergens bleek samenvoeging (hercodering) van antwoordalternatieven zinvol. Men kan in OPCAT ook optimale waarden toekennen aan de alternatieven; vraag is dan welke gewichten optimaal zijn. Bij verreweg de meeste items bleek de keuze van gelijke afstanden tussen de – aan de antwoordcategorieën – toegekende waarden (i.e. een gewicht van 1) de beste oplossing te geven. Op basis van deze resultaten kan gesteld worden dat met alle categorieën een betekenisvol verschil tussen de

kinderen aangegeven kan worden, en dat de ordening van de categorieën zoals verwacht is. Een aanpassing van de initiële scoring 3, 2, 1 en 0 bleek dus niet nodig of zinvol.

3.3 Meetmodel en psychometrische eigenschappen

Er zijn twee algemene raamwerken voor het ontwikkelen en analyseren van testen en toetsen, namelijk de klassieke testtheorie en de item respons theorie. Beide theorieën introduceren een aantal concepten en assumpties, en specificeren de relaties hiertussen in testmodellen. Bekende modellen binnen het raamwerk van de klassieke testtheorie zijn het klassieke testmodel, het Poisson model, en het *binomial error* model (zie bijvoorbeeld Gulliksen, 1950; Lord & Novick, 1968; Crocker & Algina, 1986). Daarnaast kan de generaliseerbaarheidstheorie gezien worden als een belangrijke uitbereiding op de klassieke testtheorie (Cronbach, Nanda & Rajaratnam, 1972). Binnen de item respons theorie zijn de belangrijkste modellen het 1-, 2-, en 3-parameter logistisch model, maar ook andere modellen zoals het *partial credit* model en het *rating scale* model zijn beschikbaar (zie bijvoorbeeld Rasch, 1960; Lord, 1980, Hambleton & Swaminathan, 1985; Van der Linden & Hambleton, 1997). Motoriek kleuters is geanalyseerd met behulp van het klassieke testmodel. Dit model past goed bij de aard van de data en een complexer IRT-model heeft bovendien in dit geval geen toegevoegde waarde.

Het klassieke testmodel beschrijft hoe meetfouten geobserveerde scores kunnen beïnvloeden. Het model verbindt de geobserveerde testscore van een leerling (X) aan de som van twee variabelen die niet geobserveerd zijn, namelijk de ware score (T) en de foutscore (E):

$$X = T + E$$

Omdat er twee onbekenden zijn in deze vergelijking is het niet mogelijk om de vergelijking op te lossen zonder verdere assumpties. Het klassieke testmodel veronderstelt dat: (a) ware scores en foutcores behaald door een bepaalde populatie leerlingen op één test niet gecorreleerd zijn, (b) foutcores op twee verschillende testen niet gecorreleerd zijn, en (c) de verwachte foutscore in een populatie leerlingen gelijk is aan 0. De ware score van een leerling is dus gelijk aan het verschil tussen de testscore en de foutscore. Bovendien kan aangetoond worden dat de ware score gelijk is aan de verwachte score die een leerling zou behalen indien een oneindig aantal testen zou worden afgenomen die exact hetzelfde meten. Aangezien het praktisch onmogelijk is om de ware score voor een leerling te bepalen door een oneindig aantal testen af te nemen onder precies dezelfde condities zijn er procedures ontwikkeld die het mogelijk maken om toch iets te zeggen over de gemiddelde meetfout in een specifieke populatie leerlingen. Denk hierbij bijvoorbeeld aan de coëfficiënt-alfa methode.

Testparameters die meetfout representeren en itemparameters die itemmoeilijkheid en itemdiscriminatie representeren vormen de basis van de meeste analyses binnen het klassieke testmodel. De parameters zijn gedefinieerd op een manier die gemakkelijk te begrijpen is. De parameters zijn bovendien waardevol gebleken bij de constructie van testen (zie bijvoorbeeld Henrysson, 1971; Millman & Greene, 1989). Een tekortkoming van de genoemde parameters is echter dat ze afhankelijk zijn van de steekproef op basis waarvan ze geschat zijn. De parameters beschrijven een kenmerk van een test of item namelijk uitsluitend in relatie tot een specifieke populatie. Voor meetnauwkeurigheid betekent dit dat dezelfde test betrouwbaarder is naarmate de variantie van de vaardigheid in de populatie waarin de test wordt afgenomen groter is. Voor itemmoeilijkheid betekent dit dat hogere waarden verkregen worden in steekproeven met een bovengemiddelde vaardigheid en lagere waarden in steekproeven met een beneden gemiddelde vaardigheid. Voor itemdiscriminatie, ten slotte, betekent dit dat hogere waarden verkregen worden in heterogene steekproeven en lagere waarden in homogene steekproeven (Fischer, 1974; Samejima, 1994). In de situatie dat de steekproef sterk afwijkt van de populatie waarvoor de test ontwikkeld wordt, zijn de parameters die verkregen worden na analyses met het klassieke testmodel daarom minder goed bruikbaar. Bij Motoriek kleuters hebben we hier geen last van. De observatielijst is ontwikkeld voor 4- en 5-jarigen en het onderzoek heeft ook bij deze leeftijdsgroep plaatsgevonden.

In deze paragraaf geven we, ter afsluiting van dit beschrijvende hoofdstuk, een empirische karakterisering van de subschalen op itemniveau. Andere beschrijvende gegevens komen aan de orde bij de hoofdstukken over normering (Hoofdstuk 4), betrouwbaarheid (Hoofdstuk 5) en validiteit (Hoofdstuk 6). In tabel 3.1 staat voor elk item de p -waarde, de r_{it} -waarde en de r_{ir} -waarde. De p -waarde is hier gedefinieerd als de proportie van het te behalen maximum (van 3) bij scores die variëren tussen 0 en 3. We zien dat leerlingen behoorlijk hoge scores behalen. Een p -waarde van .70 betekent bijvoorbeeld dat leerlingen gemiddeld 2.1 punten toegekend krijgen op het betreffende item. Item 15 bij Grove motoriek is een uitzondering. Waar leerlingen in de regel veel punten behalen (>2), ligt het gemiddelde aantal punten bij het laatste item van de subschaal voor grove motoriek aanzienlijk lager (<1). Dit is op zich geen probleem; de lage p -waarde impliceert slechts dat 4- en 5-jarigen vaak nog niet toe zijn aan het aspect dat bij item 15 geobserveerd wordt. Bij de r_{it} - en r_{ir} -waarden valt op dat deze ruimschoots boven de .30 liggen. Hoewel dit deels het gevolg is van de wijze waarop de items in de observatielijsten gescoord worden (i.e., vierpuntsschalen in plaats van dichotome goed-foutscores), wijzen de r_{it} - en r_{ir} -waarden op zeer goed onderscheidende items (zie Evers et al., 2010, p 40). Samenvattend laat tabel 3.1 dan ook zien dat de items in de observatielijsten geschikt zijn om zwak motorisch functionerende leerlingen te scheiden van goed motorisch functionerende leerlingen.

Tabel 3.1 Itemkenmerken Motoriek kleuters

Item	Grove motoriek			Fijne motoriek		
	p -waarde	R_{it} -waarde	R_{ir} -waarde	p -waarde	R_{it} -waarde	R_{ir} -waarde
1	.70	.75	.73	.55	.62	.59
2	.53	.60	.55	.55	.61	.59
3	.67	.72	.70	.60	.63	.64
4	.78	.82	.80	.61	.65	.65
5	.77	.80	.79	.69	.72	.72
6	.78	.81	.80	.61	.66	.63
7	.68	.72	.72	.76	.80	.79
8	.75	.79	.79	.74	.79	.77
9	.65	.68	.67	.68	.73	.69
10	.79	.82	.82	.81	.84	.83
11	.78	.82	.81	.75	.79	.78
12	.66	.71	.69	.77	.81	.80
13	.66	.72	.69	.77	.81	.79
14	.66	.71	.68	.69	.74	.71
15	.31	.40	.33	.64	.70	.66

4 Normering

4.1 Steekproeftrekking en deelname

In de schooljaren 2013/2014 en 2014/2015 zijn data verzameld met het oog op de normering van het instrument. De steekproef is getrokken met behulp van een *Simple Random Sampling* (SRS) procedure. Dat betekent dat elk element uit de steekproefkaderpopulatie een zelfde kans had om in de uiteindelijke normeringssteekproef terecht te komen. Het steekproefkader dat gebruikt is voor het onderzoek is gemaakt op basis van bestanden van DUO en bevatte 6683 scholen. Het streven was om ongeveer 900 vier- en vijfjarigen bij het onderzoek te betrekken. Er werd op gerekend dat deze leerlingen verspreid zouden zijn over ongeveer 30 scholen. Om de beoogde normeringssteekproef te realiseren zijn 90 scholen uit de lijst met 6683 scholen schriftelijk benaderd met de vraag om deel te nemen aan het onderzoek. Uiteindelijk hebben 27 scholen zich aangemeld voor deelname. Op deze scholen hebben uiteindelijk 621 leerlingen meegedaan. De gerealiseerde normeringssteekproef was dus kleiner dan vooraf beoogd werd. Daarom is bij de normering extra gelet op stabiliteit en nauwkeurigheid.

4.2 Representativiteit van de normeringssteekproef

In deze paragraaf besteden we aandacht aan de kenmerken die relevant zijn om de relatie tussen de normeringssteekproef en de populatie in kaart te brengen. Er is informatie verzameld over *regio*, *urbanisatiegraad*, *formatiegewicht* en *sekse*. De verschillende variabelen zijn als volgt gedefinieerd in het onderzoek:

- **Regio.** Bij de definitie van de variabele *Regio* is uitgegaan van de CBS-indeling naar landsdeel. Dit betekent dat er vier regio's onderscheiden zijn. Regio *Noord* omvatte de provincies Groningen, Friesland en Drenthe; regio *Oost* de provincies Overijssel, Gelderland en Flevoland; regio *West* de provincies Utrecht, Noord-Holland, Zuid-Holland en Zeeland en regio *Zuid* de provincies Noord-Brabant en Limburg.
- **Urbanisatiegraad.** De gehanteerde verdeling naar *urbanisatiegraad* of *mate van verstedelijking* is terug te voeren op de bij het CBS gebruikelijke indeling naar vijf niveaus, namelijk *zeer sterk*, *sterk*, *matig*, *weinig* en *niet verstedelijkt*. Voor het onderzoek is de gebruikelijke vijfdeling echter gereduceerd tot een tweedeling in enerzijds *niet tot matig verstedelijkt* (platteland) en anderzijds *sterk tot zeer sterk verstedelijkt* (stad). Een dergelijke tweedeling blijkt in de praktijk goed te volstaan (cf. Van Boxtel & Hemker, 2009).
- **Schooltype.** Bij de definitie van deze variabele is gebruikgemaakt van de formatiegewichten van de leerlingen binnen een school. Ten tijde van het onderzoek onderscheidde OCW drie niveaus die gebaseerd waren op het opleidingsniveau van de ouders:
 - 0.0 één van de ouders of beide ouders heeft of hebben een opleiding gehad uit categorie 3
 - 0.3 beide ouders of de ouder die belast is met de dagelijkse verzorging heeft of hebben een opleiding uit categorie 2 gehad
 - 1.2 één van de ouders heeft een opleiding gehad uit categorie 1 en de ander een opleiding uit categorie 1 óf 2

In deze indeling wordt verwezen naar de volgende categorieën in het opleidingsniveau van de ouders: 1 = maximaal basisonderwijs of (V)SO-ZMLK, 2 = maximaal LBO/VBO, praktijkonderwijs of VMBO basis- of kaderberoepsgerichte leerweg, en 3 = overig VO en hoger. Leerlingen met een formatiegewicht van 0.3 of 1.2 zijn te definiëren als achterstandsleerlingen. Scholen zijn ingedeeld naar

het percentage achterstandsleerlingen volgens een indeling in twee typen: (1) percentage achterstandsleerlingen [0, .15) en (2) percentage achterstandsleerlingen [.15, 1]. Dezelfde indeling wordt door de Inspectie van het Onderwijs gehanteerd bij de opbrengstbeoordeling in het basisonderwijs.

- **Sekse.** In de analyses omtrent *seks* is vanzelfsprekend een tweedeling naar jongens en meisjes gehanteerd.

Voor het bepalen van de populatiegegevens is gebruikgemaakt van gegevens van DUO (www.data.duo.nl) en CBS (statline.cbs.nl). De resultaten van de representativiteitsanalyse zijn te vinden in tabel 4.1. We zien dat de aantallen in onze normeringssteekproef soms afwijken van de aantallen die we in het licht van de populatiegegevens zouden verwachten. In feite is alleen de verdeling van jongens en meisjes in overeenstemming met die in de populatie. De westelijke regio's en de sterk tot zeer sterk verstedelijkte gebieden zijn ondervertegenwoordigd; de scholen met relatief veel achterstandsleerlingen zijn oververtegenwoordigd. Hoewel er geen reden is om aan te nemen dat de motoriek bij kinderen samenhangt met regio, urbanisatiegraad en formatiegewicht is dat op voorhand ook niet uit te sluiten. Daarom is in een vervolgstap een regressieanalyse verricht.

