



Universiteit Leiden

Leergedrag in kaart

Een exploratief onderzoek naar de voorspellers van adequaat leergedrag van
leerlingen binnen het voortgezet onderwijs

Margje Kooiker

Universiteit Leiden

Eerste beoordelaar: Prof. Dr. Christine A. Espin

Tweede beoordelaar: Dr. Marian J.A.J. Verhallen

Oktober 2011

Samenvatting

Leerprestaties van leerlingen en hun gedrag in de klas beïnvloeden elkaar. Leerlingen die betrokken zijn bij de les, presteren beter. Om prestaties van leerlingen te verbeteren, zouden scholen moeten inzetten op het optimaliseren van adequaat leergedrag van leerlingen. Hoewel diverse instrumenten leergedrag van leerlingen in kaart kunnen brengen, blijken deze instrumenten niet sensitief genoeg om, ook kleine, ontwikkelingen in leergedrag te meten. Om de progressie van leerlingen te kunnen beoordelen, is het belangrijk dat er een leergedragvolgsysteem komt waarmee scholen het leergedrag van leerlingen herhaaldelijk kunnen meten om zo veranderingen in leergedrag te kunnen vaststellen. In dit onderzoek is gezocht naar gedragsindicatoren die in staat zijn goed leergedrag te voorspellen. Vier tweedejaars klassen uit het voortgezet onderwijs participeerden in het onderzoek ($N = 44$). Er zijn vier gedragsindicatoren geobserveerd: actief leergedrag, aan taak, niet aan taak en storend gedrag. De mentor beoordeelde het leergedrag van leerlingen als zeer goed, goed of minder goed. Door de observaties in de klas te vergelijken met de beoordeling van de mentor, bleek de mate waarin een leerling actief leergedrag liet zien, verschilde tussen deze drie groepen. De groep die volgens de mentor zeer goed leergedrag liet zien, was ook meer aan taak dan leerlingen die volgens de mentor goed of minder goed leergedrag lieten zien. Er bleek geen verschil te zijn in de hoeveelheid tijd die de leerlingen met goed of minder goed leergedrag bezig waren met hun taak.

Inleiding

Leerprestaties van kinderen op de basisschool worden beïnvloed door hun gedrag in de klas (Algozzine, Wang, & Violette, 2011). Hoe minder storend gedrag en hoe groter de actieve of passieve betrokkenheid bij de activiteiten die de leraar initieert, des te beter de leerprestaties blijken te zijn (Chafouleas et al., 2010a; Chafouleas, Volpe, Gresham, & Cook, 2010b; Fredricks, Blumenfeld, & Paris, 2004; Morgan, Farkas, Tufis, & Sperling, 2008). Leerlingen die meer betrokken zijn tijdens de les en daarmee adequaat leergedrag laten zien, zijn meer overtuigd van de waarde van het leren en voelen zich competent (Roeser, Van der Wolf, & Strobel, 2001). Deze leerlingen tonen meer zelfvertrouwen en weten beter wat ze willen bereiken in het leven. Leerlingen die minder betrokken zijn bij de lessen laten meer storend gedrag zien. Daardoor hebben ze minder aandacht voor de instructies van de leerkracht (Morgan et al., 2008) en houden minder tijd over om te besteden aan lesactiviteiten, zoals het maken van opdrachten (Johnson, McGue, & Iacono, 2005; Miles & Stipek, 2006). De kans is vervolgens groot dat ze de leerstof niet goed begrijpen en een steeds grotere leerachterstand oplopen. Hierdoor zullen ze taken al snel als te moeilijk ervaren en is het waarschijnlijk dat ze nog meer probleemgedrag gaan vertonen om daarmee te ontkomen aan deze taken (McIntosh, 2005). Er lijkt dan sprake van een neergaande spiraal waar de leerling moeilijk zelf uit kan komen. Leerlingen met problematisch gedrag in de klas presteren minder goed en hebben een verhoogd risico om voortijdig de school te verlaten (Fredricks et al., 2004; Furlong & Christenson, 2008; Markussen, Froseth, & Sandberg, 2011; Roeser et al., 2001; Tobin & Sugai, 1999). Er zijn veel onderzoeken die spreken van de relatie tussen problematisch gedrag in de klas en leerprestaties, maar de causale relatie is onduidelijk en lijkt bidirectioneel. Leerlingen die op jonge leeftijd problematisch gedrag vertonen, lopen meer risico op het ontwikkelen van problemen met leren (e.g. Fleming, Harachi, Cortes, Abbott, & Catalano, 2004; Morgan et al., 2008) en leerlingen die in de eerste klassen van het primair onderwijs problemen hebben met leren, lopen meer risico op het ontwikkelen van probleemgedrag (e.g. McIntosh, Horner, Chard, Boland, & Good, 2006; Morgan et al., 2008)

Naast het verlies aan effectieve leertijd voor de leerling zelf, zal het storende gedrag ook gevolgen hebben voor de instructietijd van medeleerlingen. Leraren geven inderdaad aan dat ze veel tijd verliezen door storend gedrag van leerlingen. Uit een onderzoek onder leraren blijkt dat een vijfde deel van de leraren zegt twee á drie uur per week te verliezen door storend gedrag van leerlingen, een zesde deel zegt zelfs vier uren of meer per week kwijt te zijn door problematisch gedrag van hun leerlingen (Walker, Ramsey & Gresham, 2003). Hieruit blijkt

dat de aanwezigheid van probleemgedrag van leerlingen in direct verband staat met de hoeveelheid effectieve instructietijd die een leraar heeft voor zijn klas. Het verlies aan instructietijd heeft daarmee negatieve consequenties voor leerlingen die in klassen zitten waar leerlingen storend gedrag laten zien. Omdat effectieve instructietijd de sleutel is voor betere prestaties van leerlingen (Douglas, 2009), is het belangrijk dat er zo weinig mogelijk tijd verloren gaat aan storend gedrag.

Problematisch gedrag in de klas kan voor de leerling verstrekende gevolgen hebben (McIntosh, 2005, 2008). Wanneer een leerling zonder startkwalificatie de school verlaat, een zogenaamde *drop-out*, heeft hij door zijn lager scholingsniveau minder kansen op een goede baan en is er een vergrote kans op werkloosheid (Markussen et al., 2011). Tevens blijken deze jongeren 2,5 keer vaker verdacht te worden van het plegen van een misdrijf (Blom, Oudhof, Bijl, & Bakker, 2005). Niet adequaat leergedrag heeft hiermee negatieve consequenties voor de leerling zelf, zijn medeleerlingen en voor de maatschappij.

Omdat er in de maatschappij behoefte is aan goed opgeleide mensen (Fredricks et al., 2004; Ministerie van Onderwijs, Cultuur en Wetenschap, 2011a), staat het terugdringen van voortijdig schoolverlaten dan ook expliciet op de agenda van het ministerie van Onderwijs, Cultuur en Wetenschap (Ministerie van Onderwijs, Cultuur en Wetenschap, 2011a). Het verhogen van onderwijsrendement in zijn algemeenheid is momenteel een prioriteit van de regering. Nadat in 2007 initiatieven werden ontplooid om de leeropbrengsten in het basisonderwijs te verhogen, het zogenaamde *Opbrengstgericht werken*, worden er nu maatregelen genomen deze manier van werken ook in het voortgezet onderwijs in te voeren (Visscher & Ehren, 2011). Scholen zullen worden verplicht voor een aantal basisvakken leerlingvolgsystemen in te voeren waarmee de ontwikkeling van de leerlingen bijgehouden kan worden en zullen zich aantoonbaar in moeten zetten de resultaten van de leerlingen te verbeteren (Ministerie van Onderwijs, Cultuur en Wetenschap, 2011b, 2011c). Omdat het leren binnen de school onlosmakelijk verbonden is met het leergedrag van leerlingen (Miles & Stipek, 2006), zullen scholen zich moeten richten op het voorkomen of behandelen van inadequaat leergedrag. Het stimuleren van adequaat leergedrag helpt scholen om hun belangrijkste doel – leerlingen helpen schoolse vaardigheden te ontwikkelen – te bereiken (Miles & Stipek, 2006). Volgens Walker et al. (2003) kunnen scholen veel doen om storend gedrag van leerlingen in de les te verminderen. Dit moet leerlingen helpen hun leerprestaties te verbeteren (Morgan et al., 2008; Walker et al. 2003). Door vroegtijdig te interveniëren neemt het de risico op het ontstaan van grotere problemen op gebied van leren of gedrag af

(e.g. Chafouleas et al., 2010a, 2010b; Morgan et al., 2008). Dit vraagt om instrumenten waarmee effecten van die interventies kunnen worden gemeten. Om de progressie van leerlingen te kunnen beoordelen, is het belangrijk dat er een leergedragvolgsysteem komt waarmee scholen het leergedrag van leerlingen herhaaldelijk kunnen meten om zo veranderingen in leergedrag te kunnen vaststellen (Shapiro, 2004).

