

Evaluatie van de sociale dimensie van inclusief onderwijs: de ontwikkeling van een signaleringsinstrument voor leerkrachten om de sociale participatie van leerlingen met beperkingen in kaart te brengen¹

M. Koster, E. J. van den Bosch en S. J. Pijl

Samenvatting

Dit onderzoek is gericht op de constructie van de Vragenlijst Sociale Participatie (VSP), waarmee leerkrachten de sociale participatie van leerlingen met beperkingen in groep 3, 4 en 5 van het regulier basisonderwijs kunnen beoordelen. De VSP bestaat uit vier subschalen, gerelateerd aan vier hoofdthema's van sociale participatie: "vriendschappen", "contacten/interacties", "sociale zelfperceptie van de leerling" en "acceptatie door klasgenoten". Uit een Mokken Schaalanalyse blijkt dat de subschalen gematigd sterk tot sterk zijn. Het Dubbele Monotonie-Model is van toepassing op iedere subschaal. Dat impliceert dat de subschaalscores op een ordinale schaal liggen en de items invariant geordend zijn. Bovendien is de VSP betrouwbaar en zijn er sterke aanwijzingen voor de discriminante validiteit. Tevens komen uit een confirmatieve factoranalyse met behulp van LISREL aanwijzingen voor de convergente validiteit van de VSP naar voren. De fitmaten tonen dat het model van sociale participatie theoretisch standhoudt voor groep 3 en 4 en gedeeltelijk voor groep 5.

1 Inleiding

In de afgelopen decennia is het wereldwijd steeds minder vanzelfsprekend geworden om leerlingen met beperkingen onderwijs te bieden in separate vormen van speciaal onderwijs en neemt het aantal leerlingen met beperkingen in het regulier onderwijs toe (Ferguson, 2008). De groep 'leerlingen met beperkingen' is niet helder omschreven. In verschillende landen worden daar verschillende groepen leerlingen mee aangeduid. Veelal wordt als overkoepelende term leerlingen met 'special (educational) needs' gebruikt.

Ook in Nederland zijn de afgelopen decennia verschillende beleidsinitiatieven genomen om deze groep binnen het regulier onderwijs te houden en de groei van het aantal leerlingen in scholen voor speciaal (basis-)onderwijs af te remmen (Karsten, Roeleveld, Peetsma, & Vergeer, 2001). In Nederland kan de groep leerlingen met 'special needs' het best omschreven worden als de groep leerlingen in scholen voor speciaal (basis)onderwijs en de leerlingen met een leerling gebonden budget (LGF). De laatste groep wordt ook wel aangeduid met 'leerlingen met beperkingen'. Met het van kracht worden van het 'Weer Samen Naar School beleid' (WSNS) en met de Leerlinggebonden financiering (LGF) zijn barrières voor de integratie van deze leerlingen weggenomen.

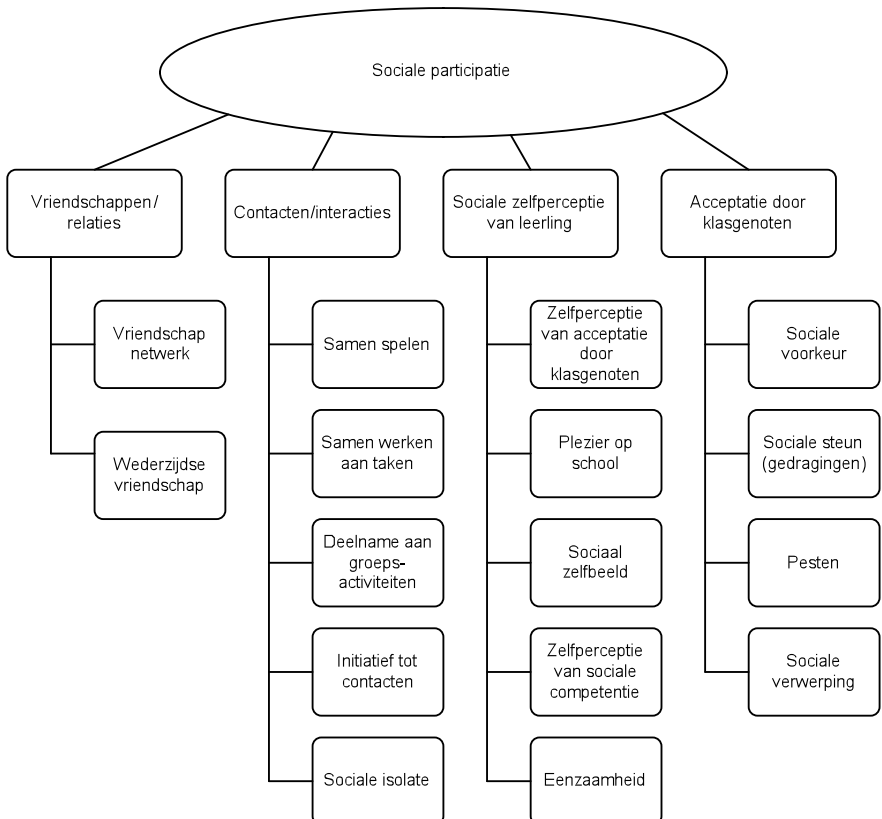
De invoering van de LGF in 2003 heeft het voor ouders van leerlingen met een leerling gebonden budget mogelijk gemaakt te kiezen tussen regulier en speciaal onderwijs. Wanneer ouders voor regulier onderwijs kiezen, is dat naast academische motieven (zoals de verwachting van hogere leerprestaties en betere kansen op vervolgonderwijs), vooral vanwege sociale motieven (Scheepstra, 1998; Sloper & Tyler, 1992). Ouders verwachten en wensen vaak dat plaatsing van hun kind binnen het regulier onderwijs leidt tot vriendschappen met kinderen zonder beperkingen. Internationale studies laten echter zien dat leerlingen met beperkingen, in vergelijking met hun medeleerlingen zonder beperkingen, in mindere mate geaccepteerd worden (Freeman & Alkin, 2000), veelal een lagere sociale status hebben in de klas (Pavri & Luftig, 2000) en vaker sociaal geïsoleerd zijn (Margalit & Efrati, 1996). Uit verscheidene studies is gebleken dat met name leerlingen met stoornissen binnen het autistisch spectrum en leerlingen met ernstige gedragsproblemen, moeite hebben om vriendschappen op te bouwen met kinderen

zonder beperkingen en meer risico lopen in een sociaal isolement te geraken (Chamberlain, Kasari, & Rotheram-Fuller, 2007). Dit is zorgelijk, mede omdat isolement in de klas van negatieve invloed kan zijn op de algehele ontwikkeling van een kind (Bagwell, Newcomb, & Bukowski, 1998).

Vanwege de nadelige gevolgen van negatieve sociale ervaringen op school, is het belangrijk dat scherp in de gaten gehouden wordt hoe het kinderen met beperkingen op sociaal gebied vergaat in de klas. Hier is een belangrijke rol voor leerkrachten weggelegd, omdat zij de aangewezen personen zijn om problemen te signaleren en, indien nodig, maatregelen te treffen. Uit onderzoek (bijv. Whitney, Smith, & Thompson, 1994) blijkt echter dat leerkrachten moeite hebben om de sociale positie van leerlingen met beperkingen in te schatten. Zij blijken geneigd te zijn een te positief beeld te hebben en negatieve voorvallen (zoals pesten) onvoldoende op te merken (De Monchy, Pijl, & Zandberg, 2004).

Om leerkrachten te ondersteunen bij het in kaart brengen van de sociale participatie van leerlingen met beperkingen, is door Koster, Nakken, Pijl en Van Houten (2009) een signaleringsinstrument ontwikkeld. Dit instrument is specifiek gericht op kinderen in groep 3, 4 en 5, omdat kinderen in deze groepen beginnen met het vormen van stabiele en hechte contacten en vriendschappen (Berndt, 2004). Als kinderen eenmaal worden verworpen door medeleerlingen en een geïsoleerde positie in de groep innemen, dan is dat heel moeilijk te veranderen (Hoza, Mrug, Pelham, Greiner, & Gnagy, 2003). Het tijdig signaleren van problemen en het treffen van eventuele maatregelen is daarom van groot belang.

De eerste stap in de ontwikkeling van het signaleringsinstrument voor leerkrachten betrof het beantwoorden van de vraag wat onder de sociale dimensie van inclusief onderwijs wordt verstaan. Voorafgaand aan de ontwikkeling van het signaleringsinstrument werd daarom een literatuurverkenning



Figuur 1. Overzicht van hoofdthema's en deelthema's binnen sociale participatie (Koster, Nakken et al., 2009).

naar de sociale dimensie van inclusie uitgevoerd.

