

Masterscriptie Orthopedagogiek

# **De betrouwbaarheid van de “Student Behavior Observation Form”**

*Onderzoek naar het betrouwbaar meten van het leergedrag van leerlingen in het voortgezet onderwijs*

Universiteit Leiden

Student: Nicole A. Maccow

Studentnummer: 0544639

Scriptie Begeleider: Prof. Dr. C. A. Espin

Tweede Beoordelaar: Dr. M. Verhallen

Datum: 28-03-2012

## **Voorwoord**

Voor u ligt mijn masterscriptie ter afronding van de opleiding Orthopedagogiek aan de Universiteit van Leiden.

Het werken aan mijn scriptie was een hele leerzame maar ook soms stressvolle ervaring. Mede door deze ervaring is mijn geloof in God enorm versterkt. Verschillende personen in omgeving hebben mij een steuntje in de rug geboden op momenten dat ik het even niet meer zag zitten. Ik wil mijn ouders bedanken die mij sinds mijn komst in Nederland er alles aan hebben gedaan om mij te ondersteunen. Bedankt voor jullie onvoorwaardelijke steun, liefde en geloof in mij. Ook wil ik mijn zus bedanken voor de vele studie weekenden . Mijn overige familieleden, vrienden en collega's wil ik ook bedanken voor alle opbeurende woorden en hulp. Tevens wil ik mijn vriend bedanken voor zijn liefde en steun. Als laatst wil ik mijn begeleiders mevrouw Espin en mevrouw Verhallen bedanken voor hun begeleiding.

Nicole Maccow

Schiedam, 2012

### **Abstract**

Het overkoepelend doel van dit onderzoek is het ontwikkelen van een betrouwbaar en valide monitoringsysteem dat het leergedrag van leerlingen op een eenvoudige wijze in kaart kan brengen. In dit onderzoek wordt onderzocht welke leergedragingen ( “aan taak”, “niet aan taak”, “storend gedrag” en “actief leergedrag”) van de Student Behavior Observation Form (SBOF) betrouwbaar zijn. Om dit vast te stellen is de interbeoordelaarsbetrouwbaarheid en test-hertestbetrouwbaarheid berekend. Ook is gekeken in welke mate het leergedrag (van de twee onderzoeksdagen) per individu verschilt en of het leergedrag verschillend is per type leerling. De steekproef bestond uit 30 brugklasleerlingen uit het voortgezet onderwijs in de leeftijd van 13 tot 15 jaar. De interbeoordelaarsbetrouwbaarheid van de leergedragingen “aan taak”, “niet aan taak” bleek hoog (>80%). Uit de correlatieanalyse bleek dat de scores van het “aan taak” en “niet aan taak” leergedrag op de twee onderzoeksdagen redelijk samen hangen. Meer dan de helft van de leerlingen lieten matig of veel verschil in het “aan taak” leergedrag zien. Dit verschilde niet per type (zeer goed, goed of minder goed) leerlingen.

## **1. Inleiding**

### *1.1 Problemen in het onderwijs*

Leerlingen die betrokken zijn bij hun schooltaken zijn vaak succesvol in hun schoolcarrière. Uit onderzoek naar de betrokkenheid in het onderwijs blijkt dat leerlingen die zich concentreren op hun schoolwerk, weinig verzuimen, zich houden aan de regels en weinig probleemgedrag vertonen meestal goed presteren (Fredricks, Blumenfeld, & Paris, 2004). Leerlingen kunnen op verschillende vlakken betrokkenheid laten zien. Betrokkenheid wordt verdeeld in drie componenten; cognitief ( het vermogen van leerlingen om hun eigen leerproces te sturen bijvoorbeeld het indelen van schoolwerk), emotioneel (de mate van identificatie met school) en gedragsmatig ( het deelnemen aan activiteiten en handelingen die betrekking hebben op school en het leren). Leerlingen die minder betrokken zijn vertonen vaker probleemgedrag (gedragingen die het leerproces verstoren zoals spijbelen, overtreden van regels) en presteren minder goed dan leerlingen die meer betrokken zijn. (Wang & Holcombe, 2010). Ze hebben vaker een ongemotiveerde werkhouding, hechten minder waarde aan hun schoolprestaties en zijn minder betrokken bij hun schoolwerk. Deze leerlingen hebben vaak externe motivatie nodig in de vorm van straffen en belonen om hun werkhouding aan te passen (Fredricks, et al., 2004).

Probleemgedrag heeft vaak ook een functie, leerlingen vragen op een niet adequate wijze aandacht van de leerkracht of proberen met hun gedrag onder hun taken uit te komen. Leerlingen die minder betrokken zijn, en probleemgedrag vertonen hebben een grotere kans op leerachterstand en het niet volledig voltooien van een opleiding (Wang & Holcombe, 2010). De afgelopen jaren is een stijging van het aantal voortijdig schoolverlaters geconstateerd en blijkt de taal- en rekenvaardigheid van leerlingen na het beëindigen van het voortgezet onderwijs beneden het verwachte niveau. Tevens zijn het aantal verwijzingen naar het speciaal onderwijs gestegen en komen meer leerlingen in aanmerking voor extra ondersteuning in de klas bijvoorbeeld leerwegondersteunend onderwijs (Traag & van der Velden, 2011). Er zijn verschillende factoren die aan de onderwijsleerproblematiek ten grondslag liggen. Zo blijkt dat ongeveer 10% tot 15% van schoolgaande kinderen kenmerken vertonen van gedragsstoornissen zoals oppositioneel-opstandig gedragsstoornissen of aandachtstekort en hyperactiviteitstoornis (Jacobs, et al., 2000). Externaliserend gedrag (zoals opstandigheid, spijbelen, druk en ongeconcentreerd gedrag) als meest problematisch ervaren door leerkrachten (Kaplan, Gheen & Midgley, 2002). Echter wordt de internaliserende gedragsproblematiek (zoals teruggetrokken en depressief gedrag) minder vaak als problematisch bestempeld. Een verklaring hiervoor is dat deze leerlingen minder op de voorgrond staan, en minder overlast bezorgen. Tevens ervaren leraren bij deze groep leerlingen minder handelingsverlegenheid dan bij het aanpakken van leerlingen die meer externaliserend gedrag vertonen (Onderwijsraad, 2010). Het externaliserend probleemgedrag is vaak chronisch van aard en heeft een

negatieve invloed op de dagelijkse klassikale leerprocessen en leerprestaties van de (mede) leerlingen. Om deze problemen terug te dringen, is het van belang om de problemen die leraren in de onderwijsleersituatie tegenkomen en onderliggende probleemfactoren in kaart te brengen (Olympia, Heathfield, Jenson, & Clark, 2002).

### *1.2 Interventies binnen het onderwijs*

Om de problemen binnen het onderwijs op een effectieve wijze op te lossen heeft men de juiste middelen nodig (McIntosh, Frank & Spaulding, 2010). Men moet over een systeem beschikken waarmee risicoleerlingen en diens leergedrag in kaart gebracht kan worden en kan bepalen welke interventie het beste aansluit. In het huidige onderwijs wordt veelvuldig gebruik gemaakt van probleemoplossende modellen. Probleemoplossende modellen zijn ontworpen om sturing te geven aan aanpak en monitoren van o.a. leerprestatie (Jacobs, et al., 2000). Een probleemoplossend model fungeert als kader voor een systematische aanpak die o.a. de vaardigheden en beperkingen van leerlingen identificeert. Tevens wordt hierin de keuze voor evidence-based interventies gemaakt en wordt de studievoortgang en de effectiviteit van interventies geëvalueerd (Hughes & Dexter, 2011). Een kritische kanttekening is echter dat de meest gebruikte modellen vaak uitsluitend gericht zijn op academische prestaties. Er is nauwelijks oog voor het functioneren op emotioneel, geestelijk en gedragmatige vlak, ondanks de wetenschap dat problemen op deze gebieden een negatieve invloed kunnen hebben op de leerprestaties van leerlingen. Door het niet onderkennen van problemen op deze gebieden ontvangen leerlingen vaak niet de nodige hulp (Chafouleas, Volpe, Gresham, & Cook, 2010b).

### *1.3 Response to Intervention (RTI)*

Het RTI is een preventie en interventie model dat toegespitst is op de ondersteuningsbehoeftes van leerlingen. De intensiteit van de interventies en monitoring stijgt naarmate er sprake is van een groter ondersteuningsbehoefte. In de eerste stap van het RTI model (tier 1), de universele screening, worden alle leerlingen drie keer per jaar gescreend. De leerlingen die tot de risicogroep behoren (25%) worden maandelijks gemonitord. Bij de tweede stap (tier 2) worden leerlingen (10-20%) waarbij een gespecialiseerde interventie wordt toegepast wekelijks gemonitord. Dit geldt ook voor de leerlingen waarbij een meer intensieve vorm van interventie wordt toegepast (5-7%). Ongeveer 2 tot 5 % van de risicogroep wordt uiteindelijk doorverwezen naar het speciaal onderwijs (Hughes & Dexter, 2011). Het RTI is effectief bevonden in het onder andere signaleren van leerachterstanden bij leerlingen en het terugdringen van het aantal voortijdig schoolverlaters. Naast het monitoren van academische vorderingen ligt de nadruk binnen het RTI ook op de aanwezigheid van emotionele- en gedragsproblemen. Tevens kan aan de hand van dit model bepaald worden of een leerling in aanmerking kan komen voor speciaal onderwijs, de daarmee samenhangende dienst- of hulpverlening, en of een bepaalde interventie aangepast

of geïntensiveerd dient te worden (Gresham, 2005).