Instrument om Leergedrag te Meten

Er zijn verschillende betrouwbare en valide instrumenten beschikbaar om probleemgedrag van leerlingen in kaart te brengen (Briesch, Chafouleas, & Riley-Tillman, 2010), maar er is nog geen instrument waarmee leergedrag van leerlingen gevolgd kan worden. Dit instrument zou sensitief moeten zijn voor kleine veranderingen en de mogelijkheid bieden om met korte tussenpozen gebruikt te worden om de effecten van interventies te kunnen bepalen (Chafouleas et al., 2010a; Espin & Yell, 1994; Espin et al., 2008; Evans & Owens, 2010; Gresham et al., 2010). Voor dat laatste is het daarom van belang dat het instrument met weinig inspanning en tijdsinvestering gebruikt kan worden. (Espin & Yell, 1994; Espin et al., 2008; Evans & Owens, 2010).

Op cognitief gebied is het belangrijk dat de prestaties van een leerling vergeleken kunnen worden met eerdere prestaties en met die van zijn medeleerlingen (Deno, 2005), maar dit lijkt ook belangrijk op het gebied van leergedrag. Voor het volgen van de cognitieve ontwikkeling van leerlingen, wordt in Amerika vaak gebruikgemaakt van *curriculum based measurement* (Deno, 2005; Espin & Yell, 1994). De procedures bij *curriculum based measurement* zijn gestandaardiseerd en er is gespecificeerd wat en hoe er gemeten wordt, hoe er gescoord wordt en hoe de gegevens geïnterpreteerd moeten worden. *Curriculum based measurement* moet aan een aantal kenmerken voldoen: 1. De toets moet betrouwbaar en valide zijn; 2. De toets moet eenvoudig af te nemen en efficiënt zijn, zodat leraren deze regelmatig af kunnen nemen om ontwikkeling bij de leerlingen te meten; 3. De data moeten gemakkelijk te interpreteren zijn, zodat de resultaten helder gecommuniceerd kunnen worden; 4. De toets moet goedkoop zijn, omdat het herhaaldelijk ingezet moet kunnen worden (Deno, 1985). De scores worden bij *curriculum based measurement* verwerkt tot een grafiek die de leraren gebruiken om de resultaten te interpreteren (Espin et al., 2008).

De informatie die scholen door middel van *curriculum based measurement* verzamelen over de voortgang van de leerling biedt hen de gelegenheid om op grond daarvan onderwijskundige beslissingen te nemen (Deno, 2003) en in gesprekken met de ouders van de desbetreffende kinderen de voortgang te bespreken (Deno, 2005). Onderzoek heeft

uitgewezen dat leerlingen van leraren die deze wijze van toetsen systematisch gebruiken beter presteren. Daarnaast blijkt dat leraren die *curriculum based measurement* gebruiken beter weten hoe hun leerlingen zich ontwikkelen en dat deze leraren regelmatig hun instructie verbeteren wanneer de leerlingen niet voldoende vooruitgaan (Fuchs, Fuchs, & Hamlett, 1993; Deno, 2003; Deno, 2005).

Leergedrag kan op een indirecte wijze in kaart gebracht worden door middel van beoordelingsschalen. Deze worden vaak als valide en betrouwbaar, efficiënt en economisch gezien (Shapiro & Clemens, 2005; Ramsay, Reynolds, & Kamphaus, 2002). Voor het invullen van een beoordelingsschaal is geen intensieve training vereist (Nock & Kurtz, 2005). De beschikbare vragenlijsten worden vaak ingevuld voor gedrag dat plaatsvindt over een langere periode, zijn niet sensitief voor veranderingen en dit maakt ze minder geschikt om leergedrag te volgen (Chafouleas, Riley-Tillman, & Christ, 2009; Riley-Tillman et al., 2005).

Een korte gedragsbeoordelingsschaal lijkt een goed alternatief te zijn om het gedrag van leerlingen te volgen en het effect van interventies te meten. Deze gedragsbeoordelingsschaal dient direct na de les ingevuld te worden wat onder dezelfde omstandigheden een aantal keren herhaald moet worden om een betrouwbaar beeld te krijgen van het geobserveerde gedrag (Chafouleas et al., 2009; Chafouleas et al., 2010a; Gresham et al., 2010). Uit verschillende onderzoeken naar het *aan taak* en *storend gedrag* van leerlingen is gebleken dat er vaak een sterke correlatie is tussen door leraren ingevulde korte gedragsbeoordelingsschalen en directe observatie door externe observatoren (aan taak: 0,81; storend gedrag: 0,87). Dat de correlatie soms matig is, geeft aan dat deze beoordelingsschalen nog doorontwikkeld moeten worden (Chafouleas et al., 2005; Riley-Tillman, Chafouleas, Sassu, Chanese, & Glazer, 2008). De resultaten lijken tot nu toe veelbelovend (Chafouleas et al., 2010a; Gresham et al., 2010). Er zijn echter nog geen korte beoordelingsschalen die specifiek leergedrag in kaart brengen.

Een meer directe wijze om het leergedrag in kaart te brengen is directe observatie (Shapiro & Kratochwill, 2000) Deze gebruikte techniek om leergedrag in kaart te brengen, levert betrouwbare en nauwkeurige informatie op, omdat het gedrag gemeten wordt op het moment dat het plaatsvindt waarmee het een geschikte variant lijkt om leergedrag van een leerling te volgen (Nock & Kurtz, 2005; Riley-Tillman, Kalberer, & Chafouleas, 2005; Shapiro & Clemens, 2005). Vijf kenmerken van systematisch direct observeren zijn: 1. het meet specifiek gedrag, 2. de definitie van dit gedrag is van te voren geoperationaliseerd, 3. de procedure van de observatie is gestandaardiseerd, 4. de tijd en plaats van observatie zijn

nauwkeurig geselecteerd, 5. het verzamelen en interpreteren van de data is gestandaardiseerd en er is een hoge interbeoordelaarsbetrouwbaarheid (Hintze & Matthews, 2004).

Het observeren van kinderen is echter tijdsintensief en om op een betrouwbare wijze te observeren is een training noodzakelijk (Nock & Kurtz, 2005; Riley-Tillman et al., 2005). Bij bijvoorbeeld de Behavioral Observation of Students in Schools (BOSS) duurt een enkele observatie ongeveer 20 minuten. Dit betekent dat je per lesuur hooguit 2,5 leerling uit een klas zult kunnen observeren. Daarnaast is, om de BOSS als observatie-instrument te mogen gebruiken, een training van 10-15 uren vereist (Volpe, DiPerna, Hintze, & Shapiro, 2005). Dit maakt dit instrument tijdsintensief en als leergedragvolgsysteem minder geschikt. Met de BOSS worden vijf leerlinggedragingen geobserveerd. Het gaat om twee gedragsindicatoren die het *aan taak*-gedrag van de leerling meten (passief en actief) en drie *niet aan taak*-indicatoren (verbaal, motorisch, passief). Het is de vraag of voor het volgen van leergedrag van leerlingen al deze vijf gedragingen nodig zijn en of niet volstaan kan worden met één of twee gedragingen die het leergedrag van de leerling op een betrouwbare wijze kunnen voorspellen. Hiermee zou de complexiteit verminderd en de tijdsduur verkort kunnen worden.