2 Theoretisch kader

Tal van begrippen worden gebruikt om de sociale dimensie van inclusie te beschrijven (Koster, Nakken et al., 2009). Drie overkoepelende termen, te weten sociale participatie, sociale integratie en sociale inclusie, worden frequent door onderzoekers gehanteerd. Er bestaat echter onduidelijkheid over de betekenis van deze termen.

Uit de literatuurverkenning, gebaseerd op 62 artikelen, afkomstig van internationale wetenschappelijke tijdschriften, bleek dat de drie genoemde termen 'sociale participatie', 'sociale integratie' en 'sociale inclusie' vaak globaal beschreven worden en dat slechts enkele onderzoekers een expliciete definitie geven. Daarnaast bleek uit de literatuurverkenning dat de drie termen door veel onderzoekers als synoniemen worden beschouwd en als zodanig worden gehanteerd. Koster, Nakken e.a. (2009) stellen voor sociale participatie te hanteren als centraal begrip. Uit de literatuurverkenning bleek vervolgens dat veel onderzoekers hun omschrijving van sociale participatie beperken tot de vriendschappen van leerlingen (zie o.a. Harper, Maheady, Mallette, & Karnes, 1999), terwijl anderen vooral het belang van interacties tussen leerlingen of het zelfbeeld van leerlingen benadrukken (zie o.a. Kamps, Dugan, Potucek, & Collins, 1999; Pavri & Luftig, 2000), of hun beschrijving van sociale participatie juist beperken tot acceptatie van leerlingen met beperkingen door leerlingen zonder beperkingen (zie o.a. Odom, 2000).

Ondanks dat de auteurs een groot aantal verschillende aspecten noemen, zijn ze het grotendeels eens over vier hoofdthema's binnen sociale participatie. Dit resulteerde in de volgende definitie van sociale participatie (waarbij de hoofdthema's onderstreept zijn):

Sociale participatie van leerlingen in het regulier onderwijs is de aanwezigheid van positieve sociale contacten/interacties tussen hen en hun klasgenoten; acceptatie van hen door hun klasgenoten, vriendschappen tussen hen en hun klasgenoten

en de perceptie van leerlingen dat zij geaccepteerd worden door hun klasgenoten (Koster, Nakken et al., 2009, pp. 135).

De vier hoofdthema's van sociale participatie kunnen beschouwd worden als theoretisch kader voor de ontwikkeling van een instrument voor leerkrachten, waarmee zij de sociale participatie van leerlingen met beperkingen in kaart kunnen brengen. De literatuurverkenning resulteerde in een model van sociale participatie met de vier hoofdthema's, die elk bestaan uit deelthema's (zie Figuur 1).

De ontwikkeling en toetsing van een van dit model afgeleid signaleringsinstrument voor leerkrachten staan centraal in dit artikel.

3 Ontwikkeling Vragenlijst Sociale Participatie

3.1 Werkwijze

In de ontwikkeling van de Vragenlijst Sociale Participatie (VSP) kunnen drie fasen worden onderscheiden. De eerste fase (3.2) bestond uit consultatie van een panel van ervaringsdeskundigen, met als doel geschikte items te selecteren. In de tweede fase (3.3) werd een definitieve selectie van items gemaakt met behulp van een grote groep respondenten, waarna in een kleine pilotstudy een eerste indruk van de psychometrische kwaliteit van de nieuwe vragenlijst verkregen werd (3.4).

3.2 Itemselectie door middel van consultatie van een panel

Het model van sociale participatie (zie Figuur 1) laat zien welke thema's gekoppeld zijn aan het begrip sociale participatie. Na de constructie van het model werden de deelthema's geoperationaliseerd in 38 concrete en meetbare items (bijvoorbeeld: 'na schooltijd speelt de leerling geregeld met klasgenoten'). Deze items vormden het begin van een vragenlijst voor leerkrachten.

De inhoud van deze items was grotendeels gebaseerd op literatuur en deels op de klinische en onderzoekservaring van de onderzoekers. Om de invloed van persoonlijke opvattingen van de onderzoekers te verkleinen, werd een panel bestaande uit veertien personen met veel praktische ervaring rondom inclusief onderwijs (onder wie

leerkrachten, ambulant begeleiders, ouders van een kind met een beperking en universitaire studenten met een beperking) tijdens een focusgroepsbijeenkomst gevraagd om kritisch te kijken naar de items. Hen werd gevraagd of de geformuleerde items in lijn waren met wat zij onder sociale participatie verstonen en of er aspecten ontbraken. Mede op basis van de opmerkingen van het panel werd een lijst van in totaal 74 items betreffende sociale participatie opgesteld (Koster et al., 2009).

3.3 Definitieve itemselectie door middel van beoordeling door grote groep respondenten

Met het oog op een betere hanteerbaarheid in de praktijk werd besloten de lijst in te korten. Om de relevantie van de items te bepalen, werd een grote groep respondenten ($n = 190$) gevraagd elk van de items te beoordelen. De groep respondenten bestond uit leerkrachten, ambulant begeleiders, ouders van een kind met een beperking, universitaire studenten met een beperking en onderwijsadviseurs. Allen hadden ervaring met inclusief onderwijs. Hen werd gevraagd op een 5-puntsschaal het belang van elk item aan te geven. De 30 items die door tenminste 75% van de respondenten als belangrijk of zeer belangrijk werden beoordeeld, werden geselecteerd voor de vragenlijst voor leerkrachten, de Vragenlijst Sociale Participatie (VSP). De VSP is gericht op leerlingen in groep 3, 4 en 5. Leerkrachten vullen de VSP in voor een leerling, door bij elk item op een 5-puntsschaal aan te geven in hoeverre het item van toepassing is op de leerling [uiteenlopend van 'dit is geheel niet van toepassing op de leerling' (1), tot 'dit is sterk van toepassing op de leerling' (5)]. Het invullen van de VSP neemt vijf tot tien minuten in beslag. Hoe hoger de score op de VSP, des te beter de sociale participatie van de leerling.

3.4 Pilotstudy naar psychometrische kwaliteiten VSP

In een pilotstudy vond een eerste onderzoek plaats naar de psychometrische kwaliteiten van de 30 items, op basis van de scores van 60 leerlingen (25 met beperkingen, 35 zonder), zoals ingevuld door de leerkracht. De betrouwbaarheid van de subschalen Vriend-

schappen/relaties en Contacten/interacties was hoog (Alpha was respectievelijk .94 en .88), terwijl de betrouwbaarheid van de subschalen Sociale zelfperceptie en Acceptatie door klasgenoten vrij laag was (Alpha was respectievelijk .69 en .64). Om de betrouwbaarheid van deze twee laatste subschalen te verbeteren werden drie items toegevoegd aan de subschaal Sociale zelfperceptie en twee aan de subschaal Acceptatie door klasgenoten. Alle vijf toegevoegde items zijn afkomstig van de oorspronkelijke lijst met 74 items. Eén item, behorend tot de Contacten/interacties subschaal, werd verwijderd omdat dit item negatief correleerde met de totaalscore op de VSP. De herziene VSP bestond derhalve uit 34 items. Elk van de vier hoofdthema's werd vertegenwoordigd door zes tot veertien items. Voorbeelden van items zijn: 'De leerling maakt deel uit van een vriendengroepje' (subschaal Vriendschappen/relaties), 'De leerling heeft zichtbaar pret met klasgenoten' (subschaal Contacten/interacties), 'De leerling heeft het gevoel bij de groep te horen' (subschaal Sociale zelfperceptie van de leerling) en 'Klasgenoten zijn bereid om de leerling te helpen, zonder daarbij betuttelend te zijn' (subschaal Acceptatie door klasgenoten).

4 Methode

4.1 Inleiding

Na afronding van de pilotstudy, werd een grootschaliger onderzoek gestart, waarbij de psychometrische kwaliteiten van de VSP nader werden onderzocht.

Door middel van een non parametrische itemrespons analyse (Mokken Schaalanalyse) is de kwaliteit van de VSP als geheel en van de vier subschalen afzonderlijk onderzocht (zie 4.4). Tevens is de discriminante validiteit van de VSP onderzocht (zie 4.5). Vervolgens is met behulp van een LISREL analyse de convergente validiteit van de VSP onderzocht (zie 4.6).