Tabel 4.1 *Uitkomsten representativiteitsanalyse*

Variabele	Categorie	Populatie	Steekproef		Toetsing			
		%	N	%	χ^2	df	p	ϕ
Regio	Noord	10.0	160	25.8	387.2	3	<.01	.79
	Oost	22.4	196	31.6				
	West	47.9	73	11.8				
	Zuid	19.7	192	30.9				
	Totaal	100.0	621	100.0				
Urbanisatie	Platteland	54.8	393	63.3	18.2	1	<.01	.17
	Stad	45.2	228	36.7				
	Totaal	100.0	621	100.0				
Schooltype	[0, .15)	20.6	191	30.8	39.4	1	<.01	.25
	[.15, 1]	79.4	430	69.2				
	Totaal	100.0	621	100.0				
Sekse	Jongen	50.5	314	50.6	<.01	1	.96	<.01
	Meisje	49.5	307	49.4				
	Totaal	100.0	621	100.0				

In een regressieanalyse gaan we na of de afwijkingen die we vonden in de representativiteitsanalyse er met het oog op de normering toe doen. Dat is het geval als de prestaties van groepen leerlingen significant van elkaar verschillen. Vanzelfsprekend moeten we bij het uitvoeren van de regressieanalyse onderscheid maken tussen grove en fijne motoriek. Tabel 4.2 laat zien in hoeverre prestatieverschillen tussen leerlingen te verklaren zijn vanuit de variabelen *regio*, *urbanisatiegraad*, *schooltype* en *seks*. De effecten zijn per

subschaal steeds in één analyse geschat. Dit betekent dat we in tabel 4.2 het unieke effect van een bepaalde variabele kunnen zien, gecorrigeerd voor de effecten van de overige variabelen.

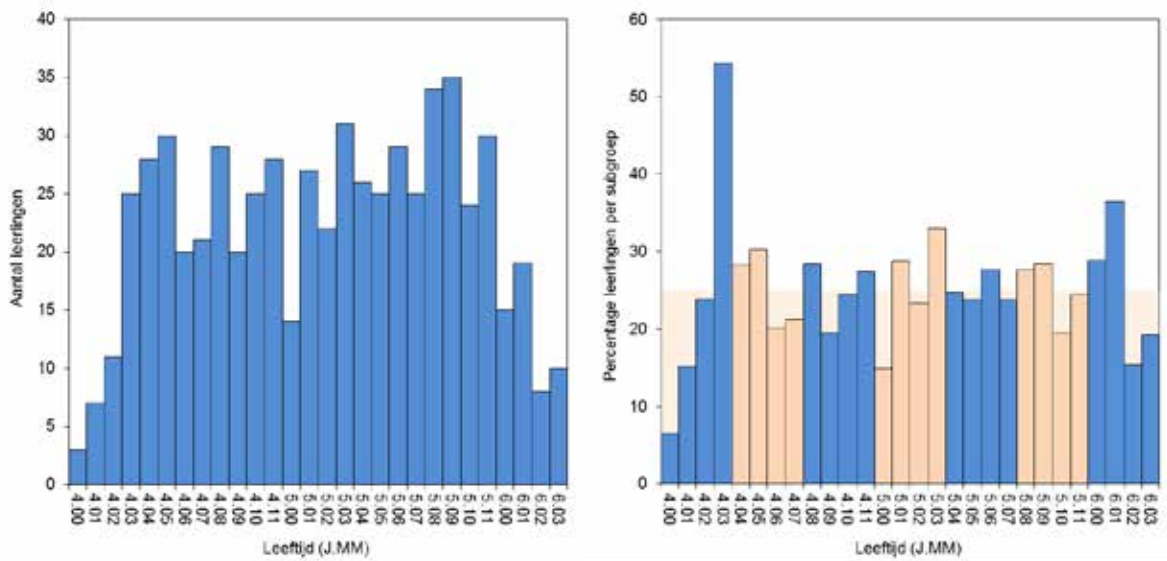
Met uitzondering van *seks* heeft bij zowel grove als fijne motoriek geen enkele variabele een betekenisvol effect op de prestaties van leerlingen. We zien weliswaar dat de plattelandsgebieden iets beter presteren dan de steden, en dat de prestaties in het oosten, westen en zuiden van Nederland iets achterblijven bij de prestaties in het noorden, maar geen van de verschillen is significant bij $\alpha = .05$. Zoals eerder aangegeven, is het verschil tussen jongen en meisjes wel relevant. In de groep met 4- en 5-jarigen hebben meisjes een betere motoriek dan jongens. Bij grove motoriek is er sprake van een klein tot gematigd verschil (Cohen's $d = .34$) en bij fijne motoriek van een gematigd tot groot verschil (Cohen's $d = .68$).

Tabel 4.2 *Uitkomsten regressieanalyse*

Variabele	Categorie	N	Grove motoriek			Fijne motoriek			
			effect	z	d	N	effect	z	d
Regio	Noord	160	----	----	----	160	----	----	----
	Oost	196	-1.38	-1.46	-.17	196	-1.59	-1.74	-0.20
	West	73	-.60	-.49	-.07	73	-2.21	-1.85	-0.28
	Zuid	192	-1.48	-1.46	-.18	192	-1.76	-1.79	-0.22
Urbanisatie	Platteland	393	----	----	----	393	----	----	----
	Stad	228	-1.37	-1.75	-.17	228	-0.62	-0.82	-0.08
Schooltype	[0, .15)	191	----	----	----	191	----	----	----
	[.15, 1]	430	.98	1.31	.12	430	-0.10	-0.14	-0.01
Sekse	Jongen	314	----	----	----	314	----	----	----
	Meisje	307	2.81	4.21	.34	307	5.47	8.46	0.68

Uit de representativiteitsanalyse en de regressieanalyse blijkt dat de normeringssteekproef geschikt is om landelijk representatieve normen op te stellen voor zowel grove als fijne motoriek. Er is geen reden om de onder- of overrepresentatie van bepaalde groepen leerlingen via statistische weging op te lossen. In bijna geen enkel geval verschillen de groepen immers significant in prestatie. In geval er wel sprake is van een significant prestatieverschil (i.e. bij seks) is de verdeling van leerlingen over de onderscheiden groepen voldoende in overeenstemming met de verdeling in de populatie. Naar één aspect is echter nog niet gekeken. Het is onduidelijk hoe de normeringssteekproef in elkaar zit qua leeftijdsopbouw. Als we één normtabel construeren voor 4- en 5-jarigen is het van belang dat de leerlingen in de normeringssteekproef zich zoveel mogelijk uniform verdelen over maandgroepen. Als er binnen de groep met 4- en 5-jarigen door toepassing van een vorm van continu normeren verfijning wordt aangebracht, is het van belang dat de leeftijd binnen elke subgroep uniform verdeeld is. Figuur 4.1 laat zien hoe de leeftijdsverdeling eruitziet. Links staat het aantal leerlingen per maandgroep. Rechts staat het percentage leerlingen binnen een subgroep als we uitgaan van een indeling in zeven “viermaandsgroepen”.

Figuur 4.1 Leeftijdsopbouw in de normeringssteekproef in aantallen (links) en percentages per subgroep (rechts)



In figuur 4.1 is te zien dat de leeftijd behoorlijk uniform verdeeld is over de maandgroepen. Alleen de laagste (≤ 4.02) en de hoogste (≥ 6.00) maandgroepen zijn ondervertegenwoordigd.

De ondervertegenwoordiging van de laagste maandgroepen is in de praktijk lastig te voorkomen. Een leerling start pas vanaf 4-jarige leeftijd in groep 1 en leerkrachten moeten een leerling eerst een paar weken observeren voordat ze de observatielijst in kunnen vullen. Daarnaast komt het voor dat leerlingen die nieuw zijn op school, tijd nodig hebben om te wennen, voordat ze een functioneringsniveau laten zien dat voor hen representatief is. Leerkrachten stellen het invullen van de observatielijst in dat geval in de regel enkele maanden uit. De hoogste maandgroepen bevinden zich in een schemergebied. Leerlingen in groep twee zijn doorgaans 5 jaar en worden in de loop van het schooljaar 6. De leerlingen vanaf 6 jaar behoren niet direct tot de doelgroep, maar door zittenblijven en een mogelijk vertraagde instroom van specifieke leerlingen (bijv. immigranten) zijn deze leerlingen toch, in zekere zin terecht, in de normeringssteekproef terechtgekomen. In het licht van de wijze waarop de dataverzameling heeft plaatsgevonden, komt de leeftijdsopbouw in onze normeringssteekproef vermoedelijk in hoge mate overeen met de leeftijdsopbouw die we landelijk in groep 1 en 2 zullen aantreffen. In de indeling met zeven viermaandsgroepen is de verdeling in zes van de zeven gevallen statistisch uniform te noemen bij $\alpha = .05$:

- 4.04-4.07 jaar $\chi^2(3, N = 99) = 3.02 \quad p = .39$
- 4.08-4.11 jaar $\chi^2(3, N = 102) = 1.92 \quad p = .59$
- 5.00-5.03 jaar $\chi^2(3, N = 94) = 6.85 \quad p = .08$
- 5.04-5.07 jaar $\chi^2(3, N = 105) = 0.41 \quad p = .94$
- 5.08-5.11 jaar $\chi^2(3, N = 123) = 2.43 \quad p = .49$
- 6.00-6.03 jaar $\chi^2(3, N = 52) = 5.69 \quad p = .13$

In de eerste viermaandsgroep gaat de aanname van een uniforme verdeling niet op:

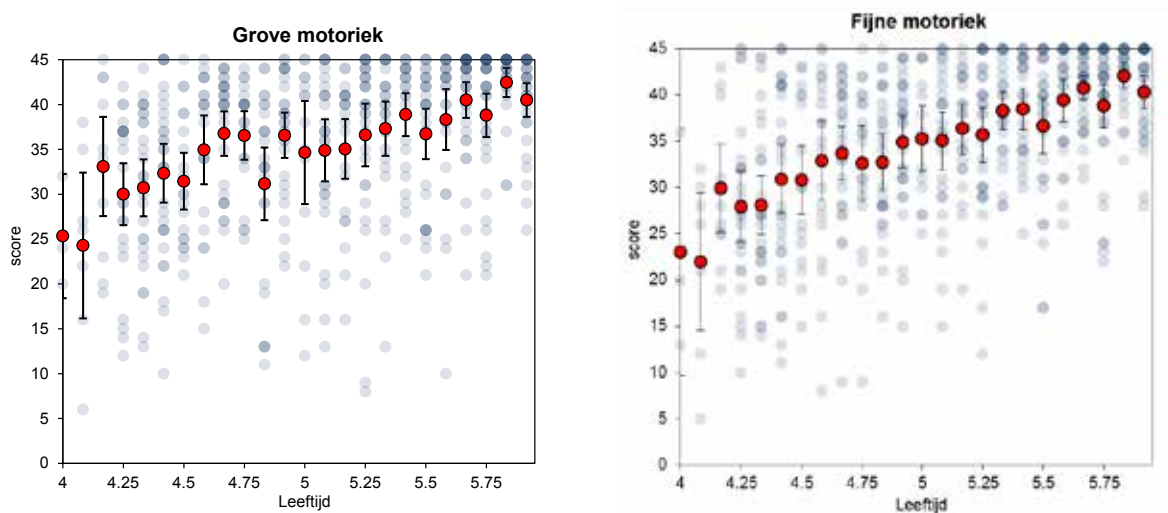
- 4.00-4.03 jaar $\chi^2(3, N = 46) = 23.91 \quad p < .01$

4.3 Prestaties als functie van leeftijd

Voorafgaand aan de vaststelling van de normen is uitgezocht in hoeverre het wenselijk is om binnen de groep met 4- en 5-jarigen subgroepen te onderscheiden met een eigen normering. Variabelen als *regio*, *urbanisatiegraad* en *schooltype* zijn niet relevant, omdat de motoriek van leerlingen zowel vanuit inhoudelijk als statistisch oogpunt gezien niet met deze variabelen samenhangt (zie paragraaf 4.2). Wel mag verwacht worden dat de motoriek van jonge kinderen zich dusdanig snel ontwikkelt dat een normering voor 4- en 5-jarigen te grofmazig is. Figuur 4.2 laat zien hoe ontwikkeling in grove en fijne motoriek verloopt bij leerlingen

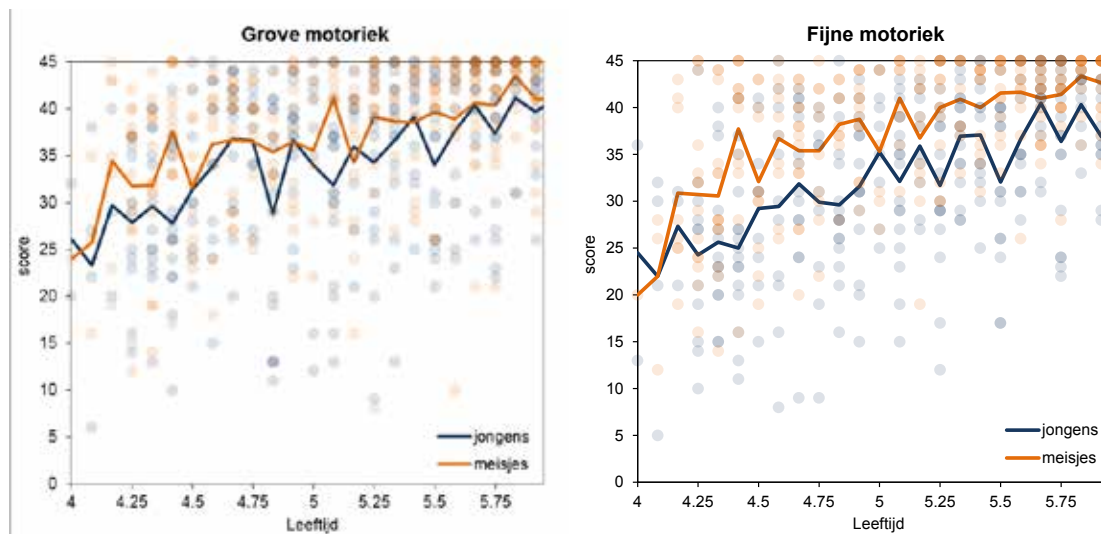
in de leeftijd van 4 en 5 jaar. De rode markeringen geven de gemiddelden en de bijbehorende 90%-betrouwbaarheidsintervallen weer. De blauwe markeringen representeren de scores zoals we die bij leerlingen geobserveerd hebben; een donkerdere markering wijst op een groter aantal leerlingen met die score. Bij zowel grove als fijne motoriek is er sprake van een duidelijke groei over de tijd. De jongste leerlingen behalen gemiddeld een score rond de 25 à 30. De oudste leerlingen behalen gemiddeld een score rond de 40. Aan de blauwe markeringen is te zien dat de spreiding afneemt over de tijd en dat het aantal leerlingen met de maximale score sterk toeneemt.