Het gebruiken van directe observatie voor het volgen van leergedrag lijkt een betrouwbare en valide manier om leergedrag van leerlingen te meten. Voor een leraar is het echter moeilijk om tegelijkertijd les te geven en systematisch te observeren. Wanneer de observatie door een externe observator gedaan wordt, is het belangrijk dat dit gebeurt vanaf een onopvallende plaats in het lokaal zodat de aanwezigheid van de observator de interactie tussen de leraar en leerlingen en leerlingen onderling zo weinig mogelijk beïnvloedt (Riley-Tillman et al., 2005; Shapiro, 2004).

Leergedrag van leerlingen kan gevolgd worden door middel van directe observatie, maar de beschikbare instrumenten zijn te tijdsintensief om gebruikt te worden als leergedragvolgsysteem. Onderzoek moet uitwijzen of er een instrument ontwikkeld kan worden met minder gedragsindicatoren en dat minder tijd vraagt om afgenomen te worden. Een korte gedragsbeoordelingsschaal waar het te observeren leergedrag is geoperationaliseerd kan een goed alternatief zijn.

Gedragsindicatoren om Leergedrag te Meten

Er bestaat nog geen eenduidigheid over wat er gemeten zou moeten worden om het leergedrag van leerlingen in kaart te brengen. Aan taak-gedrag en niet aan taak-gedrag zijn indicatoren die uit veel onderzoeken naar voren kwamen als goede voorspellers van leerprestaties van leerlingen (e.g. Chafouleas et al., 2007; Riley-Tillman et al., 2008; Volpe et

al., 2009). Ook storend gedrag blijkt een sterke negatieve relatie te hebben met leerprestaties. Hoe meer storend gedrag leerlingen vertonen, hoe lager hun prestaties vaak zijn (e.g. Volpe et al., 2009; Fredricks et al., 2004; Morgan et al., 2008).

Een regelmatig als maatstaf voor het leergedrag van leerlingen gebruikt begrip is betrokkenheid wordt regelmatig als maatstaf voor het leergedrag van leerlingen gebruikt. Leerlingen die meer betrokken zijn, presteren beter en lopen minder risico voortijdig de school te verlaten (e.g. Fredricks et al., 2004; Furlong & Christenson, 2008; Markussen et al., 2011). Binnen betrokkenheid kunnen drie categorieën onderscheiden worden: gedragsmatig (het leergedrag, aandacht voor en actieve deelname aan de activiteiten in de klas), cognitief (interesse in en opvatting over het belang van leren, het stellen van doelen en zelfregulatie) en affectief (relatie en support van ouders, leraren en klasgenoten) (Fredricks et al., 2004). Furlong en Christenson (2008) voegen de academische betrokkenheid als vierde categorie toe en splitsen daarmee de gedragsmatige categorie in tweeën: onder gedragsmatig wordt de deelname aan activiteiten in de les verstaan en onder de academische de hoeveelheid tijd die een leerling bezig is met de taken voor school. De academische en gedragsmatige betrokkenheid van leerlingen kan in kaart gebracht worden door observatie, terwijl de cognitieve en affectieve betrokkenheid minder goed observeerbaar is en beter getoetst kan worden met zelfbeoordelingsformulieren (Furlong & Christenson, 2008). Deze laatste twee categorieën zijn wel van invloed op het leergedrag van de leerling, maar zijn, omdat ze niet goed observeerbaar zijn, niet geschikt om het leergedrag te volgen.

De academische of gedragsmatige betrokkenheid van een leerling kan onderverdeeld worden in actieve (e.g. schrijven, vinger opsteken, hardop lezen, praten met leraar of medeleerling over de lesstof) en passieve (e.g. luisteren, stil lezen, kijken in een werkboek) betrokkenheid (Shapiro, 2004). Dit onderscheid wordt ook gemaakt in de *Behavioral Observation of Students in Schools*. In een onderzoek naar de betrokkenheid van basisschoolleerlingen met ADHD waarin gebruik werd gemaakt van de *Behavioral Observation of Students in Schools* bleek een duidelijk verschil in betrokkenheid tussen deze groep en een controlegroep (Vile Junod, DuPaul, Jitendra, Volpe & Cleary, 2005). De leerlingen met ADHD lieten evenveel actief leergedrag zien als de leerlingen in de controlegroep, maar waren minder tijd passief aan hun taak en lieten in plaats daarvan meer storend gedrag zien dan hun medeleerlingen zonder ADHD. Hoewel de leerlingen met ADHD voor dit onderzoek geselecteerd waren vanwege hun lage leerprestaties is niet bekend in hoeverre het geobserveerde gedrag correleerde met die leerprestaties en met het leergedrag

van de leerlingen. Verder onderzoek lijkt nodig om duidelijkheid te krijgen welke gedragingen het beste het leergedrag van leerlingen kunnen voorspellen en daarbij voldoende sensitief zijn om ontwikkeling te meten.

Dit onderzoek wil antwoord geven op de vraag: ‘Welk gedrag moet gemeten worden om het leergedrag van leerlingen te volgen?’. Wanneer helder is wat adequaat leergedrag is, kan dit gebruikt worden voor het ontwikkelen van een instrument om het leergedrag van leerlingen te volgen. In dit onderzoek is onderscheid gemaakt tussen betrokken en niet-betrokken gedragingen. Omdat leerlingen die meer betrokken zijn beter presteren (e.g. Fredricks et al., 2004; Furlong & Christenson, 2008; Markussen et al., 2011), is de verwachting dat het leergedrag van deze leerlingen door de mentor als positief wordt gezien. Om te onderzoeken of er verschil gemaakt moet worden in de mate van betrokkenheid (actief versus meer passief) zijn er twee categorieën gebruikt voor niet-betrokken en twee voor positief betrokken leerlingen. Het gedrag van een leerling die niet bezig is met de taak van de leraar, valt onder de categorie *niet aan taak*. Wanneer de leerling daarnaast de leeromgeving voor zichzelf of een ander actief verstoort, valt zijn gedrag onder *storend gedrag*. Bij positieve betrokkenheid wordt er onderscheid gemaakt tussen het bezig zijn met de taak van de leraar en het actief reageren op een stimulus van de leraar. Binnen de categorie *actief leergedrag* valt elk gedrag waarmee de leerling actief reageert op een stimulus van de leraar. Is de leerling bezig met de door de leraar opgegeven taak, dan valt zijn gedrag onder *aan taak*.

De verwachting is dat leerlingen die beter leergedrag hebben volgens hun mentoren, meer *aan taak* en *actief leergedrag* en minder *niet aan taak* en *storend gedrag* laten zien. Daarnaast is de verwachting dat leerlingen die volgens hun mentor *zeer goed* leergedrag hebben, meer *actief leergedrag* laten zien dan de rest van de klas.

Methode

Steekproef en setting

Het onderzoek vond plaats op een binnenstadsschool voor voortgezet onderwijs in een grote stad (500.000 inwoners). De wijk waarin deze school staat, kenmerkt zich door veel inwoners met een lage sociaal-economische status (Rijksinstituut voor Volksgezondheid en Milieu, 2006). Dit onderzoek was deel van een groter onderzoek gericht op de ontwikkeling van een leergedragvolgsysteem voor het voortgezet onderwijs. De klassen werden geselecteerd door de school en alle betrokken leraren stemden toe in deelname aan het onderzoek.