4.2 Design

Er is naast de VSP geen ander instrument dat het gehele concept sociale participatie omvat. Om de convergente validiteit van de vragen-

lijst te bepalen, werd voor elk van de vier hoofdthema's van sociale participatie een afzonderlijke criteriummaat geselecteerd. Als gevolg daarvan zouden naast de VSP, die door de leerkrachten moest worden ingevuld, per leerling vier instrumenten moeten worden gehanteerd. Iedere klas zou dan gedurende minimaal twee uur bezocht moeten worden, wat een grote mate van non-respons tot gevolg zou kunnen hebben. Aangezien dit de generaliseerbaarheid van de bevindingen negatief zou beïnvloeden, werd besloten het onderzoek in twee deelstudies op te splitsen. In elke deelstudie werden twee hoofdthema's gemeten en de door te brengen tijd in klassen werd daarmee gereduceerd.

4.3 Respondenten

In het onderzoek waren de groepen 3, 4 en 5 van reguliere basisscholen betrokken. In iedere deelnemende klas zat minimaal één leerling met Leerling Gebonden Financiering (LGF). Er is gekozen voor leerlingen met LGF omdat deze leerlingen door een Commissie voor Indicatiestelling zijn beoordeeld. Dat geeft enige garantie voor de aard en de ernst van de beperkingen en problemen.

De leerkracht vulde voor zowel de leerling(en) met LGF, als voor twee klasgenoten zonder LGF een VSP in, om zodoende de scores tussen leerlingen met en zonder LGF te kunnen vergelijken. De laatstgenoemde leerlingen werden ad random gekozen door de derde en zestiende leerling van de leerling-lijst te selecteren. Als de leerling met LGF de derde of zestiende leerling was op de lijst, of als leerling 3 en 16 van hetzelfde geslacht waren, werd de eerstvolgende leerling op de lijst geselecteerd. In klassen met

twee of meer leerlingen met LGF vulde de leerkracht de VSP in voor de leerling(en) met LGF en voor twee leerlingen zonder LGF.

Bij de start van beide deelstudies werd aan 300 reguliere basisscholen een uitnodiging voor deelname aan het onderzoek gestuurd. Deze scholen werden random geselecteerd uit het totaal van 2074 scholen die volgens de Centrale Financiële Instellingen (CFI), één of meer leerlingen met LGF geplaatst hadden in groep 3, 4 en/of 5. Het ging daarbij specifiek om leerlingen met ernstige spraak-taalproblematiek (ESM), gedragsstoornissen, stoornissen binnen het autistisch spectrum (ASS), verstandelijke beperkingen en motorische beperkingen. Scholen op de Waddeneilanden en in de provincies Zeeland, Limburg, Noord-Brabant en Zuid-Holland werden buiten het onderzoek gelaten, omdat de reistijd vanaf Groningen te groot was.

In de eerste deelstudie waren 75 klassen, afkomstig van 53 scholen, betrokken. In twee klassen vulde de leerkracht voor slechts één leerling (in plaats van twee leerlingen) zonder LGF een VSP in. Daarom bestaat de uiteindelijke steekproef uit 244 leerlingen, van wie 96 met LGF en 148 zonder LGF. Van deze leerlingen zaten er 48 in groep 3, 120 in groep 4 en 76 in groep 5. In het kader van het onderzoek naar de convergente validiteit, richtte de dataverzameling bij deze eerste deelstudie zich op de thema's 'sociale zelf-perceptie van de leerling' en 'acceptatie door klasgenoten'. De tweede deelstudie, waarin 105 klassen van 66 scholen betrokken waren, richtte zich op vriendschappen van leerlingen en op hun contacten en interacties met klasgenoten (zie Tabel 2 voor het aantal leerlingen per thema). In vijf klassen uit de tweede

Tabel 1

Spreiding van leerlingen met LGF over typen beperkingen en sekse in beide deelstudies

	Eerste deelstudie		Tweede deelstudie	
	# leerlingen (% totaal)	# jongens (% type beperkingen)	# leerlingen (% totaal)	# jongens (% type beperkingen)
Gedragsstoornis	13 (13.5%)	11 (84.6%)	16 (11.3%)	14 (87.5%)
Autisme Spectrum stoornis	42 (43.8%)	38 (90.5%)	55 (39.0%)	45 (81.8%)
Motorische beperking	10 (10.4%)	9 (90.0%)	25 (17.7%)	18 (72.0%)
Verstandelijke beperking	11 (11.5%)	5 (45.5%)	15 (10.6%)	5 (33.3%)
Spraak-taalproblemen	20 (20.8%)	13 (65.0%)	27 (19.1%)	19 (70.4%)
Totaal	96 (100%)	76 (79.2%)	138 (97.8%)*	101 (73.2%)

*Twee leerlingen met leerstoornissen en één chronisch zieke leerling waren betrokken in de tweede deelstudie, maar werden niet meegenomen in de analyses van dit artikel vanwege de kleine aantallen.

deelstudie vulde de leerkracht de VSP in voor twee leerlingen met LGF, maar voor slechts één leerling zonder LGF. De uiteindelijke steekproef bestaat uit 141 leerlingen met LGF en 205 leerlingen zonder LGF, van wie 114 in groep 3 zaten, 121 in groep 4 en 111 in groep 5 (zie Tabel 1).

4.4 Analyse psychometrische kwaliteit van de VSP

Om de kwaliteit van de items te beoordelen en om te bepalen in welke mate de items van elk van de vier subschalen daadwerkelijk tot die subschaal behoren, werd een item respons analyse (zie Embretson & Reise, 2000) verricht. Item respons modellen geven de relaties weer tussen scores op de items en de door de vragenlijst beoogde te meten latente vaardigheid. Deze relatie wordt, voor ieder item afzonderlijk, weergegeven door middel van de Item Response Functie (IRF). Hier worden alleen non-parametrische modellen besproken omdat parametrische modellen geen fit toonden met de data. De non-parametrische item respons modellen die voor het onderzoek werden bestudeerd, zijn gebaseerd op drie assumpties (Sijtsma & Molenaar, 2002):

- Unidimensionaliteit (alle items meten dezelfde latente trek)
- Lokale onafhankelijkheid (de score op een item wordt niet beïnvloed door scores op andere items)
- Monotoniciteit (alle item respons functies zijn monotoon niet-dalend. Dat wil zeggen dat de kans op een positief antwoord op een item toeneemt of gelijk blijft bij een hogere latente trekwaarde)

Deze drie assumpties vormen samen het Monotoon Homogeniteits-Model (MHM; Mokken, 1971). Als aan deze assumpties wordt voldaan voor een verzameling items, dan zijn de respondenten naar hun itemscores te rangordenen op de latente eigenschap (Sijtsma & Molenaar, 2002). Wanneer bovendien de item respons functies elkaar niet kruisen, houdt het strengere Dubbele Monotonie-Model (DMM; Mokken, 1971) stand. Er is dan voldaan aan de eis van 'invariante itemordering', wat inhoudt dat de ordening van de items naar itemmoeilijkheid voor alle respondenten gelijk is (Sijtsma & Molenaar, 2002).

Voor elk van de vier subschalen werd

de fit van de beide modellen beoordeeld door middel van de Mokken Schaalanalyse voor polytome items (Molenaar & Sijtsma, 2000). Loevinger Schaalbaarheidscoëfficiënt H_i geeft voor elk item aan of item i fit met het MHM ($0 \leq H_i \leq 1$). Het geeft daarnaast aan in welke mate het item bijdraagt aan het onderscheiden van individuen. Om ervoor te zorgen dat items voldoen aan de assumpties van het MHM en dat items voldoende onderscheidend vermogen hebben, wordt over het algemeen $H_i > .3$ beschouwd als minimale schaalwaarde (Sijtsma & Molenaar, 2002).

Coëfficiënt H is een gewogen gemiddelde van de H_i waarden van de items die samen een subschaal vormen en geeft de mate weer waarin leerlingen geordend kunnen worden door middel van de somscores op de items. Bij waarden van $.3 \leq H_i < .4$ wordt een subschaal als zwak beschouwd, bij waarden van $.4 \leq H_i \leq .5$ als middelmatig en bij waarden van $H_i > .5$ als sterk.

Naast de item schaalbaarheidscoëfficiënten worden nog andere diagnostische maten gebruikt om te beoordelen of de monotoniciteit assumptie en de non-intersectie assumptie van het DMM standhouden voor elk item. Deze diagnostische maten worden samengevat in criteriawaarden, waarbij een waarde ≥ 80 duidt op een grote schending van assumpties, een waarde tussen 40 en 80 twijfelachtig is betreffende assumptieschending en een waarde ≤ 40 aangeeft dat assumpties niet geschonden worden (Molenaar & Sijtsma, 2000). Om mogelijke 'differential item functioning' tussen de groep leerlingen met en de groep leerlingen zonder LGF te beoordelen, is de ordening van de scores op alle items vergeleken tussen beide groepen. Een verschillende ordening suggereert differential item ordering, wat zou impliceren dat de subschalen niet gebruikt zouden kunnen worden om leerlingen van beide groepen te vergelijken. De vier subschalen van de VSP zijn afzonderlijk geanalyseerd met als doel afwijkende items te ontdekken en om mogelijke differential item ordering te beoordelen tussen leerlingen met en zonder LGF.