Figuur 4.2 *Ontwikkeling grove en fijne motoriek bij 4- en 5-jarigen*



Uit de eerder uitgevoerde regressieanalyse bleek dat meisjes bij zowel grove als fijne motoriek beter presteren dan jongens. Er kan daarom reden zijn om voor jongens en meisjes een andere normering te hanteren. Figuur 4.3 laat zien hoe de motoriek zich bij jongens en meisjes over de tijd ontwikkelt. Zoals verwacht mocht worden op basis van de regressieanalyse is het prestatieverschil bij grove motoriek relatief klein. Bij fijne motoriek is het prestatieverschil behoorlijk groot, namelijk gemiddeld meer dan 5 punten op een scoreschaal van 0 tot 45. De prestaties van jongens en meisjes lijken niet samen te hangen met de maandgroep. Meisjes presteren over de gehele linie consistent, met ongeveer hetzelfde verschil, beter dan jongens.

Figuur 4.3 Ontwikkeling grove en fijne motoriek bij jongens en meisjes



Of er in het licht van deze uitkomsten ook aparte normeringen moeten komen voor verschillende subgroepen is de vraag. Het is wenselijk om in de normering rekening te houden met de leeftijd van leerlingen, omdat we de relatieve prestaties van oudere leerlingen anders systematisch **overschatten** en de prestaties van jongere leerlingen systematisch **onderschatten**. Als het om het gesignaleerde prestatieverschil tussen jongens en meisjes gaat, ligt dit anders. Uit de wetenschappelijke onderzoeksliteratuur blijkt dat er bij aanzienlijk meer jongens sprake is van motorische ontwikkelingsproblemen dan bij meisjes. Zo komt een *Developmental Coordination Disorder* (DCD) vaker voor bij jongens dan bij meisjes. In diverse onderzoeken worden verhoudingen genoemd van 4 meisjes tegenover 10 jongens (zie www.balansdigitaal.nl). In het perspectief van dit gegeven ligt het niet voor de hand om in de normering rekening te houden met sekse. In dat geval zouden we het percentage jongens met een afwijkende prestatie bij motoriek immers kunstmatig gelijkstellen aan het percentage meisjes met een afwijkende prestatie bij motoriek. Dit is onwenselijk.

Bij de normering van motoriek is een indeling gemaakt in zes leeftijdsgroepen. De eerste vijf leeftijdsgroepen bestrijken elk een periode van 4 maanden, beginnend bij 4.00-4.03 jaar en eindigend bij 5.04-5.07 jaar. In de laatste leeftijdsgroep zitten de leerlingen van 5.08 jaar en ouder. Vanzelfsprekend hadden we in plaats van voor een indeling in viermaands groepen ook kunnen kiezen voor een fijnere of grovere groepsindeling. Er bestaat geen theoretisch optimum. Er is enerzijds geprobeerd om de ontwikkeling van leerlingen op een realistische manier te verdisconteren in de normering. Anderzijds is gestreefd naar een zo groot mogelijke steekproefomvang per leeftijdsgroep.

Tabel 4.3 laat voor zowel grove als fijne motoriek zien hoe de leerlingen gepresteerd hebben. In de eerste vier kolommen staan respectievelijk de minimaal en maximaal haalbare score, het gemiddelde en de standaarddeviatie. Daarna volgen de skewness en kurtosis. Ten slotte wordt met behulp van de Lilliefors (Kolmogorov-Smirnov) test en de Anderson-Darling test nagegaan of aangenomen kan worden dat de scores afkomstig zijn uit een normale verdeling. We zien dat de gemiddelde score toeneemt naarmate leerlingen zich in een hogere leeftijdsgroep bevinden. Daarnaast valt op dat de scoreverdelingen over de tijd meer van een normale verdeling gaan afwijken. De scoreverdelingen zijn in toenemende mate linksscheef en meer gepiekt dan bij een normale verdeling het geval zou zijn. Dit beeld wordt bevestigd door de statistische toetsen. Met uitzondering van de eerste leeftijdsgroep is de aanname van een normale verdeling statistisch gezien niet verdedigbaar. Er is sprake van een plafondeffect.

Tabel 4.3 Beschrijvende gegevens naar leeftijdsgroep

Grove motoriek									
Leeftijd	N	min.	max.	M	SD	skewness	kurtosis	LKS	AD
4.00-4.03	46	0	45	29.57	9.33	-.45	-.63	.11 ($p = .17$)	.50 ($p = .21$)
4.04-4.07	99	0	45	32.25	8.60	-.40	-.73	.10 ($p = .01$)	1.13 ($p = .01$)
4.08-4.11	102	0	45	35.29	7.97	-1.10	.70	.16 ($p < .01$)	3.27 ($p < .01$)
5.00-5.03	94	0	45	35.46	9.33	-1.15	.51	.17 ($p < .01$)	4.17 ($p < .01$)
5.04-5.07	105	0	45	37.76	7.58	-1.29	1.22	.19 ($p < .01$)	5.47 ($p < .01$)
≥ 5.08	175	0	45	40.19	5.85	-1.64	2.21	.21 ($p < .01$)	12.49 ($p < .01$)
Totaal	621	0	45	36.21	8.46	-1.10	.53	.16 ($p < .01$)	23.98 ($p < .01$)
Fijne motoriek									
Leeftijd	N	min.	max.	M	SD	skewness	kurtosis	LKS	AD
4.00-4.03	46	0	45	27.15	9.65	.04	-.46	.09 ($p < .46$)	.32 ($p = .53$)
4.04-4.07	99	0	45	30.53	9.41	-.18	-.95	.11 ($p = .01$)	1.05 ($p = .01$)
4.08-4.11	102	0	45	33.59	8.01	-.83	.28	.13 ($p < .01$)	1.77 ($p < .01$)
5.00-5.03	94	0	45	35.62	7.67	-.79	.13	.12 ($p < .01$)	1.67 ($p < .01$)
5.04-5.07	105	0	45	38.16	6.49	-.99	.58	.15 ($p < .01$)	2.97 ($p < .01$)
≥ 5.08	175	0	45	39.47	5.95	-1.50	1.90	.18 ($p < .01$)	8.71 ($p < .01$)
Totaal	621	0	45	35.36	8.51	-.90	.14	.13 ($p < .01$)	16.32 ($p < .01$)

4.4 Normeringsmethodiek

De resultaten die we in paragraaf 4.3 hebben beschreven, hebben op twee manieren consequenties voor de wijze waarop we kunnen normeren. Ten eerste is het vanwege het plafondeffect, en de daarmee samenhangende scheefheid van de scoreverdelingen, niet opportuun om in de normtabel veel scoreklassen te onderscheiden. In feite differentieert het instrument alleen goed bij zwakkere leerlingen. Dit past bij het gebruiksdoel van de observatielijst. Uit onderzoek blijkt dat bij 5 tot 10 procent van de schoolgaande kinderen sprake is van problemen in het motorisch functioneren. De observatielijst is bedoeld om de groep met een afwijkende (lage) score bij motoriek op te sporen en de normering moet zich dan ook daarop richten. Ten tweede ligt het vanwege de geobserveerde vaardigheidsgroei en het relatief kleine aantal leerlingen in elke leeftijdsgroep voor de hand om de informatie van de verschillende leeftijdsgroepen in samenhang te analyseren en te gebruiken bij de normering. Uit onderzoek is bekend dat een dergelijke aanpak efficiëntere schatters oplevert dan de traditionele cross-sectionele aanpak, ook bij een kleiner leerlingaantal (Bechger, Hemker & Maris, 2009; Roid, 1992). In de literatuur wordt gesproken over een minimumaantal van 50 tot 70 leerlingen per groep (Zhu & Chen, 2011; Evers et al., 2010; Wilkins et al., 2005). Dit komt goed overeen met onze steekproefomvang.

De aanpak waarbij gebruikgemaakt wordt van gegevens van meerdere leeftijdsgroepen tegelijkertijd is in de literatuur bekend als *continu normeren* (zie bijvoorbeeld Zachary & Gorsuch, 1985). Er zijn verschillende varianten van continu normeren. In de regel is het proces uiteen te leggen in drie stappen. In de eerste stap wordt op basis van de geobserveerde scores voor elke leeftijdsgroep het gemiddelde en de standaarddeviatie uitgerekend. Vervolgens worden de gemiddelden en de standaarddeviaties geplot en benaderd door een polynomiale curve. Vaak wordt een lineair verband verondersteld, maar er kan ook gekozen worden voor een hogere orde polynoom. Ten slotte worden de theoretische gemiddelden en standaarddeviaties afgeleid uit de geschatte curven en gebruikt bij het construeren van normtabellen.

In beginsel is er niets mis mee om de geobserveerde scoreverdelingen voor leeftijdsgroepen glad te strijken op basis van de scoreverdelingen van andere leeftijdsgroepen. In vergelijking met de traditionele aanpak zou de methodiek zelfs accuratere normen moeten opleveren (Wilkins et al., 2005).

Normeringssteekproeven die gebaseerd zijn op leeftijd of jaargroep hangen immers nauw met elkaar samen en het is onvoordelig om die samenhang te negeren. Niettemin brengt een continue normering ook risico's met zich mee. De efficiëntie bij continue normering wordt namelijk verkregen via aannamen en het is onzeker of die aannamen geldig zijn. Er wordt bijvoorbeeld aangenomen dat binnen iedere leeftijdsgroep een normale verdeling geldt, waarbij er alleen een verschil is in gemiddelde en standaarddeviatie. Deze aanname wordt in de praktijk met de nodige regelmaat geschonden (Marsman, 2012). Daarnaast wordt verondersteld dat de gemiddelden en standaarddeviaties voor de verschillende leeftijdsgroepen even nauwkeurig te schatten zijn. Als het leerlingaantal varieert over de leeftijdsgroepen is hiervan geen sprake. Ten slotte zal geen enkele curve de geobserveerde data perfect weten te benaderen. Het is in de praktijk niet altijd duidelijk welke curve vanuit inhoudelijk en statistisch oogpunt de voorkeur verdient. Wel is bekend dat de standaardfout van het gemiddelde en de andere momenten van de scoreverdeling groter wordt naarmate we een polynoom van een hogere orde kiezen.

Uit tabel 4.3 blijkt dat de aannamen die in de regel gedaan worden bij continu normeren in ons geval niet verdedigbaar zijn. De vorm van de scoreverdeling verandert met de leeftijd en in bijna geen enkele leeftijdsgroep is de scoreverdeling als "normaal" te karakteriseren. Bovendien beschikken we niet in elke leeftijdsgroep over data van evenveel leerlingen. Mogelijk kan de gangbare aanpak werken bij het schatten van (theoretische) gemiddelden, maar bij het schatten van (theoretische) percentielpunten gaat het onvermijdelijk mis als we uitsluitend naar de gemiddelden en de standaarddeviaties kijken, en uitgaan van een normale verdeling. Daarom is gezocht naar een aanpak waarin we beter rekening kunnen houden met (a) de vorm van scoreverdelingen en (b) het aantal waarnemingen binnen een leeftijdsgroep. We zijn terechtgekomen bij een specifieke vorm van regressieanalyse die in de literatuur ook wel bekend staat als kwantielregressie (Koenker, 2005). In deze aanpak benaderen we niet de samenvattende maten, zoals als het gemiddelde en de standaarddeviatie, met een polynoom, maar schatten we op basis van de geobserveerde data het gewenste (theoretische) percentieelpunt direct als functie van de leeftijdsgroep.