Voor deelname aan dit specifieke onderzoeksproject werden die leerlingen geselecteerd van wie er minimaal negen minuten over minimaal vijf verschillende lessen per onderzoeksdag (in totaal achttien minuten) aan observatiedata verzameld was. Deze minuten werden random getrokken. In totaal voldeden de data van 45 leerlingen (24 jongens, 21 meisjes) aan deze criteria. De leeftijd van de leerlingen varieerde van 13,5 tot 16,3 jaar ($M = 14,72$, $SD = 0,67$). Van twee leerlingen was de leeftijd niet bekend. De leerlingen kwamen uit vier verschillende klassen van het tweede leerjaar: vijf uit de vmbo-kaderberoepsgerichte klassen (kbl), tien uit de vmbo-basisberoepsgerichte klassen (bbl), twaalf uit het mavo en achttien uit het havo. De leerlingen zijn geboren in Nederland (68,9%), Marokko (6,7%), Irak (4,4%), Pakistan (4,4%), Suriname (2,2%), Tunesië (2,2%), de Dominicaanse Republiek (2,2%) en Burundi (2,2%). Van twee leerlingen was het geboorteland niet bekend (4,4%).

De school waar het onderzoek plaatsvond, is een onderzoeksschool van de universiteit Leiden. Ouders hebben bij inschrijving van hun kind op school toestemming gegeven mee te werken aan onderzoeken vanuit de universiteit.

Leergedragvolgsysteem

Het overkoepelende doel van het grote onderzoek was een leergedragvolgsysteem te ontwikkelen om het leergedrag van leerlingen in kaart te brengen. Om een dergelijk systeem te ontwikkelen moet duidelijk zijn welke gedragingen specifiek gemeten moeten worden met dit systeem. Het doel van dit onderzoek was te onderzoeken welke gedragingen gemeten moeten worden om het leergedrag van leerlingen te kunnen volgen.

Afhankelijke variabelen. Voor dit onderzoek werden vier gedragingen van leerlingen geobserveerd: twee positieve gedragingen (één actief, één passief) en twee negatieve gedragingen (één actief, één passief).

- *Actief leergedrag*: de leerling is bezig met de door de leraar opgedragen taak en reageert actief (mondeling, schriftelijk of motorisch) op de leraar.
- *Aan taak*: de leerling is bezig met de door de leraar opgedragen taak.
- *Niet aan taak*: de leerling is niet bezig met de door de leraar opgedragen taak.
- *Storend gedrag*: de leerling is niet bezig met de door de leraar opgedragen taak en verstoort de leeromgeving voor zichzelf en/of anderen.

De leerlingen werden gedurende twee dezelfde lesdagen geobserveerd, in twee opeenvolgende weken. Een minuut lang werd iedere tiende seconde het gedrag van een leerling geobserveerd en gescoord in één van de vier categorieën (momentary time sampling). Tabel 1 geeft een deel van het gebruikte observatie-instrument weer. Na iedere minuut werd een volgende leerling geobserveerd; dit rouleerde gedurende de gehele les. De observatoren maakten gebruik van een mp3-speler met daarop een zelfde geluidsbestand, dat om de tien seconden een piepgeluid liet horen als teken dat het gedrag van de leerling gescoord moest worden. Alleen op dit moment keek de observator naar de leerling, om zo weinig mogelijk invloed uit te oefenen op de situatie en beïnvloed te worden.

Alle scores van de vier categorieën van de in totaal achttien minuten werden bij elkaar opgeteld en per categorie werd berekend wat het percentage van het totaal was.

De observatoren die de data voor dit onderzoek hebben verzameld, zijn door middel van het scoren van gedrag vanaf een video getraind in het gebruik van het instrument. Daarnaast werden op onderzoeksdagen door twee observatoren in dezelfde klas tegelijkertijd data verzameld om zo de interbeoordelaarsbetrouwbaarheid te kunnen bepalen. De observatoren haalden voor de vier indicatoren een gemiddelde interbeoordelaarsbetrouwbaarheid van 0,89 (Espin & Verhallen, 2011).

Tabel 1. *Deel van het gebruikte observatie-instrument*

Sec:	10	20	30	40	50	60
Actief leergedrag						
Aan taak						
Niet aan taak						
Storend gedrag						

Criteriumvariabele. Om te bepalen of de verschillende vormen van leergedragingen voorspellend waren voor het leergedrag van leerlingen, werd gebruik gemaakt van een beoordelingsformulier voor leraren. Het instrument was voor dit onderzoek ontwikkeld en

tijdens dit onderzoek voor het eerst gebruikt. De vraag die gesteld werd, was: ‘Hoe beoordeelt u het leergedrag van de leerling?’ Op het formulier stond een korte uitleg van leergedrag: ‘Onder leergedrag wordt het gedrag verstaan dat verbonden is met en invloed heeft op het leren, zoals het luisteren naar instructie en het zelfstandig kunnen werken.’ Op het formulier kon de leraar per leerling op een Likertschaal aangeven hoe hij het leergedrag van zijn leerlingen beoordeelde. De mogelijke scores varieerden van 1 (zeer goed) tot 4 (slecht). Het instrument werd door de mentoren van de verschillende klassen ingevuld.

Design

In dit onderzoek zijn vier afhankelijke variabelen (leerlinggedrag: *actief leergedrag, aan taak, niet aan taak en storend gedrag*) vergeleken met één criteriumvariabele (beoordeling leergedrag door mentor).

Procedure

Nadat de school toestemming gaf voor dit onderzoek werd er geobserveerd in zes verschillende klassen. De jaarcoördinator informeerde de leraren over de aard van het onderzoek en de data waarop de observaties zouden plaatsvinden. De observaties werden verspreid over een periode van zes lesweken verzameld en waren voor iedere klas gepland op eenzelfde lesdag in twee achtereenvolgende weken. Aan het begin van de les vertelde de leraar wie de observatoren waren. De leerlingen werden niet op de hoogte gebracht van het doel van de aanwezigheid van hen, zodat hun gedrag zo min mogelijk beïnvloed zou worden. Gedurende de gehele dag werden de leerlingen tijdens de verschillende lessen geobserveerd door dezelfde observator. De observator zat op een vaste plek om de les zo min mogelijk te verstoren. Bij lesuitval, afwezigheid van de leerling of een leerling die buiten het gezichtsveld van de observator zat, zag de observator af van scoren en sloeg hij de betreffende leerling over.

De mentoren vulden twee weken na de observaties het beoordelingsformulier in waarop ze aangaven hoe zij het leergedrag van hun mentorleerlingen beoordeelden.

Voor deze studie werden de data van 45 van de in totaal 119 geobserveerde leerlingen gebruikt. Deze 45 leerlingen waren gekozen, omdat ze minimaal negen minuten per onderzoeksdag aan observatiegegevens hadden. Als een leerling meer dan negen minuten per onderzoeksdag had, werden de minuten die voor het onderzoek gebruikt werden random getrokken. Hierbij werd rekening gehouden met het criterium dat de data uit ten minste vijf

verschillende lesuren moest komen. In totaal werden er van elk van deze 45 leerlingen achttien minuten aan observatiegegevens gebruikt.

Statistische analyse

Data-inspectie. Om de data te controleren op missende waarden, uitbijters en normaliteit werd een univariate data-inspectie gedaan. De numerieke variabelen werden op normaliteit getest met behulp van een histogram met normaalverdeling en een spreidingsdiagram. De gestandaardiseerde scheefheid en gepiekttheid werd berekend om te zien of er uitgegaan mocht worden van een normaalverdeling. Met een boxplot werd gekeken of er uitbijters waren en vervolgens werden deze cases nader onderzocht. Voor de categoriale variabelen werden de gegevens door middel van een frequentietabel gecontroleerd op de verdeling en op missende waarden.