Om te onderzoeken of de vier subschalen van de VSP in de data konden worden onderscheiden, werd de 'Automated Item Selection Procedure' (AISP; Sijtsma & Molenaar,

2002) gebruikt als dimensionaliteitsanalyse. Daarbij worden stapsgewijs steeds de items geselecteerd die het best bij de schaal passen. Wanneer de AISP verschillende subschalen onderscheidt, zal bekeken worden of de items worden toegeschreven aan de subschalen waar zij op voorhand toe behoorden.

4.5 Analyse discriminante validiteit

De discriminante validiteit van de VSP wordt onderzocht door de scores van leerlingen met LGF te vergelijken met scores van leerlingen zonder LGF. Om de verschillen te beoordelen wordt gebruik gemaakt van een t-toets. De VSP is op te vatten als een ordinale schaal, terwijl de t-toets data op intervalniveau vraagt. Het is in de psychometrie niet ongebruikelijk om ordinale data te behandelen als interval data. Nunnally (1967, pag 20 ev) geeft een aantal goede argumenten daarvoor (zie ook: Gregory, 2011). Op basis van literatuur (zie o.a. Freeman & Alkin, 2000) werd verwacht dat de leerlingen met LGF lagere scores zouden behalen dan leerlingen zonder LGF. Indien deze verwachting uitkomt, kan dit beschouwd worden als aanwijzing voor de discriminante validiteit van de VSP. In totaal werden de scores van 234 leerlingen met LGF en 346 leerlingen zonder LGF in de analyse betrokken.

4.6 Analyse convergente validiteit

Inleiding

Voor het bepalen van de convergente validiteit van de VSP werd voor elk van de vier hoofdthema's van sociale participatie een afzonderlijk criterium gebruikt. De convergente validiteit van de VSP werd beoordeeld door de scores van leerlingen op de vier subschalen te vergelijken met de scores op vier instrumenten die respectievelijk sociale zelfperceptie, acceptatie, vriendschappen en contacten/interacties meten. Ook hier (zie 4.5) zijn de VSP data opgevat als data op intervalniveau. De correlatiecoëfficiënten werden gebruikt als input voor confirmatieve factor analyses, waarbij gebruik werd gemaakt van LISREL software (LISREL 8.80, Jöreskog & Sörbom, 2007). Deze analyses vormen een omvattende benadering om de convergente validiteit van de VSP te beoordelen. Ze verschaffen inzicht in de mate waarin de data aansluiten bij het

model van sociale participatie, bestaande uit de vier hoofdthema's (zie Figuur 1).

Meetinstrumenten

Meting van sociale zelfperceptie. Het bleek noodzakelijk om twee instrumenten voor sociale zelfperceptie te selecteren, omdat er geen geschikte instrumenten beschikbaar waren die zich richtten op leerlingen in zowel groep 3, 4 als 5. Voor de groepen 3 en 4 werd gekozen voor de sociale subschaal van de Platenschaal voor Zelf-Waargenomen Competentie en Sociale Acceptatie (Van Rossum & Vermeer, 1992). De betrouwbaarheid (Alpha) van deze schaal is .78 (Van Rossum & Vermeer, 1992), wat als voldoende beschouwd kan worden voor onderzoeksdoeleinden (Nunnally, 1967). De score op de Platenschaal (bestaande uit vijf vragen), kan variëren tussen 5 en 20. Voor groep 5 werd gekozen voor de sociale subschaal van de Competentie Belevingsschaal voor Kinderen (CBSK, Veerman, Straathof, Treffers, Van den Bergh, & Ten Brink, 1997). De betrouwbaarheid van deze schaal is voldoende ($\alpha = .74$, Veerman et al., 1997). De minimumscore op de CBSK (bestaande uit zes vragen) is 6, de maximumscore is 24. Voor zowel de CBSK als de Platenschaal geldt dat des te hoger de score, des te hoger de sociale zelfperceptie.

Meting van acceptatie door klasgenoten. Acceptatie door klas- of leeftijdgenoten wordt veelal gemeten met behulp van sociometrische technieken (Berndt & Burgy, 1996), zoals de nominatieprocedure (zie Meting van vriendschappen) en de ratingprocedure. Bij ratingprocedures wordt aan alle leerlingen uit een klas gevraagd voor elk van hun klasgenoten aan te geven in hoeverre ze met hen willen spelen, werken, enzovoort. Er lijkt consensus over te bestaan dat een ratingschaal het best gebruikt kan worden bij het meten van acceptatie, terwijl de nominatiemethode het best gebruikt kan worden bij het meten van vriendschap (Larrivee & Horne, 1991). In overeenstemming hiermee werd in ons onderzoek de ratingprocedure gekozen als methode om acceptatie van leerlingen in kaart te brengen. De leerlingen werd gevraagd een ratingschaal in te vullen met de namen van al hun klasgenoten, waarbij ze op een 3-puntsschaal aangaven in hoeverre ze

zouden willen spelen met iedere klasgenoot. Ze konden kiezen uit de volgende antwoordcategorieën, die visueel werden ondersteund door smileys: 1) ja, graag ☺, 2) het maakt me niet uit ☹, en 3) nee, liever niet ☹.

Meting van vriendschappen. Diverse onderzoekers (onder wie Larrivee & Horne, 1991; Parker & Asher, 1993) beschouwen de wederkerige vriendschap nominatieprocedure als adequate methode om vriendschappen in kaart te brengen. Bij deze procedure wordt kinderen gevraagd om andere kinderen te noemen die zij beschouwen als hun (beste) vriend(in), of met wie zij graag spelen, werken, enzovoort (Larrivee & Horne, 1991). Deze procedure werd gebruikt voor het meten van vriendschap in ons onderzoek. De leerlingen werd gevraagd op te schrijven wie hun maximaal vijf beste vrienden/vriendinnen zijn in de klas (Frostd & Pijl; 2007).

Meting van contacten/interacties. Voor het in kaart brengen van contacten en interacties tussen leerlingen werd gekozen voor Gresham's Observatie Categorieën (Gresham, 1982), vanwege de eenduidigheid van de gebruikte categorieën van sociale interactie en vanwege de hoge inter-observator betrouwbaarheid (Montague & Rinaldi; 2001). Bovendien bieden de categorieën een goed overzicht van de hoeveelheid en het initiatief tot contacten/interacties (geïnitieerd door de leerling die wordt geobserveerd of geïnitieerd door een klasgenoot).

De observatoren kregen een training van drie uur, waarbij gebruik werd gemaakt van videobeelden van een klassensituatie. Na de training werd de overeenstemming tussen observatoren bepaald door middel van het berekenen van Cohen's Kappa voor 'geïnitieerde/ontvangen interactie van leerling met klasgenoten'. Cohen's Kappa was .76, wat duidt op een redelijke mate van overeenstemming (Landis & Koch, 1977).

Procedure

Zoals beschreven in paragraaf 4.3 (Respondenten) vulden de leerkrachten van de klassen die deelnamen aan het onderzoek de VSP in voor de leerling(en) met LGF en voor twee leerlingen zonder LGF.

Sociale zelfperceptie. In de 75 groepen van de eerste deelstudie werd de sociale zelf-

perceptie van leerlingen in kaart gebracht. In groep 5 vond de afname van de sociale subschaal van de CBSK groepsgewijs plaats, terwijl de sociale subschaal van de Platenschaal in groep 3 en 4 individueel (buiten de klas) werd afgenomen. Voor iedere leerling werd een sociale zelfperceptiescore berekend door de ruwe scores op te tellen op de CBSK (groep 5) of de Platenschaal (groep 3 en 4). Voor leerlingen in groep 5 kon de score variëren van 6 tot 24, voor leerlingen in groep 3 en 4 kon de score uiteenlopen van 5 tot 20.