In vergelijking met de gangbare werkwijze bij continu normeren, brengt kwantielregressie enkele belangrijke voordelen met zich mee. Ten eerste houdt de analyse er automatisch rekening mee als het leerlingaantal niet in elke leeftijdsgroep even groot is. In geval we met behulp van de kleinste-kwadratenmethode een regressielijn zouden trekken door de gemiddelden of de standaarddeviaties zou hiervoor een wegingsvariabele geconstrueerd moeten worden. Ten tweede zijn de schattingen van een kwantielregressieanalyse minder gevoelig voor *outliers* en afwijkingen van de normale verdeling. Ten slotte gaan we in een kwantielregressie direct op zoek naar de informatie waarin we het meest geïnteresseerd zijn. In leerlingvolgsystemen willen toetsontwikkelaars en leerkrachten bijvoorbeeld graag weten in welk kwintiel of deciel een leerling valt. In selectietoetsen worden grensscores soms uitgedrukt in een percentiele rang. Deze informatie is met behulp van kwantielregressie direct te achterhalen, zonder daarbij specifieke aannames te doen over de vorm van de verdeling en/of de trend in gemiddelden en standaarddeviaties.

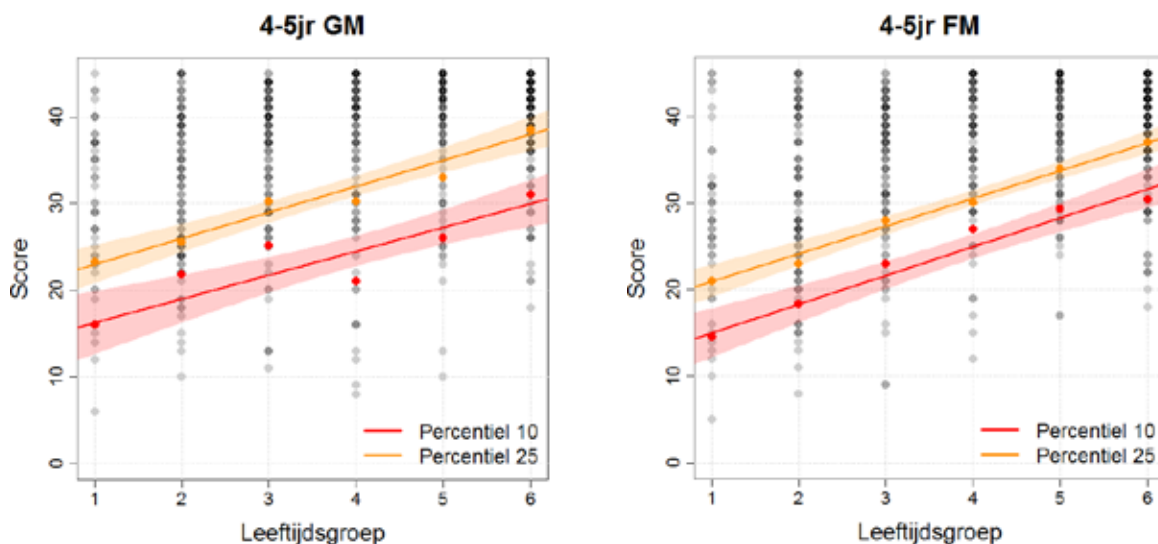
4.5 Vaststelling van de normen

Motoriek kleuters is bedoeld om leerlingen met een laag motorisch functioneringsniveau op te sporen. Er is besloten om drie scoreklassen te onderscheiden die we grofweg kunnen duiden als "niet afwijkend van de normgroep" (1), "afwijkend" (2) en "zeer afwijkend" (3). Dit betekent dat we per leeftijdsgroep twee grensscores moeten kiezen: één die de grens tussen klasse 1 en 2 representeert en één die de grens tussen klasse 2 en 3 representeert. De grensscores kunnen op verschillende manieren worden vastgelegd. Omdat het doel van de observatielijst is om leerlingen die op een laag niveau motorisch functioneren te identificeren, is de grens voor klasse 2 en 3 bij percentiel 10 gelegd. Bij het maken van het onderscheid tussen klasse 1 en 2 is uitgegaan van percentiel 25. Met behulp van kwantielregressie zijn beide percentieelpunten geschat als functie van leeftijdsgroep. Er zijn verschillende pakketten beschikbaar

waarmee we een kwantielregressieanalyse kunnen uitvoeren. Wij hebben gebruikgemaakt van het package *quantreg* in R (Koenker, 2015).

Bij het schatten van het model is een lineair verband verondersteld en gebruikgemaakt van een aangepaste versie van het algoritme van Barrodale en Roberts voor l_1 -regressie (1973). Voor meer informatie over dit algoritme, zie Koenker en d'Orey (1987, 1994). Het 95%-betrouwbaarheidsinterval voor de percentiel-10- en percentiel-25-lijn is gebaseerd op de geschatte covariantiematrix voor de parameterschattingen. Figuur 4.4 geeft een visuele representatie van de uitkomsten van de kwantielregressieanalyse. Op de achtergrond worden in het zwart, per leeftijdsgroep, de geobserveerde scores van de leerlingen weergegeven. Net als eerder in figuur 4.2 en 4.3 wijst een donkerdere markering op een groter aantal leerlingen met die score. De rode en oranje markeringen representeren respectievelijk het 10^{de} en 25^{ste} percentiel in elke leeftijdsgroep. De rode en oranje lijn geven het 10^{de} en 25^{ste} percentiel weer als functie van de leeftijdsgroep. De lijnen gaan vergezeld van een 95%-betrouwbaarheidsinterval. We zien dat de regressielijnen de geobserveerde punten behoorlijk goed weten te benaderen, zeker als we bedenken dat deze vanwege de kleine steekproeven behept zijn met een zeer grote standaardfout. De aanname dat er een lineair verband bestaat tussen de leeftijdsgroep en percentiel 10 en 25 lijkt zonder meer verdedigbaar; figuur 4.4 geeft op geen enkele manier aanleiding om voor een hogere orde polynoom te kiezen.

Figuur 4.4 Visuele weergave uitkomsten kwantielregressieanalyse. Links: grove motoriek. Rechts: fijne motoriek.



De grenzen voor de drie scoreklassen kunnen we, gegeven de uitkomsten van de kwantielregressieanalyse, afleiden uit de volgende functies:

Grove motoriek

$$P25 = 20 + 3t$$

$$P10 = 13.5 + 2.75t$$

Fijne motoriek

$$P25 = 17.8 + 3.2t$$

$$P10 = 11.7 + 3.3t$$

waarin t de leeftijdsgroep representeert. In de praktijk betekent dit dat we leerlingen op basis van de grensscores van tabel 4.4 kunnen indelen in de groepen “niet afwijkend”, “afwijkend” en “zeer afwijkend”.

Tabel 4.4 Grensscores motoriek

Percentiel	Omschrijving	Leeftijdsgroep					
		4.00-4.03	4.04-4.07	4.08-4.11	5.00-5.03	5.04-5.07	≥ 5.08
<i>Grove motoriek</i>							
25	Afwijkend	23	26	29	32	35	38
10	Zeer afwijkend	16	19	21	24	27	30
<i>Fijne motoriek</i>							
25	Afwijkend	21	24	27	30	33	37
10	Zeer afwijkend	15	18	21	25	28	31

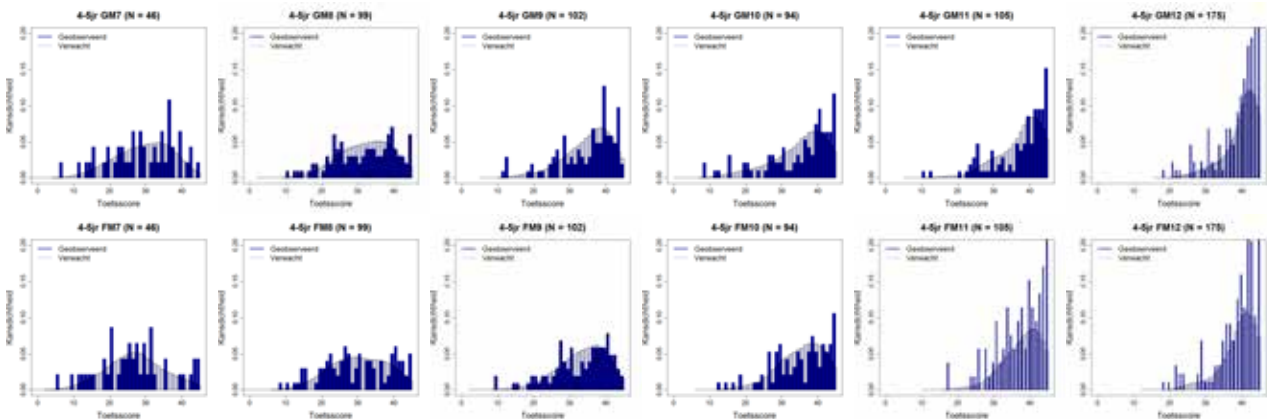
4.6 Nauwkeurigheid

Hoewel er geen reden is om aan te nemen dat de uitkomsten van de kwantielregressieanalyse niet adequaat zijn, is nog niet duidelijk hoe nauwkeurig de normering is. De nauwkeurigheid hangt af van de omvang van de leeftijdsgroepen. Evers et al. (2010) geven aan dat de groeps grootte bij voorkeur zo gekozen wordt dat de nauwkeurigheid van de normering hetzelfde is als bij een traditionele normering die, als het om minder belangrijke beslissingen op individueel niveau gaat, gebaseerd is op minimaal 200 (voldoende) of 300 (goed) leerlingen. Er is onderzocht met welk leerlingaantal in de traditionele normering onze steekproefomvang equivalent is. Doorgaans wordt gekeken naar de nauwkeurigheid van het gemiddelde, maar dat is in het geval van motoriek van ondergeschikt belang. Voor het onderscheid tussen “zeer afwijkend” en “afwijkend” kijken we immers naar percentiel 10 en 25. We hebben ons in het onderzoek dan ook gericht op het vaststellen van de nauwkeurigheid van deze twee percentiepunten in de scoreverdeling.

Het package *quantreg* (Koenker, 2015) biedt de mogelijkheid om op basis van het geschatte model waarden te voorspellen. Indien gewenst kan bij elke voorspelling ook de standaardfout bepaald worden. De standaardfout is op verschillende manieren te schatten. Er is voor gekozen om de standaardfouten te baseren op de geschatte covariantiematrix voor de parameterschattingen. Het is relatief eenvoudig om de standaardfout te bepalen voor onze normeringssteekproef. Het is aanzienlijk lastiger om dit zelfde te doen voor een hypothetische steekproef met 200 of 300 leerlingen. Dat komt doordat de scores zeer specifieke verdelingen kennen die over de tijd steeds meer gaan afwijken van een normale verdeling. In dit licht is het niet realistisch om als referentie een standaardfout te kiezen die berekend is onder aanname van een normale verdeling. Bij voorkeur berekenen we de standaardfout op basis van een (theoretische) verdeling die (a) gebaseerd is op 200 tot 300 waarnemingen en (b) qua gemiddelde, standaarddeviatie, kurtosis en skewness overeenkomt met de geobserveerde verdeling.

De standaardfouten voor de hypothetische steekproeven met 200 of 300 leerlingen zijn in twee stappen bepaald. In de eerste stap zijn voor elke leerling, op basis van zijn geobserveerde score en de meetfoutvariantie, 2500 *plausible values* gegenereerd. Een *plausible value* is een trek uit de volledige range aan scores die een leerling behaald zou kunnen hebben, gegeven zijn itemantwoorden en de betrouwbaarheid van de toets (Mislevy, 1991). Omdat we voor elke leerling meerdere *plausible values* genereren, wordt het mogelijk om de onregelmatigheden in de geobserveerde scoreverdelingen op een realistische manier glad te strijken. Figuur 4.5 laat zien hoe de (theoretische) gladgestreken scoreverdelingen zich verhouden tot de geobserveerde scoreverdelingen.