Variantieanalyse. Om te zien of de verschillende beoordelingen van de mentor groepen opleverden die qua geobserveerd leergedrag van elkaar verschilden, werd gebruikgemaakt van een variantieanalyse. Als eerste werd een MANOVA uitgevoerd. De afhankelijke variabelen waren de vier leerlinggedragingen, de onafhankelijke variabele was de leraarbeoordeling. Met de MANOVA werd getoetst of er sprake was van een significant verschil in gedrag tussen de drie verschillende categorieën van de mentorbeoordeling. Een voorwaarde voor het uitvoeren van de variantieanalyse was dat de steekproef random getrokken en de data normaal verdeeld waren.

Na de MANOVA werd voor elk van de vier afhankelijke variabelen een ANOVA uitgevoerd en door middel van een post hoc-toets werd gekeken welke gemiddelden significant van elkaar verschilden. Voor dit doel werd gekozen voor de LSD-toets. Met deze toets werden de groepen paarsgewijs met elkaar vergeleken en werd de verhoogde kans op een type 1 fout gecorrigeerd.

De power van de test werd getoetst door middel van het berekenen van de Cohen's effect size. Hierbij werd gekeken hoe groot de waarschijnlijkheid was dat de nulhypothese verworpen werd terwijl de alternatieve hypothese waar was.

Voor de toetsen werd gekozen voor een significantieniveau α van 0,05 (95%).

Resultaten

In deze sectie bevinden zich de resultaten van de toetsen die gedaan zijn om antwoord te geven op de hypothesen, voorafgegaan door een data-inspectie. De responsvariabele in dit onderzoek is de beoordeling van het leergedrag van de leerling door de leraar; de predictorvariabelen zijn de categorische variabelen *actief leergedrag*, *aan taak*, *niet aan taak* en *storend gedrag*.

Data-inspectie. Het leergedrag van twee van de 45 leerlingen werd door de mentoren als slecht beoordeeld. Omdat dit aantal te weinig is voor statistische berekeningen zijn de categorieën *minder goed* en *slecht* bij elkaar gevoegd tot één categorie *minder goed*. In tabel 2 staan de aantallen weergegeven na samenvoeging van deze categorieën. Het leergedrag van de meeste leerlingen wordt als *goed* tot *zeer goed* gezien (68,2%). De leerlingen van de verschillende klassen zijn vertegenwoordigd in alle drie de groepen. Ook de verdeling van jongens en meisjes over de verschillende groepen is nagenoeg gelijk. De gemiddelde leeftijd van de steekproef ($N = 42$, twee missings) is 14,71 jaar ($SD = 0,67$). Tussen de drie categorieën is er een klein verschil te zien van de gemiddelde leeftijd van de leerlingen. De gemiddelde leeftijd van de leerlingen die behoren tot de categorieën *zeer goed* en *goed* is iets hoger dan die van de categorie *minder goed*. Dit verschil is niet significant.

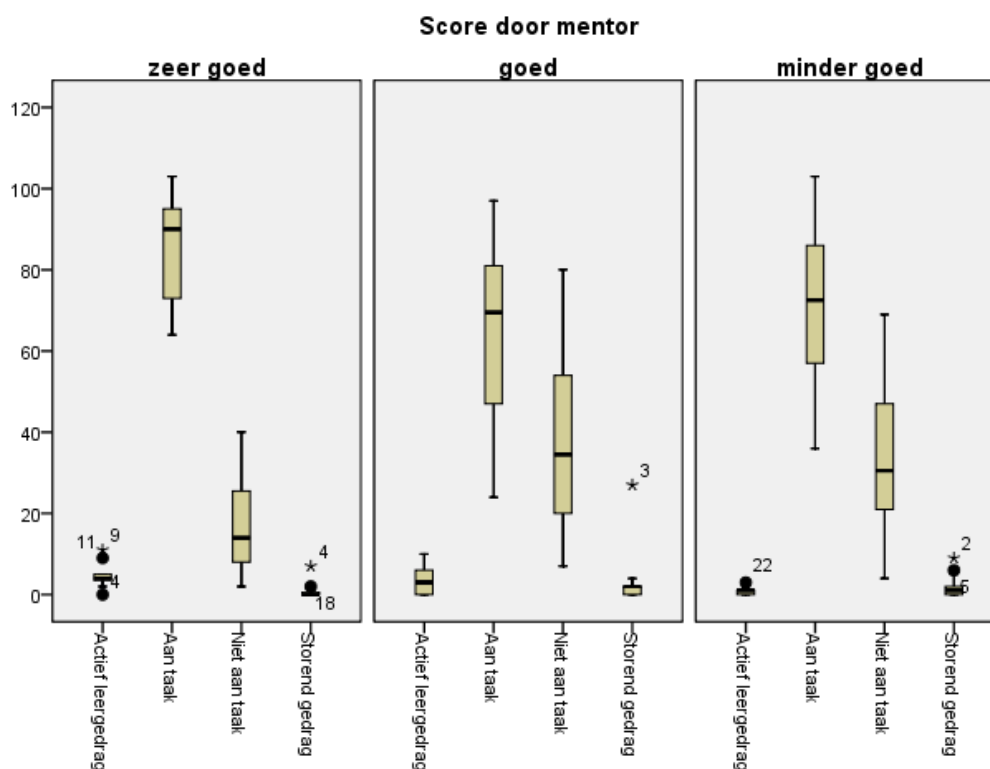
Tabel 2. *Overzicht van het leergedrag van leerlingen, beoordeeld door de mentor met een overzicht van de verdeling over geslacht, leeftijd en onderwijsvorm.*

	Geslacht		Leeftijd		Onderwijsvorm								
	Jongen		Meisje		kb		bb		tl		havo		
	n	%	n	%	<i>M</i>	n	%	n	%	n	%	n	%
Zeer goed	5	20,8	6	30,0	14,82	1	25,0	2	20,0	3	25,0	5	27,8
Goed	11	45,8	8	40,0	14,75	1	25,0	7	70,0	5	41,7	6	33,3
Minder goed	8	33,3	6	30,0	14,57	2	50,0	1	10,0	4	33,3	7	38,9
totaal	24		20		14,71	4		10		12		18	

kb = kaderberoepsgerichte leerweg; bb = basisberoepsgerichte leerweg met leerwegondersteuning; tl = theoretische leerweg.

Bij de data-inspectie bleken er vier uitbijters te zijn (figuur 1: nr. 2, 3, 4, 9). Eén van hen was een leerling uit de categorie *zeer goed* die 11 keer scoort op *actief leergedrag* ($M = 2,87$, $SD = 2,99$). Na inspectie van de overige scores in deze categorie, bleken alle waarden van 0 tot en met 11 voor te komen. Er bleek sprake te zijn van een positieve scheefheid. De meeste leerlingen vertoonden geen *actief leergedrag*. De uitbijter week één score af van de

daaropvolgende score en was van minimale invloed op het gemiddelde en is daarom niet uit de dataset verwijderd. Drie andere leerlingen scoorden 7, 9 en 27 keer bij *storend gedrag*. De score van 27 op *storend gedrag* week sterk af van de overige scores en was van grote invloed op het gemiddelde ($M = 2,00$, $SD = 4,33$; zonder deze uitbijter: $M = 1,43$, $SD = 2,07$). Deze uitbijter werd daarom niet meegenomen tijdens de berekeningen. De beide andere uitbijters werden in de dataset gehouden, omdat hun invloed op het gemiddelde minder groot was (gemiddelde zonder de drie uitbijters: $M = 1,12$).



Figuur 1. Boxplots van de scores van beoordeling van leergedrag door de mentor en de geobserveerde gedragingen van de leerlingen. De uitbijter die niet meegenomen werd in de analyses is nummer 3.