Acceptatie door klasgenoten. In de groepen 4 en 5 van de eerste deelstudie werd de ratingschaal klassikaal afgenomen, in groep 3 vond het invullen van de schaal individueel (buiten de klas) plaats: de onderzoeker las de namen van alle klasgenoten op en de leerling gaf mondeling aan hoe graag hij/zij met ieder van hen wilde spelen. Elke leerling ontving scores van alle klasgenoten op een 3-puntschaal (zie Meetinstrumenten). Optelling van alle scores resulteerde in een ruwe score voor iedere leerling. Aangezien de totaalscore die een leerling kon behalen sterk afhankelijk was van de groepsomvang, werden Z-scores berekend. Als gevolg daarvan konden scores van leerlingen zowel worden vergeleken met de scores van hun klasgenoten als met die van leerlingen uit andere klassen (met verschillende aantallen leerlingen).

Vriendschappen. De leerlingen uit de tweede deelstudie werd gevraagd op te schrijven wie hun beste vrienden/vriendinnen waren in de klas, met een maximum van vijf. De data afkomstig van de wederzijdse vriendschap nominatiemethode zijn geanalyseerd met behulp van UCINET software (Borgatti, Everett, & Freeman, 1999) voor het vaststellen van vriendschappen. Vriendschap werd gedefinieerd als wederkerige keuze, wat impliceert dat twee leerlingen elkaar als vriend nomineren (Frostd & Pijl, 2007). Voor iedere leerling werd geteld hoeveel vriendschappen hij/zij had.

Contacten/interacties. Het uitvoeren van observaties kost veel tijd en daarom werden slechts in 58 klassen van de tweede deelstudie observaties uitgevoerd. In elke klas werden zowel de leerling met LGF als een leerling zonder LGF van hetzelfde geslacht gedurende vier periodes van vijf minuten

geobserveerd. De leerling zonder LGF werd ad random gekozen door de derde en zestiende leerling van de leerling-lijst te selecteren. Als de leerling met LGF de derde of zestiende leerling was, of als leerling 3 en 16 beiden van de andere sekse waren, werd de eerstvolgende leerling van dezelfde sekse op de lijst geselecteerd. Drie keer vijf minuten vond de observatie tijdens lessen plaats, de resterende vijf minuten werden de kinderen geobserveerd in vrije situaties. Elke periode van vijf minuten was verdeeld in 30 intervallen van tien seconden. Indien meer dan één interactie plaatsvond in een periode van tien seconden, dan werd alleen de eerste interactie genoteerd. Het totaal aantal (geïnitieerde en ontvangen) interacties met klasgenoten werd voor iedere leerling opgeteld.

Tabel 2 biedt een overzicht van de instrumenten die gehanteerd zijn om de vier hoofdthema's te meten, met daarbij vermeld het aantal deelnemende leerlingen (met en zonder LGF).

De nominatiemethode en de subschaal Vriendschappen/relaties worden verondersteld indicatoren te zijn van de latente factor "Vriendschap" (ξ_1). Het observatieschema en de subschaal 'Contacten/interacties' zijn indicatoren van de latente factor "Contact" (ξ_2). De sociale subschaal van de CBSK/Platenschaal en de subschaal Sociale Zelfperceptie van de leerling worden beschouwd als indicatoren van de latente factor "Sociale zelfperceptie" (ξ_3). Tenslotte worden de ratingschaal en de subschaal Acceptatie door klasgenoten geacht indicatoren te zijn van de latente factor "Acceptatie" (ξ_4). Twee separate analyses werden uitgevoerd voor elke deelstudie (zie 4.2). De studies werden gesplitst in groep 3/4 en groep 5, omdat het instrument voor het meten van de sociale zelfperceptie van leerlingen verschilde voor

leerlingen in de groepen 3/4 en groep 5 (zie Tabel 2).

Vier indices werden gehanteerd om model fit te evalueren: de verhouding van chi-kwadraat tot de vrijheidsgraden is een eerste indicatie van model fit. Deze verhouding (χ^2/df) is, in vergelijking met chi-kwadraat of de P-waarden, minder gevoelig voor groepsomvang en voor afwijkingen van normaliteit (Rozendaal, Minnaert, & Boekaerts, 2003). In navolging van Byrne (1989), wordt een χ^2/df verhouding die gelijk is aan of lager dan 2 als een voldoende fit beschouwd. Daarnaast werden de volgende twee indices gebruikt om model fit te evalueren: root mean square error of approximation (RMSEA) en goodness of fit index (GFI). Volgens Jaccard en Wan (1996, in Vellutino, Tunmer, Jaccard, & Chen, 2007), duiden statistisch niet-significante waarden van RMSEA's die niet groter zijn dan .08 op acceptabele model fit, evenals GFIs die gelijk zijn aan of groter dan .90.

5 Resultaten

5.1 Resultaten psychometrische kwaliteit van de VSP

Op basis van de uitkomsten van de analyses zijn 10 van de 34 items vanwege hun lage H_i waarden en/of hoge criteriawaarden uit de vragenlijst verwijderd (één item uit de subschaal Vriendschappen/relaties, vijf uit de subschaal Contacten/interacties, twee uit de subschaal Sociale zelfperceptie van de leerling en twee uit de subschaal Acceptatie). De items zijn in hoofdzaak verwijderd vanwege hun lage H_i waarden en/of hoge criteriawaarden, maar er werd ook altijd gekeken naar de inhoud. De belangrijkste reden om items te verwijderen was hun zwakke onderscheidende vermogen; items waarop vrijwel alle leerlingen een hoge

Tabel 2

Instrumenten voor het meten van de hoofdthema's gekoppeld aan het aantal leerlingen

Hoofdthema	Instrument	# leerlingen met LGF	# leerlingen zonder LGF
Vriendschappen	Wederkerige nominatie methode	137	202
Contacten/interacties	Observatieschema	58	58
Sociale zelfperceptie	Sociale subschaal van CBSK	27	42
	Sociale subschaal van Platenschaal	67	99
Acceptatie door klasgenoten	Sociometrische rating scale	96	148

score behaalden, werden verwijderd, omdat ze weinig toegevoegde waarde hadden. Naast de afwijkende waarden, waren er ook inhoudelijke redenen om items te verwijderen. Zo werden bij de subschaal 'Sociale zelfperceptie' twee items verwijderd, omdat ze gericht waren op de durf van de leerling om iets te vragen of te vertellen in de klas. Deze twee items waren afwijkend van de andere items van deze subschaal, die meer gericht waren op gevoelens van de leerling (zoals blijdschap, eenzaamheid) in plaats van op iets durven. Daarnaast bleek een aantal items sterke inhoudelijke overlap te vertonen met andere items. Ook dat was (in combinatie met lage H_i waarden en/of hoge criteriawaarden) een reden om items te verwijderen.

De definitieve versie van de VSP bestaat uit 24 items, waarvan vijf behoren tot de subschaal Vriendschappen/relaties, negen tot de subschaal Contacten/interacties, vijf

tot de subschaal Sociale zelfperceptie van de leerling en vijf tot de subschaal Acceptatie door klasgenoten. In Tabel 3 staan verkorte versies van de items vermeld met de bijbehorende H_i waarden.

De uitkomsten van de analyses laten zien dat de vier subschalen gematigd sterk tot sterk zijn. Het Dubbele Monotonie-Model bleek van toepassing op iedere subschaal. De subschaalscores liggen op een ordinale schaal en de afzonderlijke items zijn invariant geordend.

Uit de analyses bleek dat bij geen van de vier subschalen differential item ordening aan de orde was. De subschaalscores zijn derhalve vergelijkbaar voor leerlingen met en zonder LGF. Bovendien wordt het model van sociale participatie, waarin de vier hoofdthema's worden onderscheiden, ondersteund door de uitkomsten van de analyses (AISP). De betrouwbaarheid van de vragenlijst als geheel blijkt hoog te zijn en ook de

Tabel 3

Verkorte versies van items Vragenlijst Sociale Participatie per subschaal en de bijbehorende H_i waarden

Verkorte versies van items	H_i
Subschaal Vriendschappen/relaties	
Deel uitmaken van een vriendengroepje	.53
Het hebben van vrienden in de klas	.54
Na schooltijd spelen met klasgenoten	.49
Het ontvangen van uitnodigingen voor spelen tijdens vakanties	.43
Het ontvangen van uitnodigingen voor verjaardagsfeestjes	.42
Subschaal Contacten/Interacties	
Pret hebben met klasgenoten	.74
Door klasgenoten buitengesloten worden bij activiteiten	.70
Door klasgenoten uitgedaagd worden	.73
Door klasgenoten gevraagd worden om samen te spelen	.73
Door klasgenoten gepest worden	.72
Meedoen met spelletjes	.69
Door klasgenoten uitgelachen worden	.64
Klasgenoten gaan vrijwillig naast de leerling zitten	.69
Samenwerken aan taken	.62
Subschaal Sociale zelfperceptie van de leerling	
Het gevoel hebben tot de groep te behoren	.56
Gevoelens van eenzaamheid	.56
Het gevoel hebben zichzelf te kunnen zijn	.47
Het gevoel hebben gepest te worden	.49
Plezier hebben in het naar school gaan	.47
Subschaal Acceptatie door klasgenoten	
Bereidheid van klasgenoten om, indien nodig, spelregels aan te passen	.53
Bereidheid van klasgenoten om bij samenwerken rekening te houden met wat de leerling wel en niet kan	.55
Bereidheid van klasgenoten om bij spel rekening te houden met houden met wat de leerling wel en niet kan	.61
Opkomen voor de leerling door klasgenoten	.52
Bereidheid van klasgenoten om de leerling te helpen	.48