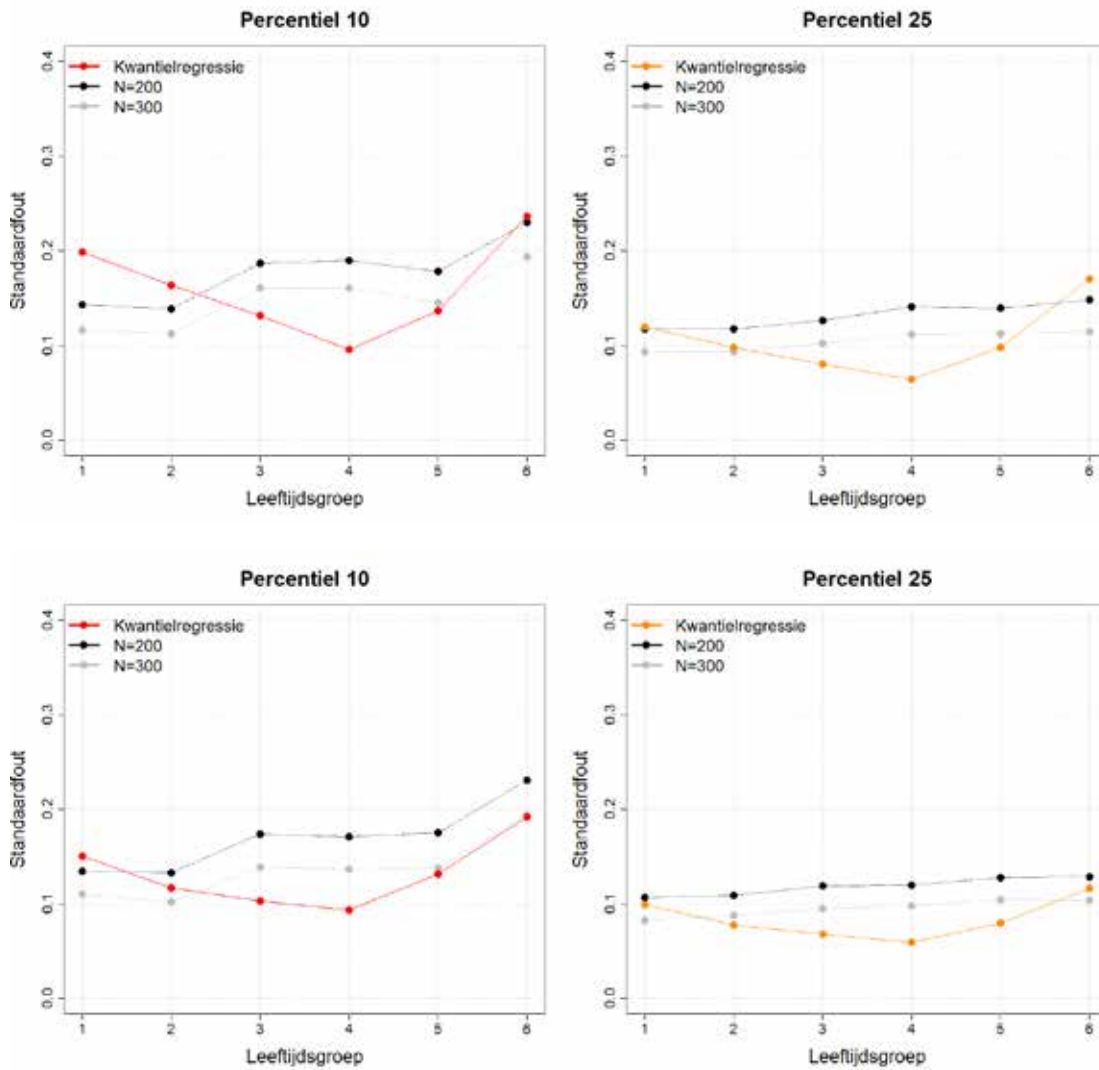
Figuur 4.5 Visuele weergave van de geobserveerde en de theoretische, gladgestreken scoreverdelingen. Boven: grove motoriek. Onder: fijne motoriek.



In de tweede stap is per leeftijdsgroep, voor zowel grove als fijne motoriek, 100 keer een steekproef getrokken met 200 of 300 leerlingen. In totaal beschikten we na deze stap dus over 2400 mogelijke scoreverdelingen, namelijk voor 2 observatielijsten \times 6 leeftijdsgroepen \times 2 steekproefgroottes \times 100 replicaties. Voor elke van deze 2400 scoreverdelingen is de standaardfout van percentiepunt 10 en 25 bepaald. Vervolgens is voor elke subschaal en leeftijdsgroep (normgroep) de gemiddelde standaardfout uitgerekend. Deze gemiddelden zijn als referentie gebruikt in de vergelijking tussen de traditionele en de continue normering. Er is bewust voor gekozen om de standaardfout voor een hypothetische steekproef met 200 of 300 leerlingen te schatten in een *bootstrap* procedure. We verkrijgen dan namelijk een goede benadering van de verwachte standaardfout en niet een standaardfout die als gevolg van een ongelukkige of gunstige steekproef toevallig een keer klein of groot uitvalt.

Figuur 4.6 laat zien hoe nauwkeurig de normering is in de verschillende leeftijdsgroepen. De rode en oranje lijnen geven de standaardfout weer voor respectievelijk percentiepunt 10 en 25 bij toepassing van een continue normering. De grijze en zwarte lijn geven de standaardfout weer als we op de traditionele manier normeren op basis van steekproeven met 200 of 300 leerlingen. Om een realistische vergelijking over leeftijdsgroepen mogelijk te maken, zijn de standaardfouten afgebeeld op een schaal met een gemiddelde van 0 en een standaarddeviatie van 1. We zien dat de nauwkeurigheid van de normering in leeftijdsgroepen 2 tot en met 5 in bijna in alle gevallen gelijk of groter is dan in een traditionele normering met 300 leerlingen. In de twee uiterste leeftijdsgroepen zijn de resultaten, zoals verwacht mag worden, iets minder gunstig. Soms zouden we een bepaald percentiepunt op basis van een traditionele normering met 200 leerlingen nauwkeuriger kunnen bepalen dan op basis van de continue normering die wij toegepast hebben. Dit geldt met name in de eerste leeftijdsgroep bij percentiepunt 10. Hoewel dit onwenselijk is, moet hierbij wel bedacht worden dat leerkrachten de observatielijst in de eerste leeftijdsgroep relatief weinig zullen toepassen (zie paragraaf 4.2). Dit betekent dat verreweg de meeste leerlingen adequaat geclassificeerd zullen worden op basis van de normering die is opgesteld. Gemiddeld gezien heeft de continue normering schatters opgeleverd die even nauwkeurig zijn als de verwachte nauwkeurigheid in een traditionele normering met 300 leerlingen.

Figuur 4.6 *Standaardfout naar leeftijdsgroep bij toepassing van een continue normering of een traditionele met 200 of 300 leerlingen. Boven: grove motoriek. Onder: fijne motoriek.*



5 Betrouwbaarheid

5.1 Generaliseerbaarheid van scores over versies

De betrouwbaarheid van Motoriek kleuters is vanuit verschillende invalshoeken te benaderen. Ten eerste kunnen we kijken naar de generaliseerbaarheid van scores over versies. We baseren de betrouwbaarheid in dat geval op de covarianties tussen de items in het instrument. In de literatuur worden verschillende methoden voorgesteld. Cronbachs alfa (1951) wordt verreweg het meest gebruikt, maar er zijn methoden beschikbaar die een betere schatting geven van de werkelijke betrouwbaarheid. Een voorbeeld daarvan is Guttman's lambda2 (Guttman, 1945). Hoewel het verschil klein is, kan aangetoond worden dat de waarde van Guttman's lambda2 dichterbij de werkelijke betrouwbaarheid ligt dan Cronbachs alfa. Eventueel kan de betrouwbaarheid ook uitgedrukt worden in termen van de greatest lower bound (Ten Berge & Sočan, 2004). In principe geeft de greatest lower bound een nog betere (onder)schatting van de betrouwbaarheid, maar een nadeel is dat de schattingen te hoog kunnen uitvallen in steekproeven kleiner dan duizend leerlingen en bij tests of toetsen met meer dan tien items (Sijtsma, 2009). Om deze reden is ervoor gekozen om Guttman's lambda2 als uitgangspunt te nemen en daarnaast ook Cronbachs alfa te berekenen. De waarden voor deze coëfficiënten staan in tabel 5.1. Ze zijn berekend voor zowel de volledige normeringssteekproef als voor de leeftijdsgroepen (normgroepen) die daarbinnen onderscheiden worden, voor grove en fijne motoriek afzonderlijk.

Tabel 5.1 Generaliseerbaarheid over versies

Leeftijd	Alfa	Lambda2	SE	90%BI	95%BI
<i>Grove motoriek</i>					
4.00-4.03	.925	.926	2.539	±4.176	±4.976
4.04-4.07	.926	.927	2.323	±3.821	±4.553
4.08-4.11	.917	.920	2.253	±3.706	±4.416
5.00-5.03	.944	.946	2.169	±3.568	±4.251
5.04-5.07	.933	.933	1.962	±3.227	±3.845
≥ 5.08	.908	.911	1.746	±2.872	±3.422
Totaal	.935	.936	2.140	±3.520	±4.194
<i>Fijne motoriek</i>					
4.00-4.03	.935	.936	2.441	±4.015	±4.783
4.04-4.07	.937	.937	2.363	±3.887	±4.631
4.08-4.11	.920	.921	2.250	±3.701	±4.410
5.00-5.03	.910	.913	2.262	±3.721	±4.434
5.04-5.07	.913	.916	1.880	±3.092	±3.684
≥ 5.08	.903	.908	1.806	±2.971	±3.540
Totaal	.936	.936	2.152	±3.541	±4.219

Tabel 5.1 laat zien dat de betrouwbaarheid zowel in de volledige normeringssteekproef als in de verschillende leeftijdsgroepen boven .90 ligt. Dit geldt niet alleen voor grove motoriek, maar ook voor fijne motoriek. Dit betekent dat de betrouwbaarheid van de twee observatielijsten als 'goed' gekwalificeerd kan worden (Evers et al., 2010).

Op basis van de betrouwbaarheidscoëfficiënten en de standaarddeviaties uit tabel 4.3 kan voor elke toetsscore de standaardmeetfout en het betrouwbaarheidsinterval uitgerekend worden. Deze informatie is te vinden in de laatste drie kolommen van tabel 5.1. Aangezien de analyses plaatsvinden binnen het

raamwerk van de klassieke testtheorie is verondersteld dat de standaardmeetfout en het betrouwbaarheidsinterval dat daaruit voortvloeit constant is voor het gehele scorebereik. Uit tabel 5.1 is af te leiden dat in geval een leerling van 4 jaar en 2 maanden bij grove motoriek een score van 25 behaald, deze bij een herhaalde observatie in 90 procent van de gevallen in het interval [20.82; 29.18] zal liggen. Voor alle andere leeftijdsgroepen en scores is het interval op exact dezelfde wijze af te leiden.

5.2 Generaliseerbaarheid van scores over leerkrachten

De motoriek van leerlingen wordt door middel van observaties in de klas bepaald. Naast de generaliseerbaarheid van scores over versies, is het daarom belangrijk om te kijken naar de generaliseerbaarheid van scores over leerkrachten. Dat kunnen we alleen doen als we bij (een deel van) de leerlingen beschikken over meer dan één beoordeling. In de opzet van het onderzoek is hier rekening mee gehouden. Aan de deelnemende scholen is gevraagd om voor een leerling twee observatielijsten in te vullen als er sprake was van een situatie waarin meer dan één leerkracht goed bekend is met het gedrag en functioneren van een groep. Dat is bijvoorbeeld het geval bij leerkrachten met een duobaan. Van de 621 leerlingen zijn er 113 door twee verschillende leerkrachten beoordeeld. De betreffende leerlingen bevonden zich in verschillende leeftijdsgroepen. In leeftijdsgroepen 2 tot en met 5 zaten 70 leerlingen die bij benadering evenredig verdeeld waren over de betreffende leeftijdsgroepen. Het aantal leerlingen in leeftijdsgroep 1 beperkte zich tot een totaal aantal van drie. In leeftijdsgroep 6 zaten 40 leerlingen. De observaties werden uitgevoerd in januari of februari. De tweede observatie vond in de regel enkele dagen na de eerste observatie plaats. De leerkrachten kregen de opdracht om de observatie onafhankelijk van elkaar uit te voeren.

De mate van generaliseerbaarheid over leerkrachten wordt in tabel 5.2 uitgedrukt als correlatie tussen de twee sets met beoordelingen. We zien dat de correlatie bij fijne motoriek iets hoger ligt dan bij grove motoriek. Duidelijke richtlijnen waaraan de correlaties moeten voldoen zijn er niet, maar in de regel worden waarden van .60 en hoger als voldoende aangemerkt. Bij Motoriek kleuters liggen de correlaties duidelijk boven deze "grenswaarde". Dit betekent dat de observatielijst de leerkracht voldoende houvast geeft bij het beoordelen van de motoriek van leerlingen. In de analyse is geen onderscheid gemaakt tussen leeftijdsgroepen, omdat het aantal leerlingen per groep te beperkt is om de correlatie betekenisvol uit te splitsen.

Tabel 5.2 Generaliseerbaarheid over leerkrachten

Observatielijst	N	Correlatie	95%BI
Grove motoriek	103	.762	[.667; .833]
Fijne motoriek	103	.825	[.751; .878]

5.3 Accuraatheid van classificaties

Motoriek kleuters wordt gebruikt om leerlingen in te delen in de groepen "niet afwijkend" (3), "afwijkend" (2) en "zeer afwijkend" (1). Op basis van de standaardmeetfouten in tabel 5.1 en de gladgestreken scoreverdelingen in figuur 4.5 kunnen we uitrekenen hoe accuraat een dergelijke classificatie is. We doen dat in de vorm van een zogeheten classificatie-/misclassificatietabel. Tabel 5.3 laat voor zowel grove als fijne motoriek zien in hoeverre de leerlingen in een bepaalde scoregroep ook op basis van hun werkelijke score in die scoregroep geassocieerd zouden worden. Bij grove motoriek zien we dat 69.1 procent van de leerlingen die met hun geobserveerde score in de categorie "zeer afwijkend" vallen, ook met hun werkelijke score in deze categorie vallen. Anders gezegd, de kans dat we een leerling met een leeftijd tussen 4.00 en 4.03 maanden terecht als zeer afwijkend bestempelen is ongeveer 69 procent. Verder laat de tabel zien dat

ruim 30 procent van de leerlingen in deze groep een score heeft die in werkelijkheid een scoregroep hoger ligt. Er zijn nauwelijks leerlingen die in werkelijkheid in de groep “niet afwijkend” vallen. De overige percentages zijn op exact dezelfde wijze te interpreteren. Zoals verwacht mocht worden op basis van de eerder gepresenteerde standaardmeetfouten neemt de accuraatheid van de classificatie iets toe naarmate leerlingen ouder worden.