De gestandaardiseerde kurtosis voldeed bij de categorieën *actief leergedrag*, *aan taak* en *niet aan taak* aan de voorwaarden van een normaalverdeling (tabel 3b). De gestandaardiseerde kurtosis van de categorie *storend gedrag* was ook na verwijdering van de uitbijter nog steeds hoog. Daarnaast was deze curve ook positief scheef ($Z_{\text{skewness}} = 5,50$). De helft (50%) van de leerlingen scoorde nul keer op *storend gedrag* en 36,4% scoorde hier één of twee keer, de overige zes leerlingen (13,6%) scoorden tussen de vier en de negen keer op *storend gedrag*.

Ook bij de categorie *actief leergedrag* was sprake van een positieve scheefheid. Bijna de helft van de leerlingen (45,5%) scoorde nul of één keer op *actief leergedrag*, 30% scoorde tussen de vier en zes keer op *actief leergedrag* en de rest (zes leerlingen) had een score tussen de zeven en elf keer op deze categorie. De curve van *aan taak* was negatief scheef ($Z_{skewness} = -1,75$). De helft van de leerlingen scoorde tussen de 77 en 103 keer op *aan taak*, de andere helft had een *aan taak*-score tussen de 24 en 77. Bij de *niet aan taak*-scores had 50% een score van 25 of minder, de andere helft scoorde tussen de 26 en 80 keer *niet aan taak* in achttien minuten. Ook hier was sprake van een positieve scheefheid.

Omdat de data niet voldeden aan de criteria voor een normaalverdeling, zullen de resultaten met voorzichtigheid geïnterpreteerd moeten worden. Gezien het explorierend karakter van dit onderzoek werd er toch voor gekozen parametrische toetsen uit te voeren.

Tabel 3a. *Overzicht van de beschrijvende gegevens van de numerieke variabelen.*

	<i>N</i>	Min.	Max.	<i>M</i>	<i>SD</i>	$Z_{skewness}$	$Z_{kurtosis}$
Actief leergedrag	45	0	11	2,87	2,99	2,89	0,05
Aan taak	45	24	103	71,60	20,69	-1,57	-0,57
Niet aan taak	45	2	80	31,53	20,66	1,95	-0,38
Storend gedrag	45	0	27	4,33	4,33	13,28	37,52

Tabel 3b. *Overzicht van de beschrijvende gegevens van de numerieke variabelen zonder de uitbijter.*

	<i>N</i>	Min.	Max.	<i>M</i>	<i>SD</i>	$Z_{skewness}$	$Z_{kurtosis}$
Actief leergedrag	44	0	11	2,93	2,99	2,79	0,42
Aan taak	44	24	103	72,18	20,55	-1,75	-0,34
Niet aan taak	44	2	80	31,45	20,90	1,96	-0,45
Storend gedrag	44	0	9	1,43	2,07	5,50	5,76

Tabel 3c. *Overzicht van gemiddelde scores leerlinggedrag per mentorscores.*

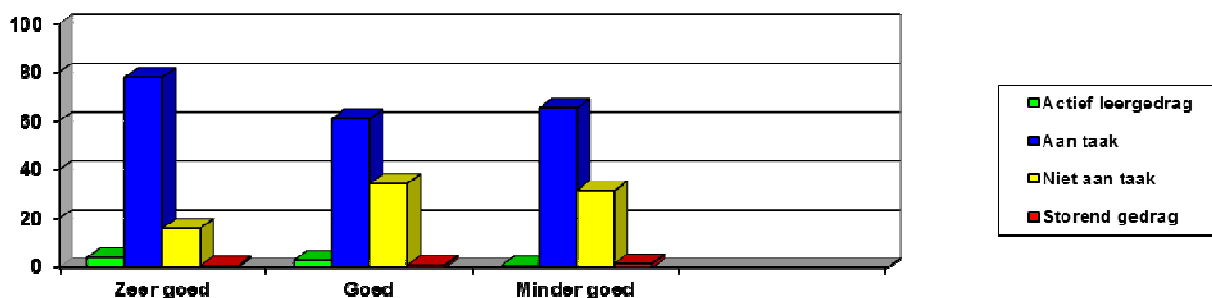
	<i>Score mentor</i>								
	<i>Zeer goed (n=11)</i>			<i>Goed (n=19)</i>			<i>Minder goed (n=14)</i>		
<i>Leerlinggedrag</i>	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%	<i>M</i>	<i>SD</i>	%
Actief leergedrag	4,64	3,04	4,27	3,47	3,20	3,21	0,86	0,86	0,80
Aan taak	84,82	13,26	78,02	65,95	21,64	61,07	70,71	20,47	65,47
Niet aan taak	17,64	12,27	16,23	37,32	22,88	34,56	34,36	19,57	31,82
Storend gedrag	0,91	2,12	0,84	1,26	1,37	1,17	2,07	2,73	1,91

Om te onderzoeken op welke manier de indicatoren gerelateerd waren aan de beoordeling van de mentor, werd een MANOVA uitgevoerd. De Levene's test voor de betrouwbaarheid gaf een significante waarde voor *actief leergedrag* ($p = 0,01$). Dit betekent dat er sprake was

van ongelijke variantie voor deze groep. Voor de andere groepen werd voldaan aan de assumptie van gelijke variantie. Ondanks dat de assumptie van gelijke variantie niet opging voor *actief leergedrag*, werd deze indicator omwille van het explorerende karakter van dit onderzoek toch meegenomen in de toetsen.

De MANOVA liet zien dat er sprake is van significante groepseffecten (Wilks' Lambda = 0,64, $F(6, 78) = 3,20, p = 0,007$). Als vervolg op de MANOVA werd voor elk van de vier gedragsindicatoren een ANOVA uitgevoerd om te zien welke van de gedragingen van elkaar verschilden en voor welke score van de mentor dat gold. De resultaten lieten een significant groepseffect zien voor drie van de vier variabelen: *actief leergedrag* ($F(2, 41) = 6,99, p < 0,01$), *aan taak* ($F(2, 41) = 3,31, p < 0,05$) en *niet aan taak* ($F(2, 41) = 3,70, p < 0,05$).

Voor deze drie gedragingen werd door middel van een LSD-toets gekeken waar de verschillen lagen en is de power van de toets gemeten door Cohen's d te berekenen. Bij *actief leergedrag* werd er een significant verschil gemeten tussen de groepen *zeer goed* en *minder goed* ($t_{(23)} = 4,45; p < 0,000; ES = 1,79$) en de groepen *goed* en *minder goed* ($t_{(21)} = 3,40; p = 0,003; ES = 1,11$). Bij *aan taak* was er een verschil tussen *zeer goed* en *goed* ($t_{(28)} = 2,61; p = 0,014; ES = 1,05$). De categorie *minder goed* scoorde gemiddeld meer *aan taak* dan de categorie *goed*. Het verschil tussen deze groepen was niet significant. Ook bij *niet aan taak* was er een significant verschil tussen de groepen *zeer goed* en *goed* ($t_{(28)} = -3,07; p = 0,005; ES = -1,07$) en de groepen *zeer goed* en *minder goed* ($t_{(23)} = -2,47; p = 0,021; ES = -1,02$). De effectgrootte bij *actief leergedrag* van het verschil tussen de groepen *zeer goed* en *minder goed* (Cohen's d is hoger dan 1,30) was zeer groot. Bij de andere verschillen kan gesproken worden van een groot (negatief) effect (waarden tussen 0,80 en 1,29, negatief of positief).