Voortgangsmonitoring is een essentieel en veelvoorkomend proces binnen het RTI model en wordt ingezet om de voortgang of prestatie op een of meerdere gebieden zoals lezen, rekenen of sociaal gedrag in kaart te brengen. Het blijkt dat voortgangsmonitoring verschillende voordelen heeft mits het adequaat geïmplementeerd wordt. Enkele voordelen zijn dat er een betere aansluiting is bij het niveau van de leerling. Ook kan men op verantwoording afleggen voor de gekozen aanpak en duidelijke verslaglegging en documentatie van de voortgang leveren. Het adequaat implementeren van de voortgangsmonitoring en de interventies kan leiden tot minder verwijzingen naar het speciaal onderwijs (Hughes & Dexter, 2011).

#### *1.4 Observeren van leergedrag*

Het implementeren van voortgangsmonitoring binnen het RTI model is echter complex (Gresham, 2005). Het blijkt dat er nog weinig eenduidigheid is over de geschiktheid van de instrumenten (Chafouleas, Volpe, Gresham & Cook, 2010b). Er is een scala aan onderzoeksmiddelen en methodes zoals interviews houden, vragenlijsten invullen en observeren die ingezet kunnen worden voor het in kaart brengen van het leergedrag, maar er zijn weinig gestandaardiseerde instrumenten (Hintze & Matthews, 2004). Beoordelingsschalen die primair bedoeld zijn om een bepaalde vaardigheid te toetsen worden vaak gewijzigd en aangepast om als voortgangsmonitoring te dienen vooral wanneer de schalen meerdere items bevatten die specifieke gedragsveranderingen in kaart kunnen brengen. Door middel van observaties wordt het gedrag van risicoleerlingen en de voortgang van interventiebehandeling gemonitord (Jacobs et al., 2000). Direct Behavior Rating (DBR) en de Systematic Direct Observation (SDO) zijn de meest gebruikte observatiemethoden binnen het onderwijs. SDO houdt in dat men op een systematische wijze, aan de hand van een observatieschema observeert hoe vaak een gedrag voorkomt. DBR bevat kenmerken van zowel de SDO (systematische observatie wijze) en een gedragsbeoordelingsschaal (Volpe, Mc Conaughy & Hintze, 2009) Vele middelen die binnen het onderwijs worden gebruikt bijvoorbeeld heen- en weerschrift, gedrags- en voortgangskarten vallen allen hieronder ( Christ, Riley-Tillman, Chafouleas & Bioce, 2010). Uit onderzoek blijkt, dat de ontwikkeling van observatieinstrumenten vertraagd wordt door een gebrek aan consensus omtrent het gebruik van passende gedragsdefinities en de observatie methoden (Pakarinen et al.,2010).

### *1.5 Geschiktheid van gedragsdefinities*

Nock en Kurtz (2005) stellen in hun onderzoek dat het bij het ontwikkelen van een observatieinstrument allereerst belangrijk te onderzoeken welk gedrag men wil observeren. Het gedrag dient volgens hen observeerbaar en meetbaar te zijn. Met observeerbaar wordt bedoeld dat het gedrag zichtbaar is (objectief). Met meetbaar wordt bedoeld dat men de frequentie waarin het gedrag zich voordoet in kaart moet kunnen brengen. Als laatste is het van belang dat het gedrag gespecificeerd wordt. Men moet van tevoren vast stellen wanneer er wel of geen sprake is van een bepaald gedrag (concretiseren). Gedragingen die minder frequent voorkomen en erg opvallend zijn, zoals vechten, blijken minder geschikt te zijn voor een observatieinstrument omdat de kans groot is dat het gedrag zich niet voor zal doen tijdens het observatiemoment.

Echter blijken gedragingen zoals het weigeren om een taak op te volgen of niet opletten geschikter omdat ze zich vaker voordoen in een leersituatie en minder opvallen. Daarnaast kiest men de algemene definities zoals “academische betrokkenheid” boven de meer specifieke gedragingen. Het gebruik van meer specifieke gedragingen zoals vinger opsteken, luisteren naar de leerkracht, opvolgen van instructies, opstaan zonder toestemming, zijn minder geschikt omdat men een groot aantal gedragingen die soms tegelijkertijd voorkomen moet scoren (Nock & Kurtz, 2005). Uit onderzoek naar de gedragstermen blijkt dat “academische betrokkenheid” (actief of passieve deelname aan lesactiviteiten) en “storend gedrag” (al het gedrag die interfereert met lessituatie) accurater gescoord wordt. De gedraging “volgzaamheid” (sociaal gewenste respons op verzoek van de leerkracht) wordt minder accuraat gescoord. Een verklaring hiervoor is dat “academische betrokkenheid” en “volgzaamheid” in de praktijk moeilijk van elkaar te onderscheiden is (Christ, Riley-Tillman, Chafouleas & Jaffery, 2011). Verder blijkt het belangrijk te zijn dat de gedragingen veranderingen inzichtelijk maakt daar waar het instrument ingezet zal worden voor het monitoren van effecten van de interventies op het leergedrag van leerlingen (Walker, Ramsey & Gresham, 2003). Het niet nauwkeurig omschrijven van gedragsdefinities, kan de betrouwbaarheid van de data negatief beïnvloeden (Riley-Tillman, Christ, Chafouleas, Briesch & Lebel, 2009). Als men middels een observatie-instrument het leergedrag en vorderingen van leerlingen in kaart wil brengen en de keuzes voor een bepaalde interventie hierop wil baseren is het van belang dat de verkregen informatie betrouwbaar is (Gresham, 2005).

### *1.6 Betrouwbaarheid van het voortgangsmonitoringssysteem*

Een observatie instrument die in het onderwijs wordt gebruikt moet voldoen aan een aantal voorwaarden. Het instrument moet eenvoudig in het gebruik zijn, dus niet tijdrovend of moeilijk scoringsysteem bevatten. Ook moet het instrument gevoelig zijn voor gedragsverandering en niet interventie afhankelijk zijn (Gresham, et al. 2010). In Nederland zorgt de Commissie Testaangelegenheden Nederland (COTAN) voor dergelijke psychometrische waardering. Volgens de COTAN dienen instrumenten aan een aantal psychometrische voorwaarden te voldoen alvorens het daadwerkelijk ingezet kan worden. Een essentiële voorwaarde voor elk instrument is de betrouwbaarheid. Zo moet een instrument dat ingezet wordt voor voortgangscontrole op individueel niveau, voldoende betrouwbaar zijn (minimaal  $r = .70$ ). Met betrouwbaarheid wordt bedoeld de herhaalbaarheid en de consistentie van de scores (NIP, 2010). Er zijn verschillende aspecten met betrekking tot de betrouwbaarheid van een instrument zoals de interbeoordelaarsbetrouwbaarheid (als verschillende observatoren hetzelfde signaleren). Men onderzoekt of de scores van de verschillende beoordelaars op verschillende meetmomenten met elkaar overeenkomen. Meestal wordt voorafgaand aan het onderzoek bepaald hoe hoog deze betrouwbaarheid moet zijn aan de hand van bijvoorbeeld het percentage overeenkomst. Vaak wordt er een percentage overeenkomst van tussen 80 en 90% gehanteerd (Volpe, et al., 2009). Er zijn verschillende onderzoeken gedaan naar de interbeoordelaarsbetrouwbaarheid van gedragsobservatie instrumenten. In het onderzoek van Martin et al. (2010) is de beoordelaarsbetrouwbaarheid van de observatie instrument, Teacher-Pupil Observation Tool (T-POT) in kaart gebracht. De T-POT bevat 75 items die gericht zijn op zowel kenmerken van de leerling als die van de leerkracht. De observatoren hadden ruime ervaring en een uitgebreide training (ongeveer 24 uur) gehad. Uit de resultaten blijkt dat de scores van de verschillende beoordelaars met elkaar overeen kwamen, het percentage overeenkomst bedroeg 78%.

In een soortgelijk onderzoek van Hintze en Matthews (2004) zijn de psychometrische waarden van Systematic Direct Observation. Er is tiendaagse onderzoek gedaan naar het “aan taak” en “niet aan taak” leergedrag en is de interbeoordelaarsbetrouwbaarheid van de scores van 55 observaties is berekend. De observatoren waren 7 psychologie studenten die ruime ervaring hadden in het observeren aan de hand van de systematic direct observation methode. Voorafgaand aan het onderzoek hebben de observatoren een 4 uur durende voorbereiding gehad. Er werd aan de hand van videobeelden geoefend met de momentary-time sampling methode, waarin in een bepaald tijdspad geobserveerd werd of bepaalde gedragingen voor kwamen. Tevens hebben de observatoren uitleg gehad van de te observeren leergedragingen en zijn voorbeelden van de twee leergedragingen gegeven. De interbeoordelaarsbetrouwbaarheid was goed met een percentage overeenkomst 90%.



Een andere vorm van betrouwbaarheid is de test-hertestbetrouwbaarheid, waarbij men onderzoekt in hoeverre de data die men verkrijgt op verschillende meetmomenten met elkaar in overeenstemming met zijn (NIP, 2010). Aan de hand van een generaliseerbaarheidstudie kan men de variabiliteit van verschillende facetten (bijvoorbeeld beoordelaar, item, tijd, methode en setting) die invloed kunnen hebben op de betrouwbaarheid in kaart brengen (Hintze & Mathew, 2004).