Tabel 4

Schaalbaarheid (H) en betrouwbaarheid (ρ) van de definitieve subschalen en gemiddelde scores (en SD) op de subschalen van leerlingen met LGF ($n = 234$) en zonder LGF ($n = 346$)

Subschalen	Aantal items	H	ρ	M (SD)	
				Leerlingen met LGF	Leerlingen zonder LGF
Vriendschappen	5	.49	.80	13.91 (6.52)	19.10 (5.49)
Contacten/interacties	9	.70	.95	18.03 (4.34)	21.78 (3.48)
Sociale zelfperceptie	5	.51	.82	18.69 (4.83)	21.14 (4.05)
Acceptatie door klasgenoten	5	.54	.83	16.89 (4.78)	18.07 (4.77)

vier subschalen zijn betrouwbaar bevonden (Rho loopt uiteen van .80 tot .95, zie Tabel 4). Zie voor een uitgebreidere beschrijving van de resultaten Koster, Timmerman, Naken, Pijl, en Van Houten, 2009.

De ruwe score op de subschaal Contacten/interacties kan uiteenlopen van 9 (9 x 1) tot 45 (9 x 5), de ruwe scores op de drie andere subschalen kunnen uiteenlopen van 5 (5 x 1) tot 25 (5 x 5). Om interpretatie van de scores te vergemakkelijken, werden deze lineair herschaald, zodat de minimumscore op alle subschalen 0 is en de maximumscore 25. De totale score op de VSP is de som van de subschaal scores en varieert derhalve van 0 tot 100. Hoe hoger de score op de VSP, des te hoger het niveau van sociale participatie.

5.2 Resultaten discriminante validiteit

De discriminante validiteit van de VSP werd bepaald door de scores van leerlingen met en zonder LGF te vergelijken. Daaruit bleek

dat leerlingen met LGF een aanzienlijk lagere score op de VSP behaalden ($M = 67.5$, $SD = 16.6$) dan hun medeleerlingen zonder LGF ($M = 80.2$, $SD = 15.1$). Het betreft een significant verschil $t(577) = 9.51$, $p < .05$; Cohen's $d = .81$, 95% betrouwbaarheidsinterval = 10.05 - 15.29). Dit impliceert dat de sociale participatie van leerlingen met LGF lager wordt beoordeeld dan de sociale participatie van hun klasgenoten zonder LGF. Aangezien deze uitkomst de aanvankelijke verwachting bevestigt, kan dit beschouwd worden als een aanwijzing voor de discriminante validiteit van de VSP.

5.3 Resultaten convergente validiteit

De convergente validiteit van de VSP werd vastgesteld door de scores van leerlingen op de vier subschalen te vergelijken met de scores op vier instrumenten die respectievelijk sociale zelfperceptie, acceptatie, vriendschappen en contacten/interacties meten. In Tabel 5 staan de correlaties vermeld voor elk

Tabel 5

Correlatie tussen scores op subschalen VSP en scores op instrumenten waarmee de vier hoofdthema's gemeten worden

Criterionvariabelen:	Aantal vriendschappen	Interacties met klasgenoten	Zelfperceptie	Acceptatie
Subschaal VSP:		Geïnitieerde interacties	Ontvangen interacties	
Vriendschappen	.49**	.16	.22*	Groep 3/4: .33** Groep 5: .29*
Contacten/interacties	.45**	.20*	.27**	Groep 3/4: .27** Groep 5: .29*
Sociale zelfperceptie	.30**	.11	.19*	Groep 3/4: .22** Groep 5: .25*
Acceptatie door klasgenoten	.25**	.04	.09	Groep 3/4: .17* Groep 5: .12

* $p \leq .05$

** $p \leq .01$

van de subschalen. De verwachting is dat de hogere correlaties gevonden worden op de diagonaal (zie Tabel 5).

Tabel 5 laat zien dat bijvoorbeeld vriendschap gemeten met de VSP redelijk samenhangt ($r = .49$) met het aantal vriendschappen gemeten met de nominatieprocedure. De correlatie van vriendschap gemeten met de VSP met interacties ($r = .16$ en $r = .22$) en met zelfperceptie ($r = .33$ en $r = .29$) is naar verwachting lager. De correlatie tussen vriendschap gemeten met de VSP met acceptatie ($r = .50$) is wat hoger dan verwacht. De correlatiematrix in Tabel 5 laat positieve correlaties zien, maar het verwachte patroon doet zich niet altijd voor. Dit kan betekenen dat de vier achterliggende factoren (vriendschap, interacties, zelfperceptie en acceptatie) niet helemaal te scheiden zijn. Een LISREL analyse kan daar duidelijkheid over verschaffen. De correlaties in Tabel 5 dienden als input voor de LISREL analyses.

De confirmatieve factoranalyse met behulp van LISREL, verschaftte aanwijzingen voor de convergente validiteit van de VSP. Zoals kan worden afgelezen uit Tabel 6, bleken voor groep 3/4 de fit indices voor alle factoren voldoende te zijn. Voor groep 5 bleek er een minder goede fit te zijn tussen model en data: voor twee factoren (ξ_1 'Vriendschap' en ξ_2 'Contact') waren de fit indices acceptabel, voor de twee andere factoren ξ_3 'Sociale zelfperceptie' en ξ_4 'Acceptatie' waren ze dat niet ($RMSEA \leq 0.08$; $\chi^2/df \leq 2$).

De eerste rijen in Tabel 6 laten zien dat vriendschap gemeten met de VSP en vriendschap gemeten met de nominatiemethode beide laden op de latente factor vriendschap (ξ_1) en dat contacten/interacties met de VSP en de criteriumvariabele interacties met klasgenoten laden op de factor contacten (ξ_2). De fit indices voldoen aan de eerder gestelde criteria (zie bijvoorbeeld: $\chi^2/df \leq 2$). Het onderste deel van Tabel 6 geeft de fit van beide instrumenten voor zelfperceptie met factor 3 en voor de instrumenten voor acceptatie met factor 4. Voor de data verzameld in groep 3/4 past het model goed, maar dat geldt niet voor de gegevens uit groep 5.

Voor een uitgebreidere beschrijving van de uitkomsten betreffende de confirmatieve factor analyse, wordt verwezen naar Koster et al. (2011).

6 Conclusie en discussie

Het bevorderen van inclusief onderwijs staat in veel landen hoog op de politieke agenda. Daarbij wordt met name het sociale aspect van inclusief onderwijs als essentieel beschouwd. Uit tal van onderzoeken blijkt echter dat leerlingen met beperkingen die regulier onderwijs volgen het niet altijd gemakkelijk hebben op sociaal gebied. Zij hebben veelal minder vrienden, worden vaker genegeerd en zijn vaker eenzaam (o.a. Freeman & Alkin, 2000; Margalit & Efrati,

Tabel 6

Fitmaten voor concurrerende modellen voor ξ_1 "Vriendschap", ξ_2 "Contact", ξ_3 "Sociale zelfperceptie" en ξ_4 "Acceptatie"

Concurrerende modellen	χ^2	df	χ^2/df	GFI	RMSEA	P of close fit
$\xi_1 \xi_2$						
Groep 3/4						
Basismodel	331.70	9	36.86	.38	1.35	.00
Vrijlaten van methode factor	6.77	8	.85	.96	.03	.52
Groep 5						
Basismodel	56.08	4	14.02	.76	.41	.00
Vrijlaten van methode factor	2.07	3	.69	.98	.0	.60
$\xi_3 \xi_4$						
Groep 3/4						
Basismodel	6.44	4	1.61	.98	.06	.36
Groep 5						
Basismodel	10.28	5	2.06	.93	.13	.11

1996). Er is een belangrijke taak voor leerkrachten weggelegd bij het tijdig signaleren van problemen op het gebied van sociale participatie. Een betrouwbaar en valide signaleringsinstrument is daarbij onontbeerlijk. In deze studie stond de operationalisering van het begrip sociale participatie en de ontwikkeling en toetsing van een signaleringsinstrument voor leerkrachten centraal.