Hoewel tabel 5.3 specifiek bedoeld is om zicht te krijgen op de lokale meetnauwkeurigheid van de observatielijsten kunnen we ook bepalen wat het gemiddelde percentage correcte classificaties is. We baseren ons dan op de diagonaal in elke classificatie-/misclassificatietabel en wegen deze voor het aantal leerlingen in de verschillende scoregroepen. Bij beide observatielijsten ligt het percentage correcte classificaties rond de 90 procent. Dit geldt voor alle leeftijdsgroepen.

Tabel 5.3 Classificatie-/misclassificatietabel voor grove en fijne motoriek uitgesplitst naar leeftijdsgroep

Leeftijd	Groep	Groep waarin de ware score voor grove motoriek valt			Groep waarin de ware score voor fijne motoriek valt		
		1	2	3	1	2	3
4.00-4.03	1	69.1	30.7	0.1	72.4	27.4	0.2
	2	5.5	70.5	24.1	7.3	66.9	25.8
	3	0.0	3.7	96.3	0.0	4.0	96.0
4.04-4.07	1	67.9	32.0	0.1	70.9	29.0	0.2
	2	4.6	72.7	22.7	6.9	68.8	24.4
	3	0.0	3.8	96.2	0.0	3.9	96.1
4.08-4.11	1	76.5	23.5	0.0	73.9	25.9	0.2
	2	4.2	72.6	23.1	5.4	67.0	27.6
	3	0.0	2.8	97.2	0.0	2.9	97.1
5.00-5.03	1	84.6	15.4	0.0	73.9	25.5	0.6
	2	5.4	75.8	18.7	7.4	61.9	30.8
	3	0.0	4.4	95.6	0.0	3.6	96.4
5.04-5.07	1	77.4	22.6	0.0	72.4	27.4	0.2
	2	4.2	78.2	17.7	5.5	67.7	26.8
	3	0.0	4.8	95.2	0.0	3.2	96.8
≥ 5.08	1	80.9	19.1	0.0	82.1	17.9	0.0
	2	3.0	74.9	22.1	4.5	69.4	26.2
	3	0.0	5.1	94.9	0.0	5.0	95.0

6 Validiteit

6.1 Inhoudsvaliditeit

De inhoudsvaliditeit van een toets of test heeft betrekking op de vraag in hoeverre de items in het instrument een welomschreven en afgebakend universum representeren van mogelijk in het instrument op te nemen items. De inhoudsvaliditeit van Motoriek kleuters wordt gewaarborgd door de wijze waarop het instrument ontwikkeld is. Om ervoor te zorgen dat de observatiepunten relevante en veelvoorkomende situaties in de praktijk vertegenwoordigen, zijn de volgende stappen gezet:

- analyse van methoden die gebruikt worden op peuterspeelzalen, kinderdagverblijven en groep 1 en 2 van het basisonderwijs waarin aandacht besteed wordt aan motoriek;
- analyse van de tussendoelen en leerlijnen van SLO voor zintuigelijke en motorische ontwikkeling;
- overleg met het Nederlands Instituut Sport en Beweging (NISB).

Op basis van bovenstaande is een toetsmatrijs ontwikkeld. In de toetsmatrijs is een onderverdeling aangebracht in hoofd- en subdomeinen. Bij de ontwikkeling van de observatielijsten is de toetsmatrijs gevolgd en is gestreefd naar een evenwichtige verdeling van observatiepunten over de onderscheiden hoofd- en subdomeinen.

In een vervolgstap zijn de ontwikkelde observatiepunten voorgelegd aan een expertgroep bestaande uit twee kinderfysiotherapeuten, een ervaren peuterleidster, een startende pedagogisch medewerkster, en drie (ex-) leerkrachten, van wie twee toetsdeskundigen en één educatief adviseur. Op basis van de opmerkingen van de expertgroep zijn enkele wijzigingen aangebracht in de aanvankelijke versie van de observatielijst voor grove en fijne motoriek. Deze versie is beproefd in een kleinschalig hardop-denkenonderzoek (Foppen, 2013). Tijdens dit onderzoek werd aan enkele peuterleidsters, pedagogisch medewerkers en leerkrachten gevraagd om een observatielijst “hardop-denkend” in te vullen. Op deze manier konden we achterhalen of de leerkrachten, leidsters en pedagogisch medewerkers de observatiepunten goed begrepen en of de observatiepunten wel direct meetbaar waren. In een aansluitend semi-gestructureerd interview konden de deelnemers opmerkingen maken over de observatiepunten en de observatielijst als geheel. Op basis van de informatie uit het hardop-denkenonderzoek is de definitieve observatielijst samengesteld.

Vanwege de grondige inhoudelijke analyse van bestaande methoden, tussendoelen en leerlijnen voor zintuigelijke en motorische ontwikkeling, en de nauwe afstemming met deskundigen uit de praktijk lijkt de inhoudsvaliditeit van Motoriek kleuters meer dan voldoende gewaarborgd. Een uitvoerige omschrijving van de theoretische en praktische uitgangspunten die ten grondslag hebben gelegen aan de constructie van de observatielijst is te vinden in hoofdstuk 2.

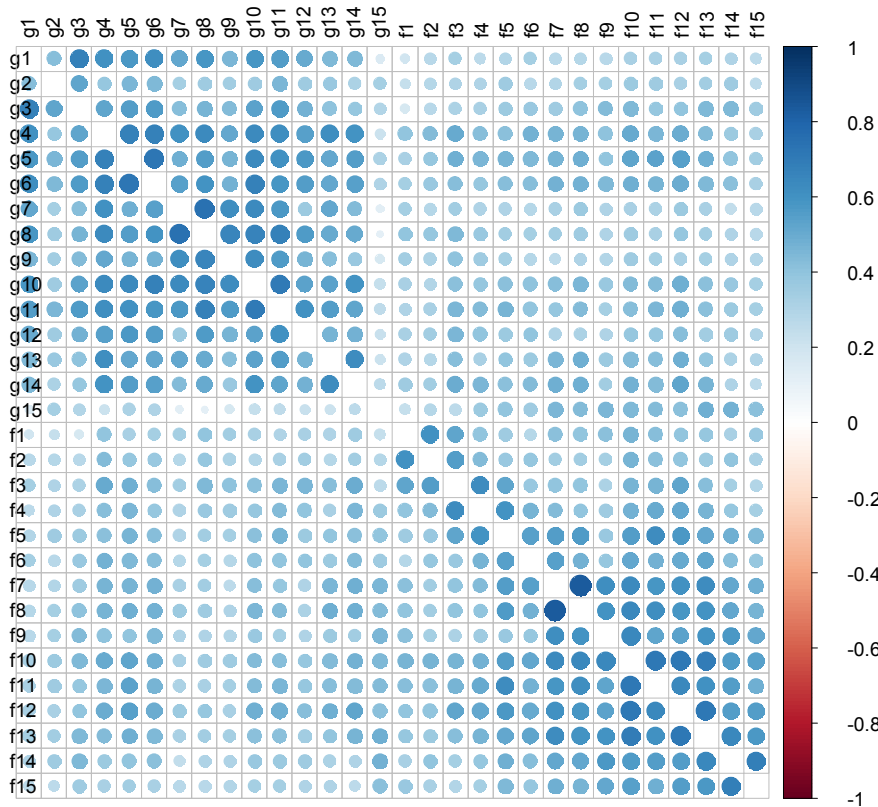
6.2 Begripsvaliditeit

De begripsvaliditeit van een instrument heeft betrekking op de vraag in hoeverre de scores op het instrument zijn toe te schrijven aan de verklarende concepten en constructen die deel uitmaken van het theoretische kader dat aan de ontwikkeling van het instrument ten grondslag heeft gelegen. Er is op verschillende manieren empirische evidentie verzameld die kan worden opgevat als onderbouwend voor de begripsvaliditeit van de observatielijst voor het motorisch functioneren van kleuters. We gaan achtereenvolgens in op de dimensionele structuur van de observatielijst, de samenhang van de observatielijst met soortgenoten, de samenhang van de observatielijst met een algemeen leerkrachtoordeel, en het functioneren van de observatielijst in verschillende subgroepen.

Dimensionele structuur

Om zicht te krijgen op de dimensionele structuur van het instrument met de subschalen voor grove en fijne motoriek is de samenhang tussen de verschillende items (of observatiepunten) in kaart gebracht. Omdat het instrument in een volledig afnamedesign is voorgelegd aan leerkrachten met betrekking tot ruim 600 leerlingen in groep 1 en 2 van het basisonderwijs zijn de inter-itemcorrelaties gemakkelijk uit te rekenen. Figuur 6.1 geeft een visuele representatie van de matrix met inter-itemcorrelaties. Op de horizontale en verticale as staan de items uit de subschalen: "g" staat voor grove motoriek en "f" voor fijne motoriek. In de figuur worden de correlaties weergegeven als een cirkel, waarbij de grootte en de kleur corresponderen met de hoogte en de richting van de correlatie. Naarmate de cirkel groter en donkerder gekleurd is, is de samenhang sterker. We zien dat de correlaties **binnen** grove en fijne motoriek hoger zijn dan **tussen** deze twee domeinen. Item 15 vormt een uitzondering; dit item vertoont relatief weinig samenhang met de overige items. Nader onderzoek liet zien dat leerlingen erg laag scoorden op dit item. Mogelijk wordt de lagere correlatie veroorzaakt door een *restriction of range*.

Figuur 6.1 Visuele representatie van de correlatiematrix voor motoriek



Om beter zicht te krijgen op de structuur van het instrument zijn in een vervolgstap de gemiddelde binnen-domein- en tussen-domeincorrelaties uitgerekend. Omdat correlaties niet zomaar bij elkaar opgeteld en gemiddeld mogen worden, is eerst Fisher's *r*-to-*z* transformatie toegepast. Dit betekent dat elke correlatie *r* is omgezet naar een *z*-score door gebruik te maken van de volgende vergelijking:

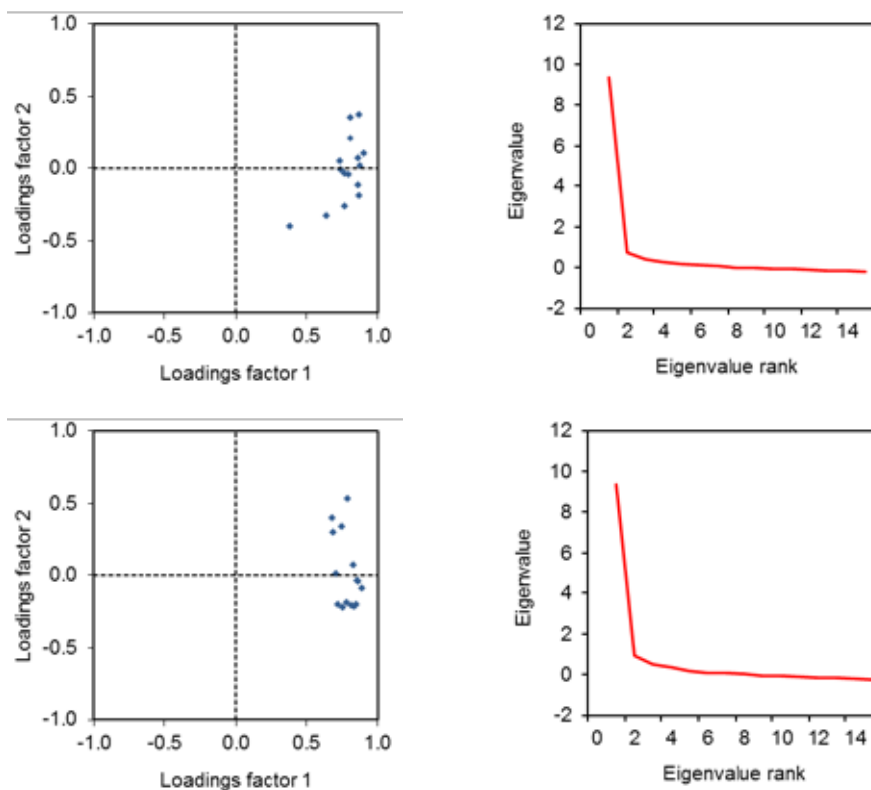
$$z = \frac{1}{2} \ln \left[\frac{(1+r)}{(1-r)} \right]$$

Vervolgens is de gemiddelde *z*-score bepaald en deze waarde is terugvertaald naar een correlatiecoëfficiënt. De items binnen het domein *grote motoriek* hangen gemiddeld .50 met elkaar samen.