Ook hebben Hintze en Mathew (2004) onderzoek gedaan naar de test-hertestbetrouwbaarheid van de Systematic Direct Observation en gekeken in hoeverre het “aan taak” en “niet aan taak” gedrag betrouwbaar was. De vuistregel die wordt gehanteerd is dat een meting bij een systematische variantie van minimaal 70% en een foutenvariantie van 30% betrouwbaar is (Leary, 2004). Hieruit blijkt dat er significante verschillen zijn tussen het gedrag van leerlingen op de verschillende onderzoeksdagen en tijdstippen. Het grootste deel van de variantie 67% in het “aan taak” en niet “aan taak” gedrag te verklaren is door de individuele verschillen tussen de participanten. De foutenvariantie was 33%. Leerlingen lieten veel verschil zien in het “aan taak” en niet “aan taak” leergedrag op de twee onderzoeksmomenten (Hintze & Matthew, 2004).

Briesch, Chafouleas en Riley-Tillman (2010) hebben ook onderzoek gedaan naar de systematisch directe observatie (SDO) en de directe observatie (DBR). In het onderzoek werd gekeken naar de academische betrokkenheid en de verschillende factoren die van invloed zijn op de test-hertestbetrouwbaarheid. Hieruit bleek dat er sprake was van een interactie te zijn tussen het individu, de dag en observatiemoment. Deze interactie verklaarde 30% van de foutenvariantie. In het geval van het DBR werd 13% van de variantie verklaard door tijdgebonden factoren en interacties. Dit betekent dat de verschillende facetten zoals dag en persoon invloed hebben op het geobserveerde gedrag, in dit geval blijkt de DBR minder betrouwbaar. Een verklaring voor dit verschil is dat er sprake kan zijn van een halo-effect, waarbij de scores beïnvloed worden door het beeld dat de observator heeft van een leerling of diens gedrag en door eerdere scores. Het halo-effect kan geminimaliseerd worden door het gedrag meteen te scoren (dus niet achteraf invullen) en door het aantal metingen per dag te minderen. Behalve de betrouwbaarheid van de data is het van belang dat een instrument voor voortgangsmonitoring geschikt is voor alle leerlingen. Naast de tijdsfactor en verschillen tussen observatoren blijkt dat leerlingkenmerken (aan taak of niet) van invloed kunnen zijn op de betrouwbaarheid van de data.

Uit een vergelijkbare studie van Chafouleas, et al. (2010a) blijken de scores van de verschillende observatoren vrij consistent wanneer het ging om leerlingen die positief leergedrag vertoonden (weinig storend gedrag en hoge mate van academische betrokkenheid). Bij leerlingen die veel variatie in het leergedrag vertoonden was er sprake van minder overeenstemming tussen de observatoren. Hierbij dient men rekening te houden met het feit dat de steekproef van deze studie bestond uit leerlingen die veel

probleemgedrag vertoonden. De geobserveerde leergedragingen van deze leerlingen op de verschillende onderzoeksmomenten waren erg uiteenlopend. Bij het gebruik van observatie-instrumenten is het dus van belang om rekening te houden met verschillende factoren die invloed kunnen hebben op de betrouwbaarheid van de verkregen informatie.

### *1.7 Huidig onderzoek*

Het doel van dit onderzoek is het ontwikkelen van een betrouwbare en valide monitoringsysteem dat gebruikt kan worden door leerkrachten of interne begeleiders om het leergedrag van leerling op een eenvoudige wijze in kaart kan brengen. Om dit te kunnen bereiken wordt ten eerste onderzocht welke gedragingen als indicator kunnen fungeren om het algemeen leergedrag van leerlingen in kaart te brengen. Een belangrijk kenmerk is dat het gedrag op een betrouwbare wijze gemeten kan worden. In dit onderzoek zal de interbeoordelaarsbetrouwbaarheid, test-hertestbetrouwbaarheid en de betrouwbaarheid van het gedrag per individu belicht worden.

De centrale onderzoeksvraag van dit onderzoek luidt;

*Kunnen de gedragingen (“aan taak”, “niet aan taak”, “storend gedrag” en “actief leergedrag”) fungeren als betrouwbare indicatoren van een observatieformulier die de voortgang van het leergedrag in kaart moet brengen?*

*Welke gedragingen brengen het algemeen leergedrag van leerlingen op een betrouwbare wijze in kaart?*

Om de centrale onderzoeksvraag te beantwoorden zijn drie deelvragen geformuleerd.

Deelvraag 1: *Is er een overeenstemming tussen de scores van de verschillende observatoren (interbeoordelaarsbetrouwbaarheid)?* De hypothese is dat de scores van de verschillende observatoren overeen zullen komen. Er zal sprake zijn van een voldoende tot goede interbeoordelaarsbetrouwbaarheid. Het percentageovereenkomst zal tussen 80 en 90% zijn.

Deelvraag 2: *Is er een verschil tussen de scores van de eerste onderzoeksdag en de tweede onderzoeksdag (test-hertestbetrouwbaarheid)?* De hypothese is dat de scores van de twee onderzoeksdagen niet significant van elkaar zullen verschillen.

Deelvraag 3: *Is per leerling sprake van consistentie tussen het “aan taak” leergedrag van de twee onderzoeksmomenten? Verschilt de mate van consistentie per type leerling?* De verwachting is dat het “aan taak”leergedrag voor alle type leerlingen consistent zal zijn.

## 2. Methoden

### 2.1 Steekproef

De dataverzameling vond plaats op een middelgrote middelbare school in de binnenstad van Den Haag. De steekproef bestond uit 30 leerlingen. Onder de leerlingen bevonden zich 11 jongens (37%) en 19 meisjes (63%) in de leeftijd van 13 tot en met 15 jaar. De gemiddelde leeftijd van de leerlingen was 14,3 jaar ( $SD = .54$ ). De scholengroep bestaat uit verschillende locaties waar onderwijs wordt gegeven op verschillende niveaus, voorbereidend middelbaar beroepsonderwijs VMBO (theoretische kader, kader gericht onderwijs en leerwegondersteunend onderwijs), Hoger Algemeen Voortgezet Onderwijs HAVO en Voorbereidend Wetenschappelijk Onderwijs VWO. Van de leerlingen waren er 40% afkomstig van het VMBO en 60% van de HAVO/VWO. In totaal hadden 24 leerlingen Nederland als geboorteland, 3 leerlingen waren in Marokko geboren, 1 leerling in Turkije, 1 leerling in Irak. Van 1 leerling was het geboorteland onbekend.

Binnen de school was besloten welke klassen zouden deelnemen aan het onderzoek. Bij aanvang van het schooljaar werden alle ouders door middel van een brief op de hoogte gesteld dat er meerdere onderzoeken zouden plaatsvinden, en waren hiermee akkoord gegaan. De deelnemers werden niet beloond voor hun deelname aan het onderzoek.

De steekproef werd niet aselekt getrokken. De steekproef werd aan de hand van de Missing Value Analysis van het computerprogramma Statistical Package for the Social Sciences (SPSS 16) geselecteerd. De cases die minimaal twee minuten observatietijd hadden gedurende drie (dezelfde)lessuren op de eerste en tweede onderzoeksdag werden gebruikt voor dit onderzoek.

### 2.2 Geobserveerde gedragingen

De gedragingen in dit onderzoek zijn afkomstig van de Pupil Observation Procedure Form (POP). De POP is ontwikkeld om informatie te verkrijgen om de effectiviteit van de leerkrachten in kaart te kunnen brengen. De POP bevat 6 items, 3 verwijzen naar het gedrag van de student (“academische respons”, “storend gedrag” en “niet aan taak”). Deze leergedragingen van de student werden als volgt omschreven. De leergedraging “academische respons” verwijst naar het gedrag waarin de leerling verbaal of motorisch reageert op een instructie van de leerkracht. Met “storend gedrag” wordt bedoeld alle gedragingen die de les voor de leerling zelf of voor omgeving verstoren. Met “niet aan taak” wordt bedoeld het niet bezig zijn met de taak of opdracht (Espin & Yell, 1994). In dit onderzoek wordt de POP gebruikt om het leergedrag van leerlingen in het voortgezet in kaart te brengen. De Student Behavior Observation Form (SBOF) is een aangepaste versie van de Pupil Observation Procedure Form (POP). Het doel van de SBOF is het zo

nauwkeurig en objectief mogelijk informatie verkrijgen betreffende het leergedrag van leerlingen en de instructies van de leerkracht (leerkracht variabelen werden in dit onderzoek buiten beschouwing gelaten). Onder leergedrag wordt verstaan het gedrag wat verbonden is met en invloed heeft op het leren, zoals het luisteren naar instructie en het zelfstandig kunnen werken. De SBOF is door de studentonderzoekers vertaald in het Nederlands en bevat 4 items (“actief leergedrag”, “aan taak”, ‘niet aan taak’, en “storend gedrag”) die gericht zijn op het leergedrag van leerlingen. Er is sprake van “actief leergedrag” (ALG) op het moment dat de leerling op eigen initiatief zichtbaar stappen onderneemt om zich de lesstof eigen te maken. Een voorbeeld hiervan is gericht vragen stellen aan de leerkracht. Het gedrag dat correspondeert met de instructie van de leerkracht zoals lezen en aantekeningen maken, wordt als “aan taak”(AT) gescoord. De leerling is “niet aan taak”(NAT) wanneer hij niet bezig is met de opdracht die voorgelegd is of wanneer de aandacht van de leerling niet gericht is op zijn werk. Alle gedragingen die het leerproces van de leerling zelf of van medeleerlingen verstoord voor aangeduid als “storend gedrag” (SG)( zie bijlagen 2). De methoden van de observaties betreft de “momentary time sampling” waarbij per interval (10 seconden) wordt genoteerd of er sprake is van één de bovengenoemde gedragingen. De scores per minuut worden berekend door met de hand de scores per tien seconde op te tellen.