Figuur 1. Gemiddelde scores van leerlinggedrag afgezet tegen beoordeling mentor

In deze data waren weinig scores op de categorieën *actief leergedrag* en *storend gedrag*. Omdat deze categorieën als behorend bij respectievelijk *aan taak* en *niet aan taak* kunnen worden beschouwd (Espin & Verhallen, 2011), voegden we de categorieën samen tot een categorie *positief leergedrag*, bestaande uit *actief leergedrag* en *aan taak* en een categorie *negatief leergedrag*, bestaande uit *storend gedrag* en *niet aan taak*. De MANOVA liet zien dat er sprake is van significante groepseffecten ((Wilks' Lambda = 0,85, $F(2, 41) = 3,66$, $p = 0,034$). De uitgevoerde ANOVA's toonden een significant verschil tussen de categorieën *zeer goed* en *goed* ($t_{(28)} = 2,60$; $p = 0,015$) en *zeer goed* en *minder goed* ($t_{(23)} = 2,50$; $p = 0,020$). Leerlingen die door hun mentor beoordeeld werden als leerlingen met *zeer goed* leergedrag, lieten meer *positief leergedrag* zien in de klas dan leerlingen die volgens hun mentor *goed* of *minder goed* leergedrag hadden. Ook nu werd er geen significant verschil gevonden tussen de categorieën *goed* en *minder goed*.

Tabel 4. *Overzicht van gemiddelde scores positief en negatief leerlinggedrag per mentorscores.*

<i>Leerlinggedrag</i>	<i>Score mentor</i>								
	<i>Zeer goed</i>			<i>Goed</i>			<i>Minder goed</i>		
	<i>M</i>	<i>SD</i>	<i>%</i>	<i>M</i>	<i>SD</i>	<i>%</i>	<i>M</i>	<i>SD</i>	<i>%</i>
Positief leergedrag	89,46	16,30	82,29	69,42	24,84	64,28	71,57	21,33	66,27
Negatief leergedrag	18,55	14,39	17,07	38,58	24,25	35,73	36,43	22,30	33,73

Discussie

De leerprestaties van leerlingen worden beïnvloed door hun gedrag in de klas (Algozzine et al., 2011). Leerlingen die taakgebonden gedrag laten zien, presteren beter (Chafouleas et al., 2010a; Chafouleas et al., 2010b, Morgan et al., 2008; Fredricks et al., 2004). Omdat scholen veel kunnen doen om storend gedrag van leerlingen te verminderen en daardoor de leerprestaties te verhogen (Walker et al., 2003), moeten scholen daarom inzetten op het optimaliseren van het leergedrag van leerlingen. Er worden diverse programma's ingezet om probleemgedrag te verminderen. In het kader hiervan wordt de behoefte aan een instrument om de ontwikkeling van het leergedrag van leerlingen te meten en te volgen steeds groter (Evans & Owens, 2010). Bestaande instrumenten om leergedrag van leerlingen in kaart te brengen, blijken niet geschikt om met regelmaat ingezet te worden (Briesch et al., 2010). Om effecten van interventies om leergedrag te verbeteren te kunnen vaststellen, is er behoefte aan een instrument dat het gedrag kan screenen, met korte tussenpozen gebruikt kan worden en gevoelig is voor veranderingen (Chafouleas et al., 2010b; Shapiro, 2004)

Er bestaat nog geen eenduidigheid wat er gemeten zou moeten worden om leergedrag van leerlingen in kaart te brengen en er is behoefte aan duidelijkheid over welke gedragsindicatoren in een leergedragvolgsysteem thuishoren (Chafouleas et al., 2010b). In dit exploratief onderzoek is gezocht naar gedragingen die adequaat leergedrag van leerlingen kunnen voorspellen. De onderzoeksvraag was: 'Welke gedrag moet gemeten worden om het leergedrag van leerlingen te volgen?'. Wanneer duidelijk is welke gedrag van leerlingen nodig is om adequaat leergedrag te voorspellen, kan er een instrument ontwikkeld worden waarmee het leergedrag van leerlingen gevolgd kan worden. Hierdoor kunnen leraren meer zicht krijgen op de ontwikkelingsstappen die een leerling zet op het gebied van leergedrag.

Op basis van literatuurstudie is gekozen voor *actief leergedrag*, *aan taak*, *niet aan taak* en *storend gedrag* als gedragsindicatoren (e.g. Espin & Yell, 1994; Fredricks et al., 2004; Furlong & Christenson, 2008; Shapiro, 2004; Miles & Stipek, 2006). Deze vier gedragingen van leerlingen zijn geobserveerd en de scores hierop vergeleken met de beoordeling van het leergedrag van de leerling door zijn mentor.

Uit de resultaten van dit onderzoek is gebleken dat *actief leergedrag* in staat is leerlingen met *zeer goed*, *goed* en *minder goed* leergedrag van elkaar te onderscheiden. Leerlingen die volgens hun mentor *zeer goed* leergedrag hebben, vertonen meer *actief leergedrag*. Deze uitkomst komt overeen met de uitkomsten van diverse onderzoeken dat leerlingen die actief meedoen met de les, beter presteren (e.g. Chafouleas et al., 2010a; Morgan et al., 2008;

Fredricks et al., 2004; Furlong & Christenson, 2005). Een actieve houding in de klas, lijkt voorspellend te zijn voor adequaat leergedrag. Wel moet deze uitkomst van dit onderzoek met voorzichtigheid geïnterpreteerd worden, omdat niet voldaan werd aan de assumptie van gelijke variantie tussen de steekproef en de populatie. Het is mogelijk dat dit komt doordat *actief leergedrag* weinig bleek voor te komen en de kans daardoor groter is dat het resultaat op toeval berust. In een volgend onderzoek met meer respondenten zou getoetst moeten worden of deze uitkomsten replicerbaar zijn. Daarvoor zou kritisch gekeken kunnen worden bij welke lessen er geobserveerd gaat worden. Bij leraargebonden lessen zal meer gelegenheid zijn tot actief leergedrag en de verwachting bestaat dat het onderscheid tussen betrokken en niet betrokken leerlingen dan beter zichtbaar is dan bij bijvoorbeeld lessen waarbij de leerlingen met behulp van de computer zelfstandig aan opdrachten mogen werken.

Leerlingen die de beoordeling *zeer goed* hebben gekregen voor hun leergedrag scoren significant vaker *aan taak* dan leerlingen in de categorieën *goed* of *minder goed*. Ook dit is conform de verwachting en lijkt te bevestigen dat leerlingen die meer betrokken zijn beter presteren (e.g. Chafouleas, et al., 2010a; Morgan, et al., 2008). Opvallend is echter dat *aan taak*-zijn niet onderscheidend werkt tussen de groepen *goed* en *minder goed*. Dit geldt ook voor het *niet aan taak*-zijn. Op grond van dit onderzoek lijkt de hoeveelheid tijd die een leerling aan zijn taak bezig is, geen goede voorspeller te zijn voor zijn leergedrag. Deze gedragsindicator lijkt onvoldoende sensitief te zijn om onderscheid te maken tussen leerlingen die goed en minder goed leergedrag laten zien. Een verklaring hiervoor zou kunnen zijn dat ten onrechte gedacht werd dat de leerling (niet) aan zijn taak bezig was. Het was bijvoorbeeld niet altijd mogelijk te zien wat een leerling deed op de computer en niet altijd was helder wat de opdracht was die de leerling op de computer moest doen. Het bekijken van een filmpje was soms adequaat leergedrag en soms ook niet. Daarnaast is *aan taak*- en *niet aan taak*-zijn minder zichtbaar dan *actief leergedrag* en *storend gedrag*. Ook voor mentoren kan het moeilijk zijn het leergedrag van hun leerlingen goed te duiden. Een leraar kan soms vooral gefocust is op het zichtbare, storende gedrag van de leerling en niet opmerken dat deze een deel van de tijd taakgericht bezig is. Dit maakt het gebruik van een korte beoordelingsschaal kwetsbaarder (Shapiro en Kratochwill, 2000).