Op basis van een literatuurverkenning werden vier thema's onderscheiden binnen sociale participatie: vriendschappen/relaties, contacten/interacties, sociale zelfperceptie van de leerling en acceptatie door klasgenoten. Deze thema's vormen de basis voor de 'Vragenlijst Sociale Participatie' (VSP), een signaleringsinstrument dat is bedoeld om leerkrachten te ondersteunen bij het maken van accurate beoordelingen van de sociale participatie van leerlingen met beperkingen. De uitkomsten van een onderzoek naar de psychometrische kwaliteiten van de VSP lijken veelbelovend. Uit een Mokken Schaalanalyse komt naar voren dat de vier subschalen van de VSP gematigd sterk tot sterk zijn. Het Dubbele Monotonie-Model is van toepassing op iedere subschaal, wat impliceert dat de subschaalscores op een ordinale schaal liggen en de items invariant geordend zijn. De betrouwbaarheid van de VSP als geheel bleek hoog te zijn en ook de vier subschalen zijn betrouwbaar bevonden. Tevens bleken er sterke aanwijzingen te zijn voor de discriminante validiteit van de vragenlijst. Een confirmatieve factoranalyse met behulp van LISREL verschaftte aanwijzingen voor de convergente validiteit van de VSP. De fitmaten voor ξ_1 (Vriendschap) en ξ_2 (Contact) waren voor alle groepen voldoende. De fitmaten voor ξ_3 (Sociale zelfperceptie) and ξ_4 (Acceptatie) waren weliswaar acceptabel voor groep 3 en 4, maar niet voor groep 5. Er kan op basis van de fitmaten geconcludeerd worden dat het model van sociale participatie theoretisch standhoudt voor groep 3 en 4, maar slechts gedeeltelijk voor groep 5.

Nader onderzoek naar de betrouwbaarheid en validiteit van de VSP, met name voor groep 5, is noodzakelijk. Daar is al een start mee gemaakt door Martens, Vanmarssenille, Bossaert en Petry (2012) die onlangs vastgesteld hebben dat de interbeoordelaars-

betrouwbaarheid van de VSP over alle subschalen varieert van .60 tot .63 (resp. de leerling met beperkingen en een aselekt gekozen leerling) en dat de interbeoordelaarsbetrouwbaarheid over de afzonderlijke subschalen varieert van .36 tot .69. Ook in het voortgezet onderwijs werden vergelijkbare interbeoordelaarsbetrouwbaarheden aangetoond (Vertessen, 2012). Daarnaast is onderzoek naar de praktische toepasbaarheid van de VSP een belangrijke volgende stap. Zo is het van belang om de betekenis van scores nader te onderzoeken, zodat voor leerkrachten duidelijk wordt wanneer een score reden tot zorg is en als risicoscore beschouwd moet worden. Daarna zou gekeken moeten worden naar welke interventies vervolgens ingezet kunnen worden om de sociale participatie van leerlingen te bevorderen.

Een belangrijke kanttekening die bij dit onderzoek geplaatst kan worden, heeft betrekking op de opsplitsing in twee deelstudies. Het zou voor het bepalen van de convergente validiteit van de VSP gunstiger geweest zijn wanneer bij alle leerlingen, alle vier instrumenten zouden zijn gehanteerd, in plaats van twee. In dat geval was het mogelijk geweest om de correlatie van alle vier latente factoren gelijktijdig te onderzoeken. Er is echter bewust gekozen voor de opdeling in twee deelstudies, om de leerlingen en leerkrachten niet teveel te belasten.

Ondanks de bovenstaande kanttekening, kan geconcludeerd worden dat de VSP een instrument is dat veelbelovend lijkt. In het kader van Passend Onderwijs wordt verwacht dat het aantal leerlingen met beperkingen dat regulier onderwijs volgt in de komende jaren verder toe zal nemen. Uit tal van internationale studies (Freeman & Alkin, 2000; Frostad & Pijl, 2007), maar ook uit Nederlands onderzoek (Koster, Pijl, Nakken & Van Houten, 2010) blijkt dat de sociale participatie van leerlingen met beperkingen, in vergelijking met leerlingen zonder beperkingen, minder positief verloopt. Plaatsing van leerlingen met beperkingen in het reguliere onderwijs leidt niet automatisch tot het ontstaan van vriendschappen, of tot een toename van acceptatie (Koster et al., 2010). Het tijdig signaleren van problemen op het gebied van sociale participatie is, gezien de ontwikkelingen in

het onderwijs, derhalve van groot belang. De VSP kan leerkrachten helpen om problemen tijdig te signaleren. Dit is extra waardevol, omdat uit verschillende studies (De Monchy et al., 2004; Scheepstra, Nakken, & Pijl, 1999; Whitney, Smith, & Thompson, 1994) is gebleken dat leerkrachten de sociale participatie van leerlingen met beperkingen te positief inschatten en dat ze negatieve gedragingen (zoals pesten) onvoldoende opmerken. Wanneer leerlingen geïsoleerd raken of worden verworpen door klasgenoten, heeft dat negatieve gevolgen voor hun verdere functioneren (zie bijv. Asher & Coie, 1990; Ollendick, Weist, Borden, & Greene, 1992). Het is van belang dat leerkrachten tijdig een juiste inschatting maken van de sociale participatie van een leerling, zodat ze, indien nodig, passende maatregelen kunnen treffen.

Doordat de VSP uit concrete items bestaat die gericht zijn op de dagelijkse praktijk, wordt verwacht dat dit de accuraatheid van de oordelen van leerkrachten bevordert. Met name risicoscores zullen de alertheid van leerkrachten vergroten, waarna ze aanvullende maatregelen kunnen treffen. Het gevolg is dat leerlingen die moeilijkheden ervaren op het gebied van sociale participatie, eerder worden opgemerkt.

Een groot voordeel van de VSP ten opzichte van andere instrumenten is dat het recht doet aan de veelzijdigheid van sociale participatie, door inzicht te verschaffen in de vier thema's van sociale participatie. Door de subschaalscores in ogenschouw te nemen, kunnen leerkrachten meer inzicht krijgen in de verschillende thema's van sociale participatie. Verwacht wordt dat een dergelijk perspectief zal resulteren in een meer gebalanceerde kijk op de sociale participatie van leerlingen, wat startpunten biedt voor gedifferentieerde interventies.

Het volgen van de sociale ontwikkeling van leerlingen zou net zo vanzelfsprekend moeten zijn als het volgen van de leerontwikkeling van leerlingen. Leerprestaties van leerlingen worden veelal bijgehouden in leerlingvolgsystemen. Voor het sociaal functioneren van leerlingen is dit nog veel minder gebruikelijk. Gezien de ontwikkeling naar passend onderwijs wordt het structureel volgen van de sociale ontwikkeling van leerlingen alleen maar

belangrijker en zou het goed zijn wanneer scholen hier een visie op zouden ontwikkelen.

Literatuur

- Asher, S.R., & Coie, J.D. (1990). *Peer rejection in childhood*. Cambridge: Cambridge University Press.
- Bagwell, C.L., Newcomb, A.F., & Bukowski, W.M. (1998). Preadolescent friendship and peer rejection as predictors of adult adjustment. *Child Development, 69*, 140-153.
- Berndt, T.J. (2004). Children's friendships: shifts of a half-century in perspective on their development and their effects. *Merrill-Palmer Quarterly, 50*, 206-223.
- Berndt, T.J., & Burgoyne, L. (1996). Social Self-Concept. In B.A. Bracken (Ed.), *Handbook of Self-Concept* (pp. 171-209). New York: John Wiley & Sons, Inc.
- Borgatti, S.P., Everett, M.G., & Freeman, L.C. (1999). *UCINET for Windows*. Natick: Analytic Technologies.
- Byrne, B.M. (1989). Multigroup comparisons and the assumption of equivalent construct validity across groups: methodological and substantive issues. *Multivariate Behavioral Research, 24*, 503-523.
- Chamberlain, B., Kasari, C., & Rotheram-Fuller, E. (2007). Involvement or isolation. The social networks of children with autism in regular classrooms. *Journal of Autism and Developmental Disorders, 37*, 230-242.
- De Monchy, M., Pijl, S.J., & Zandberg, T.J. (2004). Discrepancies in judging social inclusion and bullying of pupils with behaviour problems. *European Journal of Special Needs Education, 19*, 317-330.
- Embretson, S.E., & Reise, S.P. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.
- Ferguson, D.L. (2008). International trends in inclusive education: The continuing challenge to teach each one and everyone. *European Journal of Special Needs Education, 23*, 109-120.
- Freeman, S.F.N., & Alkin, M. (2000). Academic and social attainments of children with mental retardation in general education and special education settings. *Remedial and Special Education, 21*, 3-18.