### *2.3 Leraren beoordelingsformulier*

De leraren beoordelingsformulier (LBF) is ontwikkeld door de studentonderzoekers om de beoordeling van de leerkracht met betrekking tot het leergedrag van leerlingen in kaart te brengen. Dit formulier is een aanvulling op de SBOF en bevat een 4-punt Lickertschaal met de antwoordcategorieën, 1 (zeer goed), 2 (goed), 3 (minder goed) en 4 (slecht). De mentor van de desbetreffende klas vulde het beoordelingformulier eenmalig in voor alle leerlingen in zijn mentorgroep in aan het einde van de laatste onderzoeksdag.

### *2.4 Procedure*

Voorafgaand aan het onderzoek had de thesisbegeleider vanuit de universiteit contacten gelegd met de onderzoeksbegeleider van de school. Er waren afspraken gemaakt met betrekking tot de tijdsplanning, duur en de inhoud van het onderzoek.

De observatoren zijn voorafgaand aan het onderzoek voorbereid op de universiteit. Voorafgaand aan de voorbereiding werd de oorspronkelijke (Engelse) versie van de leerling gedrag observatieformulier vertaald in de Nederlandse taal en zijn de items van het oorspronkelijke formulier aangepast. Vervolgens werd besproken hoe de verschillende gedragingen geïnterpreteerd konden worden en werd er geoefend aan de hand van een video met beelden uit een klassensituatie (Amerikaanse basisschool). Hierna hebben

de observatoren samen met twee thesisbegeleiders getoetst in hoeverre de interbeoordelaarsbetrouwbaarheid voldoende (percentage overeenkomst > 80%) was om zelfstandig te gaan observeren. De interbeoordelaarsbetrouwbaarheid kwam op één student na overeen met de thesisbegeleiders. De observatoren waren vijf masterstudenten orthopedagogiek van de universiteit Leiden.

De data werd verzameld op twee dagen per week gedurende twee weken. Het leergedrag van de leerlingen werd geobserveerd evenals de instructies van de leerkrachten. De positie van de observatoren in de klas werd van tevoren bepaald, de observatoren namen plaats daar waar het zicht op de klas voldoende was doch uit beeld om zodanig de les niet te verstoren. Bij aanvang van elk lesuur werd door de leerkracht een plattegrond van de klas gemaakt en werden alle leerlingen op de plattegrond genummerd. De observatoren noteerden de nummers van de betreffende leerling voor iedere observatiemoment of lesuur.

Nadat bovengenoemde voorbereidingen waren getroffen konden de observatoren beginnen met de observaties. Om de 10 seconde werd het geobserveerde gedrag in een desbetreffend vakje aangevinkt. Door middel van een mp3 speler werd het signaal afgegeven wanneer de 10 seconden waren verstreken. Elke leerling werd gedurende één minuut geobserveerd tijdens de les. Nadat alle leerlingen één keer waren geobserveerd startten de observatoren opnieuw bij de eerst geobserveerde leerling.

Elke leerling werd gedurende minimaal 3 minuten per lesuur geobserveerd. Per minuut kon een maximale score van 6 behaald worden en een minimum score van 0. De totale scores van de verschillende leergedragingen werden berekend door per minuut de scores per leergedrag op te tellen. De totale score van het geobserveerde gedrag (ruwe score) per minuut werd vervolgens in een SPSS bestand ingevoerd.

Om de interbeoordelaarsbetrouwbaarheid te bepalen werd in tweetallen geobserveerd. Elke observator werd gekoppeld aan een klas. Er werd gedurende lessen van 50 minuten geobserveerd, en tijdens ongeveer 7 lessuren per dag. Bij aanvang van de lessen gaven de mentoren uitleg aan de leerlingen. Indien leerkrachten vragen hadden, konden ze deze na de les stellen aan de observatoren. Vragen of opmerkingen met betrekking tot het onderzoek konden bij de thesisbegeleiders worden neergelegd. Bij een percentage overeenkomst van > 80% is er sprake van een hoge betrouwbaar tussen de scores van de observatoren.

## 2.5 Data-analyse

Dit onderzoek bestond uit twee onderdelen. In het eerste onderdeel werd gekeken naar de interbeoordelaarsbetrouwbaarheid en de betrouwbaarheid tussen scores van de twee onderzoeksdagen (betrouwbaarheid van de gehele steekproef). De betrouwbaarheid tussen de verschillende beoordelaren werd berekend aan de hand van het percentage overeenkomst (percentage agreement). Hiervoor is de volgende formule gebruikt:  $\text{percentage overeenkomst} = \frac{\text{overeenstemmende score}}{(\text{overeenstemmende score} + \text{verschil})} * 100$  (Briesch, et al., 2010). Er werd eerst per lesuur berekend wat het percentage overeenkomst per lesuur. Vervolgens werd het percentage overeenkomst tussen de observatoren per leergedrag berekend evenals het totale percentage overeenkomst.

Na de dataselectie werden de beschrijvende statistieken berekend en de normaliteit en de verdeling van de data vastgesteld. De data werd aan de hand van de Spearman's Correlatietoets (non-parametrische) geanalyseerd omdat de data niet normaal verdeeld was.

Vervolgens werd gekeken naar de verschillen in het geobserveerd gedrag per individu. Om de mate van verschil te bepalen werden de leergedragingen "aan taak" en "actief leergedrag" van zowel onderzoeksdag 1 als onderzoeksdag 2 samengevoegd (nieuwe variabele "aan taak") en omgezet in percentages.

Vervolgens werd het verschil tussen de percentages berekend. Na het berekenen van het verschil werd deze nieuwe variabele onderverdeeld in drie categorieën: "weinig verschil" (10% of minder verschil tussen dag 1 en dag 2), "matig verschil" (verschil tussen 11 en 20 %) en "veel verschil" (21 % of meer verschil). De scores van de leerkracht waren teruggebracht van 4 naar 3 categorieën, omdat slechts één leerling tot de categorie slecht behoorde. 1 "zeer goed", 2 "goed" en 3 "minder goed". Daarna werd aan de hand van de beoordeling van de leerkracht bepaald of de mate van verschil per type leerlingen overeenkomstig was.

### 3. Resultaten

#### 3.1 Interbeoordelaarsbetrouwbaarheid per lesuur

De eerste deelvraag heeft betrekking op de interbeoordelaarsbetrouwbaarheid. Er is onderzocht of de scores van de verschillende observatoren met elkaar overeenkomen.

De interbeoordelaarsbetrouwbaarheid werd berekend door de overeenkomst per lesuur te delen door het totale aantal scores. In tabel 1 is de verdeling per lesuur en per observator weergegeven. Uit de resultaten bleek dat het percentage overeenkomst erg hoog is (88%). Het eerste observatorkoppel had een percentage overeenkomst van 81% en het tweede koppel 98%. Een uitzondering hierop was het percentage overeenkomst van de wiskundeles, deze bedroeg 74%.

Tabel 1

*Ruwe scores overeenkomst, verschil en percentage overeenkomst scores tussen observatoren*

| <b>Lesuur/</b>    | <b>Vak</b> | <b>overeenkomst</b> | <b>verschil</b> | <b>%</b> |
|-------------------|------------|---------------------|-----------------|----------|
| Observator 1 en 2 |            |                     |                 |          |
| 1/                | Nederlands | 102                 | 12              | 89%      |
| 2/                | Engels     | 108                 | 12              | 91%      |
| 3/                | Wiskunde   | 150                 | 54              | 74%      |
| 4/                | Nederlands | 96                  | 22              | 81%      |
| Totaal            |            | 456                 | 100             | 82%      |
| Observator 1 en 3 |            |                     |                 |          |
| 5/                | Nederlands | 188                 | 4               | 98%      |
| 6/                | LIS        | 152                 | 16              | 91%      |
| 7/                | LIS        | 159                 | 14              | 92%      |
| Totaal            |            | 499                 | 34              | 93%      |

### 3.2 Interbeoordelaarsbetrouwbaarheid per leergedrag

In het vorige onderdeel werd gekeken naar de interbeoordelaarsbetrouwbaarheid per lesuur. In dit onderdeel is de interbeoordelaarsbetrouwbaarheid per leergedrag berekend.

In tabel 2 wordt de betrouwbaarheid per leergedrag verdeeld over de verschillende lesuren en het totale gemiddelde per leergedrag weergegeven. Het leergedrag actief leergedrag bleek het minst betrouwbaar (43%). Het “aan taak” leergedrag was het meest betrouwbare (99%), gevolgd door “niet aan taak” (87%). Het totale percentage van het leergedrag “storend gedrag” was redelijk betrouwbaar (71%). In tabel 2 is te zien dat het overeenkomst van “aan taak” leergedrag beduidend hoger ligt (tussen 91% en 100%) dan de overige leergedragingen. En er was sprake van weinig verschil tussen de lesuren. Het “actief leergedrag” was het minst betrouwbare, de overeenkomst lag tussen 0% en 100%. Dit gold ook voor het “storende gedrag”. De overeenkomst van het “niet aan taak”leergedrag tijdens lesuur 2 was nogal laag (40%), maar was ook heel weinig gescoord (namelijk 4 keer door observator 1 en 10 keer door observator 2). Hierdoor dienen deze resultaten met enige voorzichtigheid geïnterpreteerd te worden. Gezien de verdeling en het overeenkomst percentage van de leergedragingen kan gesteld worden dat de interbeoordelaarsbetrouwbaarheid wisselend betrouwbaar is.