De samenhang binnen het domein fijne motoriek is even sterk. De gemiddelde tussen-domeincorrelatie is gelijk aan .38. Deze uitkomst ondersteunt de eerder gemaakte keuze voor twee aparte subschalen voor grove en fijne motoriek.

In aansluiting op de berekening van de inter-itemcorrelaties is voor zowel grove als fijne motoriek een exploratieve *twee-factor principal axes* factoranalyse uitgevoerd. Bij beide subschalen ligt er duidelijk één dominante factor ten grondslag aan de prestaties van leerlingen; in beide gevallen is meer dan 90 procent van de variantie te verklaren vanuit de eerste factor. De factorladingen en de eigenwaarden bevestigen dit beeld. In figuur 6.2 is te zien dat alle items hoog laden op de eerste factor en aanzienlijk lager op de tweede factor. In de *screeplots* is bovendien een duidelijk knik waarneembaar na de eerste eigenwaarde. Er is, met andere woorden, geen enkele reden om aan te nemen dat de subschalen van Motoriek kleuters ieder appelleren aan meer dan één vaardigheid. Factoranalyse als puur formeel proces laat echter geen uitspraken toe over de vraag of het construct dat de observatielijst beoogt te meten inderdaad samenvalt met het construct *motoriek* zoals dat in het wetenschappelijk forum wordt bedoeld. In combinatie met de inhoudelijke waarborgen en de nauwe afstemming met deskundigen uit de praktijk lijkt het evenwel verdedigbaar om aan te nemen dat de observatielijst het motorisch functioneren van leerlingen meet in twee dimensies: grove en fijne motoriek.

Figuur 6.2 Screeplots en factorladingen voor grove (boven) en fijne (onder) motoriek



Relatie met soortgenoten

De mate waarin Motoriek kleuters gerelateerd is aan soortgenoten is een andere aanwijzing voor de begripsvaliditeit van het instrument. Om dit te kunnen onderzoeken moet niet alleen de observatielijst worden afgenomen, maar ook een of meer andere instrumenten die hetzelfde construct pretenderen te meten. Er zijn vrij veel instrumenten beschikbaar voor het beoordelen van de motorische ontwikkeling. Vier- en vijfjarigen vormen echter lang niet altijd de doelgroep en veel van deze instrumenten kunnen om deze reden dan ook niet als soortgenoot aangemerkt worden. Ditzelfde geldt voor instrumenten waarvan de validiteit onvoldoende is onderzocht. De mogelijkheden om soortgenootonderzoek uit te voeren zijn dan ook

beperkt. Met het oog op de validering is ervoor gekozen om de *Movement Assessment Battery for Children* als soortgenoot in te zetten, omdat dit één van de weinige instrumenten is die veelvuldig voldoende betrouwbaar en valide is gebleken, ook als het om leerlingen in groep 1 en 2 van het basisonderwijs gaat (Henderson, Sugden & Barnett, 2007; Schoemaker et al., 2012; Ellinoudis et al., 2010).

De *Movement ABC* bevat een checklist die bedoeld is om kinderen met motorische problemen op te sporen. De checklist bestaat uit 30 items die door de leerkracht, ouder of begeleider ingevuld worden. De items zijn onderverdeeld in twee secties. Deel A heeft betrekking op bewegingen in een stilstaande of voorspelbare omgeving. Deel B heeft betrekking op bewegingen in een dynamische of onvoorspelbare omgeving. Iedere sectie bestaat uit drie onderdelen die elk vijf items bevatten. Voor ieder item moet de leerkracht op een 4-puntschaal de motorische vaardigheid van de leerling beoordelen (0 = goed uitvoerbaar, 3 = nog lang niet uitvoerbaar). De score voor motoriek is gelijk aan de som van de scores op de dertig items. Hoe hoger deze score is, hoe zwakker de leerling motorisch functioneert.

De *Movement ABC* is bij 112 leerlingen in de steekproef afgenomen. Alle leeftijdsgroepen waren in deze onderzoeksgroep vertegenwoordigd. De onderzoeksgroep bestond voor 51,8 procent uit jongens en voor 48,2 procent uit meisjes. Bij de *Movement ABC* behaalden jongens gemiddeld een score van 23.17 ($SD = 13.56$) en meisjes een score van 14.35 ($SD = 11.86$). Uitgedrukt in een effectgrootte is dit verschil gelijk aan .69. Bij onderdeel A bleek het verschil tussen jongens en meisjes iets groter te zijn (Cohen's $d = .78$) dan bij onderdeel B (Cohen's $d = .54$). Dit patroon vertoont grote gelijkenis met het patroon dat we eerder vonden bij Motoriek kleuters (zie paragraaf 4.2); bij grove motoriek bleek er sprake te zijn van een klein tot gematigd verschil tussen jongens en meisjes (Cohen's $d = .34$) en bij fijne motoriek van een gematigd tot groot verschil (Cohen's $d = .68$). De correlatie tussen de *Movement ABC* en Motoriek kleuters is gelijk aan .86. De *Movement ABC* en het nieuw ontwikkelde instrument zijn dus in hoge mate aan elkaar gerelateerd en lijken inderdaad soortgenoten te zijn.

Relatie met een algemene beoordeling van de motoriek

De mate waarin de scores van de leerlingen samenhangen met gelijktijdig beschikbare prestatiegegevens is een derde aanwijzing voor de begripsvaliditeit van de observatielijsten. Om dit te kunnen onderzoeken is aan de deelnemende leerkrachten en pedagogisch medewerkers gevraagd om vóór het invullen van de observatielijsten hun indruk van het motorisch functioneren van de leerlingen uit te drukken in een cijfer van 0 tot 10. Een hoog cijfer wijst op goed motorisch functioneren en een lager cijfer op zwakker motorisch functioneren. We mogen een hoge correlatie verwachten in geval (a) de leerkrachten en pedagogisch medewerkers de items in de observatielijsten begrijpen en (b) het construct motoriek "op het oog" juist geoperationaliseerd is. Tabel 6.1 geeft de correlatie tussen het algemene leerkrachtoordeel en de beoordelingen op basis van de observatielijsten. Er is een uitsplitsing gemaakt naar leeftijdsgroep.

Tabel 6.1 Samenhang tussen Motoriek kleuters en een algemene beoordeling van de motoriek

Leeftijd	N	Grove motoriek		Fijne motoriek		Grove + fijne motoriek	
		Correlatie	95%BI	Correlatie	95%BI	Correlatie	95%BI
4.00-4.03	46	.739	[.571; .847]	.714	[.534; .832]	.786	[.642; .876]
4.04-4.07	99	.761	[.663; .833]	.789	[.701; .853]	.828	[.754; .881]
4.08-4.11	102	.732	[.627; .811]	.586	[.442; .701]	.750	[.650; .824]
5.00-5.03	94	.770	[.672; .841]	.806	[.721; .867]	.846	[.776; .895]
5.04-5.07	105	.686	[.569; .776]	.750	[.652; .823]	.780	[.692; .845]
≥ 5.08	175	.652	[.558; .730]	.614	[.512; .699]	.729	[.651; .792]
Totaal	621	.738	[.700; .772]	.714	[.673; .751]	.786	[.754; .814]

Tabel 6.1 laat zien dat de scores op Motoriek kleuters in hoge mate overeenstemmen met de intuïtie van de leerkrachten en pedagogisch medewerkers. De correlaties zijn zowel in de afzonderlijke leeftijdsgroepen als bij grove en fijne motoriek als behoorlijk hoog aan te merken.

Prestaties van relevante subgroepen

Een vierde aanwijzing voor de begripsvaliditeit van Motoriek kleuters kan verkregen worden door te kijken naar verschillen tussen relevante groepen leerlingen. In het hoofdstuk over de normering van de observatielijst zijn we hier al uitgebreid op ingegaan. Zoals op theoretische gronden verwacht mocht worden, hebben variabelen als regio, urbanisatiegraad en schooltype geen effect op het motorisch functioneren van leerlingen. Leeftijd en sekse hebben dat wel. Uit wetenschappelijk onderzoek blijkt dat de motoriek zich snel ontwikkelt bij 4- en 5-jarigen en dat meisjes, met name bij fijnmotorische taken, beter presteren dan jongens. Dit beeld wordt bevestigd in het onderhavige onderzoek (zie figuur 4.2 en 4.3).

Itemkwaliteit

Een vijfde en laatste aanwijzing voor de begripsvaliditeit van de observatielijsten kan gevonden worden in de kwaliteit van de items. In de laatste paragraaf van hoofdstuk 3 zijn we hier al uitgebreid op ingegaan. We zagen dat de p -waarden bij grove en fijne motoriek (zie tabel 3.1), op één item na, tussen .53 en .81 lagen. De gemiddelde p -waarde is bij grove motoriek .70 en bij fijne motoriek .68. De 'moeilijkheidsgraad' van de subschalen ligt daarmee precies op het niveau dat werd nagestreefd bij de constructie. De r_{it} - en r_{ir} -waarden liggen alle ruimschoots boven .30 met een gemiddelde rond .70. Dit wijst erop dat de items in de observatielijsten zeer geschikt zijn om zwak motorisch functionerende leerlingen te scheiden van goed motorisch functionerende leerlingen.

Op basis van de analyses die zijn verricht met betrekking tot de begripsvaliditeit van de observatielijsten voor motoriek kan het volgende geconcludeerd worden:

- Grove en fijne motoriek hangen sterk met elkaar samen, maar zijn tegelijkertijd ook dusdanig verschillend dat ze bij voorkeur afzonderlijk van elkaar beoordeeld worden.
- Aan zowel de subschaal voor grove motoriek als de subschaal voor fijne motoriek ligt één dominante vaardigheid ten grondslag.
- De scores die leerlingen behalen op de observatielijsten hangen sterk samen met de scores die zij behalen op een soortgenoot, namelijk de *Movement ABC*.
- Leerkrachten en pedagogisch medewerkers begrijpen de items in de observatielijst en vinden dat de observatielijst 'op het oog' de juiste inhoud bevat; dat blijkt uit de sterke samenhang tussen een algemeen leerkrachtoordeel en de somscores op de observatielijst.
- De observatielijsten zijn in staat om veranderingen in het motorisch functioneren van jongens en meisjes te detecteren.
- De items in de observatielijsten zijn zeer geschikt om zwak motorisch functionerende leerlingen te scheiden van goed motorisch functionerende leerlingen.

7 Samenvatting

In deze wetenschappelijke verantwoording staat het observatie-instrument Motoriek kleuters centraal, een digitaal af te nemen en te scoren observatielijst voor leerkrachten die bedoeld is om het motorisch functioneren van kinderen in groep 1 en 2 in het reguliere basisonderwijs in kaart te brengen. In de meetpretentie van het instrument wordt onderscheid gemaakt naar grove en fijne motoriek.

In hoofdstuk 2 werden de uitgangspunten voor de testconstructie besproken, in het bijzonder een algemeen ontwikkelingspsychologisch en onderwijskundig perspectief, met een focus op het normale motorische functioneren in de betreffende leeftijdsgroep. Ontwikkelingspsychologisch gezien ligt de nadruk op dynamische systeemtheorie zoals deze door Thelen en collega's is uitgewerkt. Hierbij werd op basis van de literatuur en een analyse van veelgebruikte methoden en instrumenten een onderscheid tussen grove en fijne motoriek gelegitimeerd als basis voor de structuur van de observatielijst. Onderwijskundig gezien werd vooral gebruikgemaakt van een analyse van leerlijnen en tussendoelen zoals deze in veel methoden worden gehanteerd en zijn uitgewerkt door de SLO.

In hoofdstuk 3 is in meer detail terug te vinden hoe de inhoud van Motoriek kleuters werd verantwoord op basis van een toetsmatrijs met hoofd- en subdomeinen. Dit hoofdstuk gaat ook in op een aantal belangrijke procedures die deel uitmaakten van het constructieproces:

- Ontwikkeling van observatiepunten in een constructiegroep van ter zake deskundigen.
- Inzetten van experts in de beoordeling van de observatiepunten.
- Uitvoeren van een uitgebreid proefonderzoek.
- Inzetten van een evaluatieformulier bij de deelnemende observatoren.
- Inzetten van hardop-denkonderzoek.
- Selectie van observatiepunten op basis van heldere psychometrische en inhoudelijke criteria, waarbij de ontworpen toetsmatrijs maatgevend was.

Daarnaast staat in dit hoofdstuk het instrument in zijn definitieve samenstelling inhoudelijk beschreven, naast een beknopte karakterisering van het gehanteerde meetmodel en de kwaliteit van de uiteindelijk opgenomen items.

In hoofdstuk 4 werd de wijze van normeren (in termen van percentielscores en de daaraan ontleende classificatie) gelegitimeerd. Op grond van de aangetroffen verschillen naar leeftijd en sekse werd gekozen voor zes normgroepen naar leeftijd ('viermaandsgroepen'); daarbij worden voor jongens en meisjes dezelfde normen gehanteerd. De normen zelf werden geconstrueerd op basis van een continu normeringsmodel waarbij non-parametrische kwantielregressie werd toegepast. Het hoofdstuk bevat de uiteindelijke normeringstabellen, in termen van de grensscores van de classificatiecategorieën in de zes normgroepen. Dat deze normering adequaat is, werd aangetoond op basis van representativiteitsanalyses en regressieanalyses met betrekking tot de vraag of de normeringssteekproef een adequate afspiegeling vormt van de populatie, en op basis van nauwkeurighedsanalyses met betrekking tot de continue normering (in vergelijking met een klassiek normeringsmodel).