- Frosted, P., & Pijl, S.J. (2007). Does being friendly help in making friends? The relation between the social position and social skills of pupils with special needs in mainstream education. *European Journal of Special Needs Education, 22*, 15-30.
- Gregory, R.J. (2011). *Psychological testing*. Boston: Pearson.
- Gresham, F.M. (1982). Social interactions as predictors of children's likeability and friendship patterns: a multiple regression analysis. *Journal of Behavioral Assessment, 4*, 39-54.
- Harper, G.F., Maheady, L., Mallette, B., & Karnes, M. (1999). Peer tutoring and the minority child with disabilities. *Preventing School Failure, 43*, 47-52.
- Hoza, B., Mrug, S., Pelham, W.E., Greiner, A.R., & Gnagy, E.M. (2003). A friendship intervention for children with attention-deficit/hyperactivity disorder: Preliminary findings. *Journal of Attention Disorders, 6*, 87-98.
- Jöreskog, K.G., & Sörbom, D. (2007). *Lisrel 8.80 for Windows (Computer software)* Lincolnwood, IL: Scientific Software International, Inc.
- Kamps, D.M., Dugan, E., Potucek, J., & Collins, A. (1999). Effects of cross-age peer tutoring networks among students with autism and general education students. *Journal of Behavioral Education, 9*, 97-115.
- Karsten, T., Peetsma, T., Roeleveld, J., & Vergeer, M. (2001). The Dutch policy of integration put to the test: differences in academic and psychosocial development of pupils in special and mainstream education. *European Journal of Special Needs Education, 16*, 193-205.
- Koster, M., Minnaert, A.E.M.G., Nakken, H. Pijl, S.J., & Houten, E.J. van (2011). Assessing social participation of students with special needs in inclusive education: validation of the Social Participation Questionnaire. *Journal of Psychoeducational Assessment, 29*, 199-213.
- Koster, M., Nakken, H., Pijl, S.J., & Houten, E.J. van (2009). Being part of the peer group: a literature study focusing on the social dimension of inclusion in education. *International Journal of Inclusive Education, 13*, 117-140.
- Koster, M., Pijl, S.J., Nakken, H., & Houten, E.J. van (2010). Social participation of students with special needs in regular primary education in the Netherlands. *International Journal of Disability, Development and Education, 57*, 59-75.
- Koster, M., Timmerman, M.E., Nakken, H., Pijl, S.J., & Houten, E.J. van (2009). Evaluating social participation of pupils with special needs in regular primary schools: examination of a teacher questionnaire. *European Journal of Psychological Assessment, 25*, 213-222.
- Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174.
- Larrivee, B., & Horne, M.D. (1991). Social status: a comparison of mainstreamed students with peers of different ability levels. *The Journal of Special Education, 25*, 90-101.
- Margalit, M., & Efrati, M. (1996). Loneliness, coherence and companionship among children with learning disorders. *Educational Psychology, 16*, 69-79.
- Martens, S., Vanmarsenille, C., Bossaert, G., & Petry, K. (2012). De betrouwbaarheid en validiteit van de SPQ in het Vlaamse basisonderwijs. KU Leuven: Masterproef.
- Mokken, R.J. (1971). A theory and procedure of scale analysis with applications in political research. The Hague, The Netherlands: Mouton.
- Molenaar, I.W., & Sijtsma, K. (2000). *User's manual MSP5 for Windows*. Groningen, The Netherlands: iecProGAMMA.
- Montague, M., & Rinaldi, C. (2001). Classroom dynamics and children at risk: a followup. *Learning Disability Quarterly, 24*, 75-83.
- Nunnally, J.C. (1967). *Psychometric Theory*. New York: McGraw-Hill Book Company.
- Odom, S.L. (2000). Preschool inclusion: what we know and where we go from here. *Topics in Early Childhood Special Education, 20*, 20-27.
- Ollendick, T.H., Weist, M.D., Borden, M.C., & Greene, R.W. (1992). Social status and academic, behavioural, and psychological adjustment: a five-year longitudinal study. *Journal of Consulting and Clinical Psychology, 60*, 80-87.
- Parker, J.G., & Asher, S.R. (1993). Friendship and friendship quality in middle childhood: links with peer group acceptance and feelings of loneliness and social dissatisfaction. *Developmental Psychology, 29*, 611-621.
- Pavri, S., & Luftig, R. (2000). The social face of inclusive education: are students with learning disabilities really included in the classroom? *Preventing School Failure, 45*, 8-14.

- Rossum, J.H.A., van, & Vermeer, A. (1992). Het meten van waargenomen competentie bij basisschoolkinderen door middel van een platen-test. *Bewegen & Hulpverlening*, 9, 198-212.
- Rozendaal, J.S., Minnaert, A., & Boekaerts, M. (2003). Motivation and self-regulated learning in secondary vocational education: information-processing type and gender differences. *Learning and Individual Differences*, 13, 273-289.
- Scheepstra, A.J.M. (1998). *Leerlingen met Downs syndroom in de basisschool*. Groningen: Stichting Kinderstudies.
- Scheepstra, A.J.M., Nakken, H., & Pijl, S.J. (1999). Contact with classmates: the social position of pupils with Down's Syndrome in Dutch mainstream education. *European Journal of Special Needs Education*, 14, 212-220.
- Sijtsma, K., & Molenaar, I.W. (2002). Introduction to nonparametric item response theory (Vol. 5). Thousand Oaks, CA: Sage.
- Sloper, T., & Tyler, S. (1992). Integration of children with severe learning difficulties in mainstream schools: evaluation of a pilot study. *Educational and Child Psychology*, 9, 35-45.
- Veerman, J.W., Straathof, M.A.E., Treffers, P.D.A., Van den Bergh, B.R.H., & Ten Brink, L.T. (1997). *Handleiding Competentie Belevingschaal voor Kinderen*. Lisse: Swets & Zeitlinger.
- Vellutino, F.R., Tunmer, W.E., Jaccard, J.J., & Chen, R. (2007). Component of reading ability: multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading*, 11, 3-32.
- Vertessen, N. (2012). *Het onderzoeken van de bruikbaarheid van de SPQ bij GON leerlingen in het secundair onderwijs*. KU Leuven: Masterproef.
- Whitney, I., Smith, P.K., & Thompson, D. (1994). Bullying and children with special educational needs. In P.K Smith & S. Sharp (Eds.), *School Bullying: insights and perspectives* (pp. 213-240). New York: Routledge.

Manuscript aanvaard op: 11 januari 2013

Noot

- 1 Aan de hier gerapporteerde studies is een belangrijke bijdrage geleverd door prof. dr. Han Nakken, overleden in 2010.

Auteurs

Marloes Koster is werkzaam als gedragsdeskundige bij de Prof. W.J. Bladergroenschool te Groningen. **Els van den Bosch** is als docent/onderzoeker verbonden aan de afdeling Orthopedagogiek van de Rijksuniversiteit Groningen. **Sip Jan Pijl** werkt als hoogleraar aan de afdeling Orthopedagogiek van de Rijksuniversiteit Groningen en is tevens parttime hoogleraar bij de NTNU Universiteit van Trondheim in Noorwegen.

Correspondentieadres: Els van den Bosch, Rijksuniversiteit Groningen, Faculteit GMW, afdeling Orthopedagogiek, Grote Rozenstraat 38, 9712 TJ Groningen. E-mail: e.j.van.den.bosch@rug.nl

Abstract

Evaluating the social dimension of inclusive education: the development of a teacher questionnaire to assess the social participation of students with special needs

This study addresses the construction and examination of a teacher questionnaire, the Social Participation Questionnaire (SPQ), to assess the social participation of students with special needs in Grades 1, 2 and 3 of regular primary education. The SPQ consists of four subscales, related to four key themes of social participation: "friendships", "contacts/interactions", "student's social self-perception" and "acceptance by classmates". A Mokken Scale Analysis showed that the subscales appeared intermediate to strong. As the double monotonicity model turned out to fit well for each subscale, the subscale scores are on an ordinal scale and the statements are invariantly ordered. Furthermore, the SPQ was found to be reliable and evidence was obtained for its discriminant validity. Finally, a confirmatory factor analysis using LISREL provided evidence for the SPQ's convergent validity. The fit indexes suggest that the model of social participation is theoretically viable for Grades 1 and 2 but only partly viable for Grade 3.