Tabel 2

*Percentage overeenkomst tussen observatoren per lesuur en leergedrag*

| <b>Lesuur</b> | <b>1</b> | <b>2</b> | <b>3</b> | <b>4</b> | <b>5</b> | <b>6</b> | <b>7</b> | <b>%</b> |
|---------------|----------|----------|----------|----------|----------|----------|----------|----------|
| leergedrag    |          |          |          |          |          |          |          |          |
| ALG           | 89%      | 0%       | 0%       | 0%       | 100%     | 0%       | 0%       | 43%      |
| SG            | 100%     | 100%     | 0%       | 100%     | 66%      | 0%       | 20%      | 71%      |
| NAT           | 87%      | 40%      | 76%      | 0%       | 98%      | 92%      | 81%      | 87%      |
| AT            | 93%      | 96%      | 100%     | 91%      | 99%      | 95%      | 98%      | 99%      |



### *3.3 Samenvoegen leergedragingen*

De leergedragingen actief leergedrag en storend gedrag bleken minder betrouwbaar te zijn. Gezien de lage betrouwbaarheid van deze twee leergedragingen en het feit dat er minder vaak op deze leergedragingen werd gescoord, is er besloten om deze twee samen te voegen ( actief leergedrag met “aan taak” en “storend gedrag” met “niet aan taak”). Het samenvoegen van “actief leergedrag” en “aan taak” had geen negatieve invloed op de betrouwbaarheid. Wanneer deze leergedragingen werden samengevoegd bleef de betrouwbaarheid nog hoog (96%). Dit gold ook voor de leergedragingen “storend gedrag” en “niet aan taak”. Bij het samenvoegen steeg het overeenkomst percentage van het “niet aan taak” leergedrag van 87% naar 89%.

### *3.4 Test-hertestbetrouwbaarheid van de scores van onderzoeksdag 1 en 2.*

In dit onderdeel wordt onderzocht of de scores van onderzoeksdag 1 en onderzoeksdag 2 met elkaar samenhangen (test-hertestbetrouwbaarheid). Ten eerste werd de normaliteit van de data vastgesteld. Vervolgens werd aan de hand van een Spearman’s Correlatietoets gekeken of er een samenhang is tussen de scores van de twee onderzoeksdagen.

In tabel 3 zijn de beschrijvende statistieken van de verschillende leergedragingen van onderzoeksdag 1 en onderzoeksdag 2 weergegeven. Het “aan taak” leergedrag was op onderzoeksdag 1 en 2 het meest voorkomend (59%). Het minst voorkomende leergedraging op onderzoeksdag 1 was “storend gedrag” (5%). Op onderzoeksdag 2 was “actief leergedrag” het minst voorkomend (2%). Gezien de verdeling van de percentages kan gesteld worden dat er weinig verschil is tussen de scores van onderzoeksdag 1 en 2.

De vuistregel van de normaal verdeling is dat data bij benadering normaal verdeeld is wanneer de waarden van de scheefheid en gepiekttheid tussen +2 en -2 vallen (P. Kroonenberg, persoonlijke communicatie, 1 november 2009). In tabel 3 is te zien dat de variabelen “aan taak” en “niet aan taak” van onderzoeksdag 1 normaal verdeeld zijn. Op dag 2 was de variabele “niet aan taak” als enige normaal verdeeld. Dit betekent dat de scores op deze leergedragingen allen rond het gemiddelde lagen. Er was geen sprake van extreme waardes. De overige variabelen waren niet normaal verdeeld. De leergedragingen “actief leergedrag” en “storend gedrag” werden minder vaak gescoord waardoor een of meerdere afwijkende scores een groter invloed hadden op de verdeling. Opvallend was dat er bij de leergedraging “actief leergedrag” op dag 2 er sprake was van een hoge gepiekttheid (23.39), een groot aantal scores waren veel lager dan het gemiddelde. Op basis van de beschrijvende statistieken kan geconstateerd worden dat de data deels normaal verdeeld was. De data zal aan de hand van de Spearman’s Correlatietoets geanalyseerd worden.

Tabel 3

*Beschrijvende gegevens van de verdeling van leergedragingen op onderzoeksdag 1 en 2*

|                   | <i>N</i> | <i>M</i> | <i>SD</i> | %   | Scheefheid | Gepiektheid |
|-------------------|----------|----------|-----------|-----|------------|-------------|
| Onderzoeksdag 1   |          |          |           |     |            |             |
| Actief leergedrag | 30       | 2.76     | 4.34      | 8%  | 1.90       | 3.27        |
| Aan taak          | 30       | 20.97    | 6.88      | 59% | .47        | -.60        |
| Niet aan taak     | 30       | 10.07    | 5.35      | 28% | .02        | -.31        |
| Storend gedrag    | 30       | 1.67     | 3.09      | 5%  | 2.65       | 8.18        |
| Onderzoeksdag 2   |          |          |           |     |            |             |
| Actief leergedrag | 30       | .73      | 2.05      | 2%  | 2.65       | 23.39       |
| Aan taak          | 30       | 21.23    | 9.14      | 59% | 4.63       | -1.23       |
| Niet aan taak     | 30       | 2.17     | 8.15      | 34% | -.15       | -.91        |
| Storend gedrag    | 30       | 1.63     | 2.54      | 5%  | 2.24       | 5.70        |

Om te onderzoeken of er sprake is van een overeenkomst tussen de meting van het eerste onderzoeksmoment en het tweede onderzoeksmoment is er een Spearman's Correlatietoets uitgevoerd.

Op basis van de resultaten van de correlatieanalyse is te constateren dat er sprake is van een redelijke samenhang tussen de scores van de eerste en de tweede onderzoeksdag; "aan taak"  $r(28) = .49$ ;  $p < .01$ , niet "aan taak"  $r(28) = .55$ ;  $p < .01$ , en storend gedrag  $r(28) = .47$ ;  $p < .01$ . Dit gold niet voor de scores van actief leergedrag  $r(28) = .20$ ;  $p < .01$ . Hier bleek dat er weinig samenhang was tussen de scores van de eerste en tweede onderzoeksdag. Kijkend naar het feit dat de leergedragingen "actief leergedrag" en "storend gedrag" niet normaal verdeeld waren dienen deze resultaten met enige voorzichtigheid geïnterpreteerd worden.

Doordat de gemiddelden van actief leergedrag en storend gedrag erg laag waren ten opzichte van de leergedragingen "aan taak" en "niet aan taak" werd besloten om "actief leergedrag" samen te voegen met de leergedraging "aan taak". Het samenvoegen van deze leergedragingen zorgde voor een stijging van de samenhang  $r(28) = .74$ ;  $p < .01$ .

### *3.5 Overeenkomst leergedrag per individuele leerling*

In het eerste deel van dit onderzoek werd het percentage overeenkomst tussen de scores van de observatoren berekend (interbeoordelaarsbetrouwbaarheid). Vervolgens is onderzocht of er sprake was van een samenhang tussen de scores van de onderzoeksdagen (test-hertestbetrouwbaarheid). Om beter zicht te krijgen op de betrouwbaarheid van het leergedrag, wordt in dit onderdeel per individuele leerling gekeken of het leergedrag van de twee onderzoeksdagen met elkaar overeenkomen (consistentie). Vervolgens wordt middels de uitkomsten van het lerarenbeoordelingsformulier gekeken of de mate van consistentie per type leerling verschilt. De verwachting is dat het “aan taak” leergedrag voor alle type leerlingen consistent zal zijn.

Om de mate van consistentie te bepalen werd het verschil tussen het percentage “aan taak”gedrag op onderzoekdag 1 en onderzoekdag 2 per leerling berekend, ( zie tabel 4). Bij 47% van de leerlingen was er sprake van een consistent beeld. Ze lieten minder dan 10% verschil zien tussen het “aan taak” leergedrag van dag 1 en dag 2. Bij deze groep leerlingen was er sprake van een wisselende beoordeling van het leergedrag. Van de 14 leerlingen die een consistent leergedragspatroon lieten zien, werd het leergedrag van 3 leerlingen met zeer goed beoordeeld door de mentor. Het leergedrag van 6 leerlingen werd volgens de mentor goed en van 5 leerlingen minder goed.

Bijna een derde deel (33%) van de leerlingen liet een redelijk consistent beeld zien. De scores van het “aan taak” leergedrag van de twee meetmomenten verschilde tussen de 11 en 20%. Van de 10 leerlingen waarbij er sprake was van een redelijk consistent patroon, werd het leergedrag van 2 leerlingen als zeer goed beoordeeld door de mentor. Het leergedrag van 5 leerlingen werd als goed ervaren en van 3 leerlingen als minder goed.

Bij 20% van de leerlingen was er sprake van aanzienlijk verschil (meer dan 21%) tussen de scores van de twee onderzoeksdagen. Het leergedrag van deze leerlingen was matig consistent. Van 20% van de leerlingen werd het leergedrag als zeer goed beoordeeld en van 47% als goed. De mentor beschouwde het leergedrag van 1 leerling als zeer goed, van 3 leerlingen als goed en van 2 leerlingen als minder goed. Aan de verwachting dat het leergedrag niet zal verschillen per type leerling wordt niet voldaan.

Tabel 4

*Weergave consistentie leergedrag van onderzoeksdag 1 en 2 en beoordeling van de leerkracht*

| <b>Verdeling leerlingen en mate van verschil tussen onderzoeksdag 1 en 2</b> |                 |                |               |           |
|--|-----------------|----------------|---------------|-----------|
| Beoordeling leergedrag   | Weinig Verschil | Matig Verschil | Veel Verschil | totaal    |
| Zeer Goed  | 3               | 2              | 1             | 6 (20%)   |
| Goed   | 6               | 5              | 3             | 14 (47%)  |
| Minder Goed  | 5               | 3              | 2             | 10 (33%)  |
| totaal   | 14 (47%)        | 10(33%)        | 6 (20%)       | 30 (100%) |

#### 4. Conclusie en discussie

De Student Behavior Observation Form (SBOF) is ontwikkeld om het leergedrag van leerlingen in het voortgezet onderwijs in kaart te brengen. In dit onderzoek is getracht om de betrouwbaarheid van de verschillende gedragingen van het SBOF vast te stellen. Om dit te kunnen bereiken werd de interbeoordelaarsbetrouwbaarheid en de test-hertestbetrouwbaarheid berekend. Tevens is er per leerling gekeken of de scores van het “aan taak” leergedrag per onderzoeksdag verschillend was en of de consistentie van het gedrag verschilde per type leerling.