Uit hoofdstuk 5 blijkt dat de betrouwbaarheid van Motoriek kleuters in termen van interne consistentie van de subschalen voor grove en fijne motoriek goed is (dat wil zeggen, voor alle normgroepen groter dan .90). Tevens werd daarbij inzicht gegeven in de betreffende standaardmeetfout en de betrouwbaarheidsintervallen. Aan de hand van classificatie-misclassificatietabellen lieten we zien wat de consequenties hiervan zijn voor de accuraatheid van de classificatie in de categorieën 'niet afwijkend', 'afwijkend' en 'zeer afwijkend'. De generaliseerbaarheid van scores over observatoren (i.e. twee leerkrachten in een duobaan) bleek zeer behoorlijk (met correlaties tussen twee onafhankelijke beoordelingen groter dan .75).

In hoofdstuk 6 besteedden we ruim aandacht aan de inhouds- en begripsvaliditeit van Motoriek kleuters (criteriumvaliditeit is bij dit instrument niet aan de orde). Op grond van de uitgebreide explicitering van de theoretische uitgangspunten en inhoudelijke verantwoording en door de toepassing van zorgvuldige

constructieprocedures en -onderzoek lijkt de inhoudsvaliditeit van het instrument ruim voldoende gewaarborgd. De dimensionele structuur van het instrument als belangrijk aspect van de begripsvaliditeit is goed te onderbouwen op basis van de correlaties tussen de items onderling (tussen en binnen de subschalen voor grove en fijne motoriek) en met de somscores op de subschalen. Factoranalyses lieten zien dat in elke subschaal één dominante factor ten grondslag ligt aan de scores. Dit wordt bevestigd door de uitstekende itemkwaliteit in termen van itemrestcorrelaties. Ook de 'moeilijkheidsgraad' van de items sluit aan bij de uitgangspunten en verwachtingen van de constructeurs. De samenhang tussen Motoriek kleuters en een soortgenoot (de Movement ABC) bleek met een correlatie van .86 hoog. Sekseverschillen bleken ten aanzien van de Movement ABC hetzelfde patroon te laten zien als in Motoriek kleuters. Scores op Motoriek kleuters bleken sterk samen te hangen met het algemene oordeel van de leerkrachten over het motorisch functioneren van de geobserveerde kinderen. We kunnen daaruit afleiden dat de leerkrachten een goed begrip hebben van de observatiepunten in Motoriek kleuters en dat het motorisch functioneren op een correcte wijze is geoperationaliseerd. Zoals op theoretische gronden verondersteld mocht worden, hebben variabelen als regio, urbanisatiegraad en schooltype geen effect op het motorisch functioneren van leerlingen. Leeftijd en sekse hebben dat – in overeenstemming met de verwachtingen – wel.

8 Literatuurverwijzingen

- Alkema, E., Dam, E. van, Kuipers, J., Lindhout, C. & Tjerckstra, W. (2006). *Méer dan onderwijs: theorie en praktijk van het onderwijs in de basisschool*. Assen: Van Gorcum B.V.
- Bechger, T. Hemker, B.T., & Maris, G. (2009). *Over het gebruik van continue normering*. Arnhem: Cito.
- Bilo, R.A.C. & Voorhoeve, H.W.A. (2008). *Kind in ontwikkeling. Een handreiking bij de observatie van jonge kinderen*. Amsterdam: Reed business.
- Boxtel, H. van., Hop, M. & Hemker, B. (2014). Wetenschappelijke verantwoording VISEON voor kleuters. Arnhem: Cito.
- Van Kuyk, J., Breebaart, D., & op den Kamp, M. (2012). *Piramide. Educatieve methode voor kinderen van nul tot zeven jaar. Wetenschappelijke verantwoording*. Arnhem: Cito.
- Cito. (2015). *Motoriek kleuters. Cito Volgstelsysteem primair en speciaal onderwijs*. Arnhem: Cito.
- Cito. (2012). *Piramide. CD werkdocumenten*. Arnhem: Cito
- Clark, J. E. & Whitall, J. (1989). What is Motor Development? The Lessons of History. *Quest*, vol. 41, no. 3, pp. 183-202, DOI 10.1080/00336297.1989.10483969
- Corstens-Mignot, M. A. A. M. G., Cup, E. H. C., & Hartingsveldt-Bakker, M. J. van. (2002). *Standaard Observatie Ergotherapie Schrijven en Sensomotorische Schrijfvoorwaarden; SOESSS theorie en praktijk*. Utrecht: Lemma.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart and Winston.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16, 297–334.
- Cronbach, L.J., Gleser, G.C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.
- Empelen, R. van. Nijhuis-van der Sanden, R. & Hartman, A. (2000). *Kinderfysiotherapie*. Maarssen: Elsevier Gezondheidszorg.
- Eggen, T.J.H.M. & Sanders, P.F. (1993). *Psychometrie in de praktijk*. Arnhem: Cito.
- Ellinoudis, T., Evaggelinou, C., Kourtessis, T., Konstantinidou, Z., Venetsanou, F., & Kambas, A. (2010). *Reliability and validity of age band 1 of the Movement Assessment Battery for Children—Second Edition. Research in developmental disabilities*, 32, 1046-1051.
- Evers, A., Lucassen, W., Meijer, R. & Sijtsma, K. (2010). *COTAN Beoordelingssysteem voor de kwaliteit van tests*. Amsterdam, NIP/COTAN.
- Fischer, G.H. (1974). *Einführung in die Theorie psychologischer Tests*. Bern: Huber.
- Foppen, M. (2013). *Kwalitatief onderzoek naar de Observatielijsten Motoriek voor Peuters en Kleuters*. Arnhem: Cito.

- Gesell, A. & Amatruda, C. Strunk. (1947). *Developmental Diagnosis*. New York: Hoeber.
- Greven, J. & Letschert, J. (2006). *Kerndoelenboekje*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley.
- Guttman, L. (1945). A basis for analyzing test-retest reliability. *Psychometrika*, 10, 255–282.
- Ten Berge, J.M.F., & Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, 69, 613–625.
- Hambleton, R.K., & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer.
- Henderson, S.E., Sugden, D.A., & Barnett, A.L. (2007). *Movement assessment battery for children-2: Movement ABC-2: Examiner's manual*. Pearson.
- Henrysson, S. (1971). Gathering, analyzing and using data on test items. In: R.L. Thorndike (Ed.), *Educational measurement*. Washington, DC: American Council on Education.
- Hermanns, J., Nijnatten, C., Verheij, F. & Reuling, M. (2008). *Handboek jeugdzorg*. Houten: Bohn Stafleu van Loghum.
- Hettinga, H. (2011). *Schrijven is meer dan bewegen alleen*. 's-Hertogenbosch: Uitgeverij Malmberg, Praxisbulletin 2011-2012, nummer 4.
- Johnson, M.H. (2011). *Developmental cognitive neuroscience: an introduction*. Chichester (UK): Wiley-Blackwell.
- Keulen, M. & Eerd-Smetsers, van C. (2007). *Van kleutertekening tot schrijven*. Utrecht/Zutphen: ThiemeMeulenhoff.
- Koenker, R. (2005). *Quantile Regression*. Cambridge University Press, Cambridge.
- Koenker, R. (2015). Quantreg: Quantile regression. R package version 5.19.
<https://cran.r-project.org/web/packages/quantreg/>
- Koenker, R.W. & d'Orey (1994). Computing regression quantiles. *Applied Statistics*, 43, 410–414.
- Kurtz, L. A. (2009). *De motoriek van kinderen met dyspraxie, autisme, ADHD en leerstoornissen*. Huizen: Pica.
- Kwast, A. D. (2006). *Schrijven kun je leren: de theorie van het schrijven*. Alkmaar: Stichting Stylus et Cultura.
- Lord, F.M., & Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.
- Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.

- Marsman, M., Maris, G. & Bechger, T. (2012). Don't Tie Yourself to an Onion: Don't Tie Yourself to Assumptions of Normality. In: T.J.H.M. Eggen & B.P. Veldkamp (Eds.), *Psychometrics in practice at RCEC*. Enschede: RCEC.
- Largo, R.H., Kunda, S., & Thun-Hohenstein, L. (1993). Early motor development in term and preterm children. In: A.F. Kalverboer, B. Hopkins & R. Geuze (Eds.), *Motor development in early and later childhood: Longitudinal approaches* (pp. 247-265). Cambridge: Cambridge University Press.
- Louwe, J. L. (2012). *Het OntwikkelingsVolgModel: verantwoording en achtergrondinformatie*. Amersfoort: uitgeverij Agiel.
- Millman, J., & Greene, J. (1989). The specification and development of tests of achievement and ability. In: R.L. Linn (Ed.). *Educational Measurement* (3rd ed., pp. 335–366). Washington, DC: American Council on Education.
- Mislevy, R. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56, 177-196.
- Netelenbos, J.B. (1998). Motorische ontwikkeling van kinderen. *Handboek 1: introductie*. Amsterdam: Boom.
- Netelenbos, J.B. (2000). Motorische ontwikkeling van kinderen. *Handboek 2: theorie*. Amsterdam: Boom.
- Noordstar, J. (2009). *Motoriek in de basisschool*. Baarn: Bekadidact.
- Nouws, L. (1985). *Van kleuterkrabbel tot schrijftaal*. Tilburg: Uitgeverij Zwijsen b.v.
- Onderwijsraad (2010). *Naar een nieuwe kleuterperiode in de basisschool*. Den Haag: Onderwijsraad.
- Piek, J.P. (2006). *Infant motor development*. Champaign: Human Kinetics.
- Poot, H. (2009). Bewegen en zintuigelijk ervaren. In: Singer, E. & Kleerekoper, L. (2009). *Pedagogisch kader kindercentra 0-4 jaar*. Maarsen: Elsevier Gezondheidszorg.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research.
- Roid, G. H. (1992, April). *Smoothing score distributions using moments: Applications in the norming of tests*. Paper presented at the annual meeting of the National Council on Measurement in Education, San Francisco, CA.
- Sameroff, A. (2010). A unified theory of development: a dialectic integration of nature and nurture. *Child Development*, 81, 6-22.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, 74, 107–120.
- Samejima, F. (1994). Estimation of reliability coefficients using the test information function and its modifications. *Applied Psychological Measurement*, 18, 229–244.
- Schoemaker, M. M., Niemeijer, A. S., Flapper, B. C., & Smits-Engelsman B. (2012). Validity and reliability of the Movement Assessment Battery for Children-2. Checklist for children with and without motor impairments. *Developmental Medicine & Child Neurology*, 54, 368-375.

- Schweitzer, H. (2000). *Schrijven zonder pen: ideeën voor de ontwikkeling van de schrijfmotoriek*. Baarn: uitgeverij Bekadidact.
- SLO. (2008). *TULE-Bewegingsonderwijs, Inhouden en activiteiten bij de kerndoelen van 2006*. Enschede: SLO.
- Someren, M.W., Barnard, Y.F. & Sandberg, J.A.C. (1994). *The think aloud method: a practical guide to modelling cognitive processes*. Londen: Academic Press.
- Thelen, E., & Smith, L.B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
- Van Boxtel, H., & Hemker, B.T. (2009). *Wetenschappelijke verantwoording van de Intelligentietest Eindtoets Basisonderwijs*. Arnhem: Cito.
- Van der Linden, W.J., & Hambleton, R.K. (1997). *Handbook of modern item response theory*. New York City, NY: Springer-Verlag.
- Wilkins, C., Rolffhus, E., Weiss, L., & Zhu, J. J. (2005, April). *A new method for calibrating translated tests with small sample sizes*. Paper presented at the 2005 annual meeting of the American Educational Research Association, Montreal, Quebec, Canada.
- Zachary, R. A., & Gorsuch, R. L. (1985). Continuous norming: Implications for the WAIS-R. *Journal of Clinical Psychology, 41*, 86-94.
- Zhu, J. & Chen, H. (2011). Utility of Inferential Norming With Smaller Sample Sizes. *Journal of Psychoeducational Assessment, 29*, 570-580.

Websites

CBS statline.cbs.nl
DUO www.data.duo.nl
www.balansdigitaal.nl
<http://tule.slo.nl>
<http://www.slo.nl/primair/themas/jongekind/doelen/>

Cito helpt je inzicht te krijgen in je ontwikkeling en mogelijkheden. Door kennis, vaardigheden en competenties objectief meetbaar te maken en de ontwikkeling er van te volgen, kun je het beste uit jezelf halen, verantwoorde keuzes maken en beter richting geven aan je toekomst. Cito draagt daaraan bij door wereldwijd werk te maken van goed en eerlijk toetsen, vanuit de kernwaarden kundig, toonaangevend, integer, innovatief en betrokken.

Cito

Amsterdamseweg 13
Postbus 1034
6801 MG Arnhem
T (026) 352 11 11
www.cito.nl

Fotografie: Ron Steemers