De eerste onderzoeksvraag was gericht op de interbeoordelaarsbetrouwbaarheid en luidde; *Is er een overeenstemming tussen scores van de verschillende observatoren op de Student Behavior Observation Form?* De verwachting hierbij was dat er sprake is van een hoge interbeoordelaarsbetrouwbaarheid (percentage overeenkomst van minimaal 80%). Uit de resultaten bleek dat er een hoge mate van overeenstemming is tussen de scores van de drie observatoren. Het algemene percentage overeenkomst was 88%. Er was weinig verschil tussen de verschillende lessen, het percentage overeenkomst per lesuur was hoog (tussen 81 en 98%). Opvallend was dat het percentage overeenkomst van de wiskundeles lager was (74%). Dat dit percentage lager was dan bij de overige lessen is waarschijnlijk te wijten aan een verschil in interpretatie van het storend gedrag en het actief leergedrag.

Naast het berekenen van de interbeoordelaarsbetrouwbaarheid per lesuur, werd ook gekeken naar de interbeoordelaarsbetrouwbaarheid per leergedrag. Hieruit bleken de scores van de observatoren op de leergedragingen “aan taak” en “niet aan taak” overeen te komen. Het percentage overeenkomst van het

“aan taak” leergedrag was hoog (99%) evenals “niet aan taak” leergedrag (87%). Het percentage overeenkomst van het “storend gedrag” bleek wat lager (74%), maar was nog redelijk betrouwbaar. Het actief leergedrag bleek beduidend minder betrouwbaar en heeft een laag percentage overeenkomst (43%). Door de lage frequentie waarin deze laatste gedragingen gescoord werden, dienen de bevindingen met enige voorzichtigheid geïnterpreteerd te worden. Ondanks dat de interbeoordelaarsbetrouwbaarheid van ieder van de leergedragingen afzonderlijk niet even hoog waren, kan gesteld worden dat de interbeoordelaarsbetrouwbaarheid van de leergedragingen, “aan taak” en niet “aan taak” goed is. Deze bevinding komt overeen met de resultaten uit het onderzoek van Hintze et al. (2004) en Martin et al. (2010). Uit beide studies bleek de interbeoordelaarsbetrouwbaarheid erg hoog te zijn. In het onderzoek van Hintze en Matthews (2004) is onderzoek gedaan naar het “aan taak” en “niet aan taak” leergedrag en is het percentage overeenkomst van 55 observaties berekend. De observaties werden gedaan door 7 psychologie studenten die ruime ervaring hadden in het observeren. Voorafgaand aan het onderzoek volgden de observatoren een 4 uur durende training. Het percentage overkomst bedroeg 90%. De interbeoordelaarsbetrouwbaarheid van dit huidig onderzoek is goed, ondanks dat de observatoren geen ervaring hadden met het observeren en een kortere (3 uur) voorbereiding hebben gehad. In het onderzoek van Volpe, diPerna, Hintze en Shapiro (2005) werd het volgen van een uitgebreide training of opleiding aanbevolen. Echter bleek uit een soortgelijke onderzoek dat de mate van voorbereiding minimaal invloed heeft op de betrouwbaarheid van de data (LeBel, Kilgus, Briesch & Chafouleas, 2010). Het viel op dat er in het huidig onderzoek minder vaak gescoord werd op de leergedragingen “actief leergedrag” en “storend gedrag”. Dit is wellicht te wijten aan het feit dat de definities van de leergedragingen niet optimaal zijn geoperationaliseerd. Alle leergedragingen werden beschreven en werden een aantal praktische voorbeelden per definitie gegeven. Desondanks bleek tijdens de observaties nog twijfels te bestaan over de definities van met name actief leergedrag en storend gedrag. Riley-Tillman et al. (2009) stellen in hun onderzoek dat de definities zo nauwkeurig mogelijk omschreven moeten worden, opdat het gedrag op een consistente wijze gescoord kan worden. Volgens hen is het van belang om alle mogelijke voorbeelden en niet voorbeelden van de verschillende leergedragingen in de voorbereiding op te nemen. In het huidige onderzoek zijn tijdens de voorbereiding zijn maar enkele voorbeelden behandeld.

Naast de interbeoordelaarsbetrouwbaarheid is ook onderzoek gedaan naar de samenhang tussen de scores van de twee onderzoeksdagen. De tweede onderzoeksvraag luidde; *Is er een overeenkomst tussen de scores van onderzoeksdag 1 en onderzoeksdag 2?* De hypothese hierbij was dat er sprake was van een samenhang tussen de scores van de eerste en de tweede onderzoeksdag. Op basis van de resultaten van de correlatieanalyse kan gesteld worden dat er bij drie leergedragingen ( “aan taak”, “niet aan taak” en “storend gedrag” ) sprake was van een matig significante samenhang tussen de scores van de eerste en de

tweede onderzoeksdag (“aan taak”  $r(28) = .49; p < .01$ , “niet aan taak”  $r(28) = .55; p < .01$  en “storend gedrag”  $r(28) = .47; p < .01$ ). De scores van de leergedraging “actief leergedrag” van de twee onderzoeksdagen vertoonden minimale samenhang  $r(28) = .20; p < .01$ . Het samenvoegen van de variabelen “aan taak” en “actief leergedrag” zorgde voor een stijging van de samenhang,  $r(28) = .74; p < .01$ . Deze bevindingen zijn gedeeltelijk in overeenstemming met de resultaten uit eerder onderzoek. Uit eerder onderzoek blijkt dat verschillende facetten invloed kan hebben om de betrouwbaarheid van de data. In het onderzoek naar de Systematic Direct Observation werd 67% van de variantie verklaard door individuele verschillen in het gedrag van de participanten. De foutenvariantie was 33%. De metingen van dit onderzoek bleken minder betrouwbaar te zijn (betrouwbaar bij foutenvariantie van <30%). Het “aan taak” en “niet aan taak” leergedrag van de deelnemers verschilden op de twee onderzoeksmomenten (Hintze & Matthews, 2004). Een soortgelijk onderzoek naar het SDO liet echter een foutenvariantie van 30% zien, de metingen waren betrouwbaarder (Briesch, et al., 2010). Hieruit kan geconcludeerd worden dat er verschillende factoren zijn die de betrouwbaarheid van de data kunnen beïnvloeden. Zo bleken de leergedragingen het aan taak en niet aan taak leergedrag het meest betrouwbaar. In het onderzoek van Hintze en Matthews (2004) worden ook aan de hand van twee leergedragingen geobserveerd. Hierin bleek het leergedrag per observatiemoment, ochtend of middag te verschillen. Het aantal leergedragingen en het moment kan van invloed zijn geweest op de scores van het huidig onderzoek.

Verder is gekeken in hoeverre het “aan taak” leergedrag consistent was en of dit verschilde per type leerling. Om dit te bepalen was per leerling bepaald of er sprake is van weinig, matig of veel verschil. Daarna werd aan de hand van de beoordeling van het leergedrag door de mentor gekeken of de mate waarin leerlingen verschil in het leergedrag lieten zien, afhankelijk was van het type leerling. Hierbij was de derde deelvraag geformuleerd; *Is er sprake van consistentie tussen het “aan taak” leergedrag die tijdens de twee onderzoeksdagen per leerling is geobserveerd en verschilde mate van consistentie per type leerling.* De verwachting hierbij was dat het “aan taak” leergedrag van de twee onderzoeksdagen voor alle type leerlingen consistent zou zijn. De bevindingen waren niet in overeenstemming met de verwachting. Het geobserveerde “aan taak” leergedrag bleek niet altijd even consistent te zijn. Dit gold voor alle type leerlingen. Een verklaring hiervoor kan zijn dat er veel verschil was tussen de lessen van de twee onderzoeksdagen. Over het algemeen waren de leerlingen tijdens de eerste onderzoeksdag, vaker aan taak dan de tweede onderzoeksdag zie tabel 3. Tijdens de onderzoeksdagen viel het op dat de lessen op de eerste dag gestructureerder verliepen dan op de tweede onderzoeksdag. Er zijn een aantal factoren die mogelijk van invloed waren namelijk; onduidelijke instructie van de leerkracht (geen uitleg bij aanvang van de les m.b.t. samenwerking tussen leerlingen), onduidelijkheid met betrekking tot de definitie van de leergedragingen of de context waarin het gedrag wordt geplaast. Een voorbeeld hiervan

is het praten tijdens de les, het ene moment mogen de leerlingen praten dan zijn ze aan taak, het andere moment wordt hetzelfde gedrag als storend beschouwd.

De centrale onderzoeksvraag luidde; *Kunnen de gedragingen (“aan taak”, “niet aan taak”, “storend gedrag” en “actief leergedrag”) fungeren als betrouwbare indicatoren van een observatieformulier die de voortgang van het leergedrag in kaart moet brengen?* Ten eerste blijken de gedragingen, “aan taak” en niet “aan taak” het meest betrouwbaar te zijn. Er is sprake van een goede interbeoordelaarsbetrouwbaarheid (boven 80%). Ten tweede is test-hertestbetrouwbaarheid voldoende. De scores van de gedragingen “aan taak”, “niet aan taak” en “storend gedrag” op de verschillende onderzoeksmomenten hangen redelijk samen. Er is nog gekeken naar de mate van consistentie in het leergedrag van individuele leerlingen en of deze per type leerling verschilt. Het blijkt dat meer dan de helft van de leerlingen matig, tot veel verschil laten zien in het leergedrag. Dit verschilde niet per type leerling.

Op grond van de resultaten kan gesteld worden dat de leergedragingen “aan taak”, “niet aan taak” en “storend gedrag” betrouwbare indicatoren zijn om het leergedrag van leerlingen uit het voortgezet onderwijs in kaart te brengen.

#### *4.1 Beperkingen van het huidig onderzoek en aanbevelingen voor vervolgonderzoek*

Een aantal beperkingen aan dit onderzoek is noemenswaardig. De steekproefgrootte was relatief klein, van de oorspronkelijke steekproef (119) bleek de data van 30 leerlingen geschikt voor de analyses. Doordat de gegevens op één school zijn verzameld zijn de bevindingen niet generaliseerbaar en kan men door de homogeniteit van de doelgroep geen uitspraak doen over de hele populatie. In een vervolgonderzoek is het belangrijk dat de steekproef groter en heterogener is, dit in verband met de generaliseerbaarheid van de bevindingen.

Nog een beperking was dat nog niet alle psychometrische waarden van de gebruikte instrumenten (SBOF en LBF) zijn vastgesteld. Er is nog geen informatie beschikbaar over de validiteit. In dit huidig onderzoek is de betrouwbaarheid (interbeoordelaarsbetrouwbaarheid en de test-hertestbetrouwbaarheid) van de SBOF onderzocht.

In een breder onderzoek van de Universiteit Leiden zal de validiteit van de SBOF vastgesteld worden. Tevens is het van belang dat de andere vormen van betrouwbaarheid zoals de paralleltestbetrouwbaarheid worden vastgesteld. Een vergelijkbaar observatieinstrument kan gebruikt worden om een indicatie te geven van de betrouwbaarheid. Uit het onderzoek bleek dat de betrouwbaarheid van een aantal gedragingen (“aan taak”, “niet aan taak” en “storend gedrag”) goed was, desondanks is het van belang

om een gedegen onderzoek te doen naar de overige psychometrische waarden van de SBOF. Dit is vooral een aanbeveling als de SBOF binnen een probleem oplossingsmodel geïmplementeerd wordt en het ingezet wordt om de voortgang van leerlingen te monitoren. Aan de hand van de resultaten kan men een keuze maken voor bepaalde interventies of verwijzen naar speciaal onderwijs. Het over-of onderrapporteren van bepaald gedrag kan negatieve implicaties hebben voor het onderwijstraject van een leerling (Riley-Tillman, Christ, Chafouleas, Boice-Mallach & Briesch, 2011)

Tevens zijn andere aspecten van betrouwbaarheid en de validiteit die vastgesteld moeten worden. Er is naast het observeren gebruik gemaakt van het leerkracht beoordelingsformulier (LBF) om vast te stellen welke categorie leerlingen een consistent leergedrag patroon laat zien. In een vervolgonderzoek kan de validiteit en betrouwbaarheid van de LBF vastgesteld worden door o.a. te kijken of de beoordeling van de leerkracht overeenkomt met het geobserveerde leergedrag. Tevens kan er nader onderzocht worden waar mogelijke discrepanties liggen. Het is mogelijk dat de leerkracht het leergedrag van een leerling als minder goed beoordeelt terwijl de leerling wel vaak “aan taak” is. Dit kan berusten op systematische fout zoals bias van de leerkracht of van de observator. Als een leerling bijvoorbeeld minder goed presteert kan een leerkracht eerder geneigd zijn om het leergedrag negatief te beoordelen terwijl de leerling wel vaak “aan taak” is.

Daarnaast is het van belang dat de leergedragingen nader onderzocht worden. In dit huidige onderzoek is onderscheid gemaakt tussen vier verschillende leergedragingen. Op basis van de resultaten kan worden geconcludeerd dat het uitsluitend observeren van het “aan taak” en “niet aan taak” leergedrag essentiële informatie oplevert. Tevens blijkt dat deze leergedragingen het meest betrouwbaar zijn. Desalniettemin is het belangrijk verder te onderzoeken of het uitsluitend observeren van de “aan taak” of “niet aan taak” leergedragingen een objectief beeld van het algemeen leergedrag weergeeft. Een ander aspect dat in het huidig onderzoek niet aan de orde kwam is de duur en frequentie van de observaties en of deze invloed hebben op de betrouwbaarheid van de data. Gezien het doel van dit instrument (voortgang monitoren) is het van belang dat men onderzoek doet naar de duur en frequentie van de observaties. Men wil met minimaal aantal observaties een betrouwbaar beeld krijgen van het leergedrag en die vordering van de leerling. Ook wordt aanbevolen om in een vervolgonderzoek vast stellen op welke wijze de verkregen informatie geïnterpreteerd kan worden.



#### 4.2 *Implicaties voor de praktijk*

De volgende implicaties vloeien voort uit de bevindingen van dit onderzoek. Middels een observatieformulier zoals de SBOF kan men het leergedrag zoals dat zich in de klas voordoet en de vorderingen in kaart brengen. Om een compleet en betrouwbaar beeld te scheppen van het leergedrag is het belangrijk om andere aspecten van het leren te onderzoeken (bijvoorbeeld leerstrategieën) en na te gaan hoe deze aspecten het leergedrag beïnvloeden. Tevens zal men verschillende bronnen moeten raadplegen zoals ouders en leerlingen (Kievit, Tak, & Bosch, 2002).

De voorbereidingen voorafgaand aan de observaties (training, materiaal) en het observeren kunnen relatief tijdrovend zijn. Er zijn verschillende onderzoeken gedaan naar het effect van training (geen, kort of intensief) op de accuraatheid van de observaties. Hieruit bleek geen significant verschil in de nauwkeurigheid van het scoren van het leergedrag (LeBel, et al., 2010). Desalniettemin is enige vorm van voorbereiding van belang. Een gedegen voorbereiding kan fouten zoals niet correct scoren van de leergedragingen minimaliseren (Volpe, et al., 2005). Een duidelijk handleiding waarin de definities, instructies en procedures met betrekking tot dit instrument worden vastgelegd is van belang om de procedure te standaardiseren.

Om de resultaten van een observatieformulier zoals de SBOF in kaart te brengen is het van belang om een systeem te ontwikkelen waarin leerkrachten of pedagogen de scores op een eenvoudige wijze kunnen invoeren en de resultaten (leergedrag) inzichtelijk gemaakt kunnen worden voor alle betrokkenen (ouders en leerlingen). In de praktijk blijkt dat men in het onderwijs onvoldoende middelen (kennis, tijd en budget) heeft voor dure systemen en het inhuren van expertise (Evans & Owens, 2010).

In een onderzoek van Jacobs, et al. (2000) wordt aanbevolen om een stapsgewijze screening uit te voeren. Het blijkt effectiever om de observatieprocedure te faseren. Dit houdt dat in dat alleen leerlingen die binnen de risicogroep vallen meerdere malen per schooljaar worden geobserveerd.

## Literatuurlijst

- Briesch, A. M., Chafouleas, S. M., & Riley-Tillman, T.C. (2010). Generalizability and dependability of behavior assessment methods to estimate academic engagement: a comparison of systematic direct observation and direct behavior rating school. *Psychology Review*, 39, 408-421.
- Chafouleas, S.M., Briesch, A.M., Chris Riley-Tillman T.C., Theodore, J.C., Black, A.C. & Kilgus, S.P. (2010a). An investigation of the generalizability and dependability of direct behavior ratings Item scales (DBR-SIS) to measure academic engagement and disruptive behavior of middle school students. *Journal of School Psychology*, 48, 219-246.
- Chafouleas, S. M., Volpe R. J., Gresham, F.M. & Cook, C. R. (2010b). School-based behavioral assessment within school problem-solving models: Current status and future directions *Psychology Review*, 39, 343-349.
- Christ, T.J., Riley-Tillman, T.C. , Chafouleas, S.M. & Jaffery, R. (2011). Direct behavior rating. An evaluation of alternate definitions to asses classroom behavior. *School Psychology Review*, 40, 181-199.
- Christ, T.J., Riley-Tillman, T.C., Chafouleas, S.M. & Bioce, C.H. (2010). Direct Behavior Rating (DBR) generalizability and dependability across raters and observations. *Educational and Psychological Measurement*, 70, 825-843.
- Espin, C. A. & Yell, M. L. (1994). Critical indicators of effective teaching for preservice teachers: Relations between teaching behavior and ratings of effectiveness. *Teacher Education and Special Education*, 17, 154-169.
- Evans, S. W. & Owens, J. (2010). Behavioral assessment within problem-solving models: Finding relevance and expanding feasibility. *School Psychology Review*, 39, 427–430.
- Fredricks, J. A., Blumenfeld, P. C. & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, 74, 59-109.
- Gresham, F. M. (2005). Response to intervention: An alternative means of identifying students as emotionally disturbed. *Education and Treatment of Children*, 28, 328–344.
- Gresham, F. M., Cook, C. R., Collins, T., Dart, E., Rasetshwane, K., Truelson, E. & Grant, S. (2010). Developing a Change-Sensitive Brief Behavior Rating Scale as a Progress Monitoring Tool for Social Behavior An Example Using the Social Skills Rating System- Teacher Form. *School Psychology Review*, 39, 364–379.

- Hintze, J.M. & Matthews, W.J. (2004). The generalizability of systematic direct observations Across time and setting: A preliminary investigation the psychometrics of behavioral observation. *School Psychology Review*, 33, 258-270.
- Hughes, C.A. & Dexter, D.D. (2011). Response to intervention. A research based summary. *Theory into Practice*, 50, 4-11.
- Jacobs, J.R., Bogs, S. R., Eyberg, S.M., Edwards, D., Durning, P., Querido, J.G., Funderburk, B.W. & McNeil, C.B. (2000). Psychometric properties and reference point data for the revised edition of the school observation coding system. *Behavior Therapy*, 31, 695-712.
- Kaplan, A., Gheen, M. & Midgley, C. (2002). Classroom goal structure and student disruptive behaviour British. *Journal of Educational Psychology* 72, 191–211.
- Kievit, Th., Tak, J. A & Bosch J. D.(2002). *Handboek Psychodiagnostiek voor de hulpverlening aan kinderen*. (6<sup>e</sup> herz. dr.) Utrecht: Tijdstroom
- Leary, M. (2004) *Introduction to behavior research method*. (4e dr.). Washington D.C.: Pearson.
- LeBel, T., Kilgus, S.P., Briesch, A.M., Chafouleas, S. (2010). The impact of training on the teacher-completed direct behavior ratings (DBRs). *Journal of positive behavior interventions*, 12, 55-63.
- Martin, P. A., Daley, D., Hutchings, J., Jones, K., Eames, C. & Whitaker, C. J. (2010). The Teacher–Pupil Observation Tool (T-POT): Development and testing of a new classroom observation measure. *School Psychology International* , 31, 229-249.
- McIntosh, K., Frank, J.L., Spaulding, S.A. (2010). Establishing research based trajectories of office discipline referrals for individual student. *School Psychology Review*, 39, 380-394.
- NIP.(2010). *Cotan beoordelingssysteem voor de kwaliteit van tests*. Geraadpleegd op 5 maart 2011 via <http://www.psynip.nl/website/wat-doet-hetnip/tests/beoordelingsprocedure/beoordelingsprocedure>
- Nock, M.K. & Kurtz, S.M.S. (2005) Direct behavioral observation in school settings: Bringing science to practice. *Cognitive and Behavioral Practice*, 12, 359-370.
- Olympia, D.E, Heathfield, L.T., Jenson, W.R. & Clark, E. (2002). Multifaceted functional behavior assessment for students with externalizing behavior disorders. *Psychology in the Schools*, 39, 139-155.

- Onderwijsraad. (2010). *De school en leerlingen met gedragsproblemen*. Geraadpleegd op februari 10, 2011 via <http://www.onderwijsraad.nl/upload/publicaties/569/documenten/school-en-leerlingen-met-gedragsproblemen.pdf>
- Pakarinen, E., Lerkkanen, M., Poikkeus, A., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H. & Nurmi, J. (2010). A Validation of the classroom assessment scoring system in Finnish kindergartens. *Early Education and Development, 21*, 95-124.
- Riley-Tillman, T. C., Chafouleas, S.M., Christ, T. J., Briesch, A.M., & LeBel, T. J. (2009). The impact of wording and behavioral specificity on the accuracy of Direct Behavior Ratings (DBRs). *School Psychology Quarterly, 24*, 1-12.
- Riley-Tillman, T. C., Christ, Th., Chafouleas, S. M., Boice-Mallach C.H., & Amy Briesch. (2011). The Impact of observation duration on the accuracy of data obtained from Direct Behavior Rating (DBR). *Journal of Positive Behavior Interventions, 3*, 119-128.
- Traag, T., R. van der Velden & Rolf, K.W. (2011). Early school-leaving in the Netherlands: The role of family resources, school composition and background characteristics in early school leaving in lower secondary education. *Irish Education Studies, 30*, 45-62.
- Volpe, R., DiPerna, J.C., Hintze, J.M., Shapiro, E. S. (2005). Observing students in classroom settings: A review of seven coding schemes. *School Psychology Review, 34*, 454-474.
- Volpe, R.J., McConaughy, S.H., Hintze, J.M. (2009). Generalizability of classroom behavior problem and on-task scores from the Direct Observation Form. *School Psychology Review, 38*, 382-401.
- Walker, H. M., Ramsey, E., & Gresham, F. M. (2003). Heading off disruptive behavior: How early intervention can reduce defiant behavior-and win back teaching time. *American Educator, winter*, 6-21.
- Wang, M., & Holcombe, R. (2010). Adolescents' perceptions of school environment, engagement and academic achievement in middle school. *American Educational Research Journal, 47*, 633-662.

**Bijlagen I**

**LGG observatieformulier**

Datum: ..... Dag: ..... Tijd: .....

Leraar: ..... Les: .....

Leerling

Leerkracht

|         |
|---------|
| LL Code |
|         |

|      |    |    |    |    |    |    |
|------|----|----|----|----|----|----|
| sec: | 10 | 20 | 30 | 40 | 50 | 60 |
| ALG  |    |    |    |    |    |    |
| SG   |    |    |    |    |    |    |
| NAT  |    |    |    |    |    |    |
| AT   |    |    |    |    |    |    |

|    |  |
|----|--|
| PO |  |
| NO |  |
| CO |  |

|         |
|---------|
| LL Code |
|         |

|      |    |    |    |    |    |    |
|------|----|----|----|----|----|----|
| sec: | 10 | 20 | 30 | 40 | 50 | 60 |
| ALG  |    |    |    |    |    |    |
| SG   |    |    |    |    |    |    |
| NAT  |    |    |    |    |    |    |
| AT   |    |    |    |    |    |    |

|    |  |
|----|--|
| PO |  |
| NO |  |
| CO |  |

|         |
|---------|
| LL Code |
|         |

|      |    |    |    |    |    |    |
|------|----|----|----|----|----|----|
| sec: | 10 | 20 | 30 | 40 | 50 | 60 |
| ALG  |    |    |    |    |    |    |
| SG   |    |    |    |    |    |    |
| NAT  |    |    |    |    |    |    |
| AT   |    |    |    |    |    |    |

|    |  |
|----|--|
| PO |  |
| NO |  |
| CO |  |

|         |
|---------|
| LL Code |
|         |

|      |    |    |    |    |    |    |
|------|----|----|----|----|----|----|
| sec: | 10 | 20 | 30 | 40 | 50 | 60 |
| ALG  |    |    |    |    |    |    |
| SG   |    |    |    |    |    |    |
| NAT  |    |    |    |    |    |    |
| AT   |    |    |    |    |    |    |

|    |  |
|----|--|
| PO |  |
| NO |  |
| CO |  |

|         |
|---------|
| LL Code |
|         |

|      |    |    |    |    |    |    |
|------|----|----|----|----|----|----|
| sec: | 10 | 20 | 30 | 40 | 50 | 60 |
| ALG  |    |    |    |    |    |    |
| SG   |    |    |    |    |    |    |
| NAT  |    |    |    |    |    |    |
| AT   |    |    |    |    |    |    |

|    |  |
|----|--|
| PO |  |
| NO |  |
| CO |  |

|         |
|---------|
| LL Code |
|         |

|      |    |    |    |    |    |    |
|------|----|----|----|----|----|----|
| sec: | 10 | 20 | 30 | 40 | 50 | 60 |
| ALG  |    |    |    |    |    |    |
| SG   |    |    |    |    |    |    |
| NAT  |    |    |    |    |    |    |
| AT   |    |    |    |    |    |    |

|    |  |
|----|--|
| PO |  |
| NO |  |
| CO |  |

## Bijlagen II: leerkracht beoordelingsformulier

### 'Hoe beoordeelt u het leergedrag van de leerling?'

*Onder leergedrag wordt het gedrag verstaan wat verbonden is met en invloed heeft op het leren, zoals het luisteren naar instructie en het zelfstandig kunnen werken.*

1= zeer goed

2= goed

3= minder goed

4= slecht

| Naam leerling: | Score leergedrag: | 1 | 2 | 3 | 4 |
|----------------|-------------------|---|---|---|---|
| 1.             |                   | 0 | 0 | 0 | 0 |
| 2.             |                   | 0 | 0 | 0 | 0 |
| 3.             |                   | 0 | 0 | 0 | 0 |
| 4.             |                   | 0 | 0 | 0 | 0 |
| 5.             |                   | 0 | 0 | 0 | 0 |
| 6.             |                   | 0 | 0 | 0 | 0 |
| 7.             |                   | 0 | 0 | 0 | 0 |
| 8.             |                   | 0 | 0 | 0 | 0 |
| 9.             |                   | 0 | 0 | 0 | 0 |
| 10.            |                   | 0 | 0 | 0 | 0 |
| 11.            |                   | 0 | 0 | 0 | 0 |
| 12.            |                   | 0 | 0 | 0 | 0 |
| 13.            |                   | 0 | 0 | 0 | 0 |
| 14.            |                   | 0 | 0 | 0 | 0 |
| 15.            |                   | 0 | 0 | 0 | 0 |
| 16.            |                   | 0 | 0 | 0 | 0 |
| 17.            |                   | 0 | 0 | 0 | 0 |
| 18.            |                   | 0 | 0 | 0 | 0 |
| 19.            |                   | 0 | 0 | 0 | 0 |
| 20.            |                   | 0 | 0 | 0 | 0 |
| 21.            |                   | 0 | 0 | 0 | 0 |
| 22.            |                   | 0 | 0 | 0 | 0 |
| 23.            |                   | 0 | 0 | 0 | 0 |
| 24.            |                   | 0 | 0 | 0 | 0 |