Een andere verklaring voor de resultaten van dit onderzoek zou kunnen zijn dat er gebruikgemaakt is van slechts één mentorbeoordeling als criteriumvariabele. De kans op bias (e.g. halo-effect) is hierdoor groter en de score is sterk gekoppeld aan de beoordelaar (Chafouleas et al., 2010a). De kans is aanwezig dat de beoordeling van de leraar beïnvloed

werd door wat er kort voordat hij de beoordeling gaf was voorgevallen. Om deze invloed te verminderen, zou hier bij de instructie expliciet aandacht aan besteed moeten worden (Shapiro & Kratochwill, 2000).

Daarnaast is uit onderzoek naar het beoordelingsgedrag van leraren in het kleuteronderwijs gebleken, dat leraren beter in staat zijn het gedrag van hun leerlingen te ordenen dan een absoluut oordeel te geven (Briesch et al., 2010). Verder geven Shapiro en Clemens (2005) aan dat informatie verkregen van leraren niet als geheel valide kan worden gezien. Indirecte instrumenten - zoals gedragsvragenlijsten - geven vooral informatie over de perceptie van de beoordelaar en niet over het actuele niveau waarop iemand presteert. Uit onderzoek naar het gebruik van korte gedragsbeoordelingsschalen die direct na een observatie ingevuld worden, blijkt dat er wel een hoge correlatie is tussen de scores van leraren op de korte gedragsbeoordelingsschalen en de directe observatie van een externe onderzoeker (Riley-Tillman et al., 2008). Verder onderzoek naar het gebruik van de leraarbeoordeling voor leergedrag moet aantonen hoe robuust en betrouwbaar dit beoordelingsformulier is.

Op grond van dit onderzoek met deze data kan voorzichtig geconcludeerd worden dat *actief leergedrag* in staat is onderscheid te maken tussen het leergedrag van leerlingen in het voortgezet onderwijs. Uit dit onderzoek bleek dat hoe actiever leerlingen betrokken zijn bij de lesactiviteiten, hoe beter hun leergedrag is volgens hun mentor. *Actief leergedrag* is daarmee een goede voorspeller van leergedrag van leerlingen, maar minder goed dan gehoopt. Daarnaast is het *aan taak*-gedrag van leerlingen goed in staat onderscheid te maken tussen leerlingen met *zeer goed* leergedrag en *goed/minder goed* leergedrag, maar niet tussen *goed* en *minder goed* leergedrag. Dit maakt dat uit dit onderzoek het *aan taak* zijn niet sensitief genoeg blijkt te zijn en daarmee geen goede voorspeller is van leergedrag van leerlingen.

Beperkingen van het onderzoek en suggesties voor vervolgonderzoek

Dit onderzoek is uitgevoerd op een binnenstadsschool in een wijk met overwegend inwoners met een lage sociaaleconomische status. De resultaten van dit onderzoek zijn dan ook niet representatief voor het gehele voortgezet onderwijs in Nederland en het onderzoek zou ook uitgevoerd moeten worden op andere scholen. Ruim tweederde van de leerlingen uit de steekproef is in Nederland geboren, maar er zijn geen gegevens bekend over de thuistaal. Bij een volgend onderzoek zouden meer achtergrondgegevens verzameld moeten worden om een beter beeld te krijgen van de samenstelling van de steekproef.

Van de categorieën *actief leergedrag* en *storend gedrag* waren te weinig data beschikbaar. De verwachting is dat wanneer er meer geobserveerd wordt bij leraargebonden lessen en minder bij lessen waar leerlingen zelfstandig achter de computer hun werk moeten doen, de actieve gedragsindicatoren meer gescoord zullen worden en dat beter zichtbaar is of leerlingen wel of niet met de taak bezig zijn.

De definities van de categorieën *aan taak*, *actief leergedrag*, *niet aan taak* en *storend gedrag* zouden meer toegesneden moeten worden op de vorm waarin de lessen worden gegeven. Het zou goed zijn wanneer duidelijk omschreven is wat er gedaan wordt als leerlingen werken op de computer (wel/niet observeren) en bij wat voor soort les er het beste geobserveerd kan worden. Interessant zou zijn wanneer er in vervolgonderzoek gezocht zou worden naar mogelijke verschillen tussen lessen (leraargebonden, zelfstandig werken met of zonder computer).

Het beoordelingsformulier is zonder instructie aan de mentoren gegeven. Het is niet bekend of de leraren in dit onderzoek gekeken hebben naar het leergedrag van hun leerlingen van het laatste uur dat ze deze klas hadden of dat het gaat om het leergedrag over een langere periode. Ook is niet bekend of ze hierbij gebruik hebben gemaakt van de leerresultaten van de leerlingen. De verschillende categorieën (*zeer goed*, *goed*, *minder goed* en *slecht*) zijn niet toegelicht en kunnen dus door de diverse mentoren verschillend geïnterpreteerd zijn. Omdat er slechts één score per klas bekend is, was het niet mogelijk de interbeoordelaarsbetrouwbaarheid te meten. Om de robuustheid van dit instrument te meten is meer onderzoek nodig. Het is noodzakelijk te weten in hoeverre het invullen van dit instrument afhankelijk is van andere variabelen, zoals tijd, plaats en beoordelaar. Daarnaast was de definitie van leergedrag wellicht te beperkt en te weinig concreet voor de leraren. Het is belangrijk dat definities duidelijk zijn en nauwkeurig genoeg beschreven (Shapiro & Kratochwill, 2000).

Als criteriumvariabele is gekozen voor een beoordeling door de mentor. Bij volgend onderzoek zou dit formulier met instructie gegeven kunnen worden aan alle leraren van de leerling (of die leraren waar geobserveerd is). De meningen van leraren over leerlingen kan verschillen (Gregory & Weinstein, 2008) en het gedrag van leerlingen kan verschillen per leraar (Gregory & Thompson, 2010). In vervolgonderzoek zouden naast de beoordeling van de mentor andere criteriumvariabelen gebruikt kunnen worden, zoals cijfers van de leerling, de hoeveelheid werk die een leerling verricht tijdens de les, absentie, et cetera.

De rol van de leraar is in dit onderzoek buiten beschouwing gelaten. Er is geobserveerd bij diverse leraren, maar hun invloed op het leergedrag van de leerlingen is niet meegenomen in dit onderzoek. Uit onderzoek blijkt dat leraren invloed hebben op het ontstaan of beperken van probleemgedrag (e.g. Bergin & Bergin, 2009; Espin & Yell, 1994; Gregory & Thompson, 2010; Leflot, Van Lier, Onghena, & Colpin, 2010). Espin en Yell (1994) observeerden drie gedragingen van leerlingen om de effectiviteit van leraargedrag te kunnen bepalen. Deze drie gedragingen waren *actief leergedrag*, *niet aan taak* en *storend gedrag*. De mate van *niet aan taak* zijn van leerlingen bleek een grote voorspeller te zijn van de effectiviteit van leraargedrag en ook *storend gedrag* blijkt een gemiddelde voorspeller te zijn. Hoe effectiever het leraargedrag was, hoe meer de leerlingen met hun taak bezig waren en hoe minder storend gedrag er gezien werd. Dit lijkt bevestigd te worden door de resultaten van het onderzoek naar de betrokkenheid van kinderen met ADHD ((Vile Junod et al., 2005). Deze leerlingen vertoonden evenveel actief leergedrag als hun medeleerlingen uit de controlegroep, maar op het moment dat er geen actieve bijdrage van hen verwacht werd, lieten ze meer storend gedrag zien. Een leraar die zijn leerlingen stimuleert, creëert een hogere betrokkenheid wat zorgt voor betere leerprestaties (Fredricks et al., 2004) en ook leraren die minder negatieve opmerkingen maken, zien in hun klas meer taakgebonden gedrag van leerlingen (Leflot et al., 2010).

Naast het gedrag van de leraar hebben ook de verwachtingen van de leraar en de leerling ten opzichte van elkaar invloed op het leerlinggedrag in de klas (Gregory & Thompson, 2010; McKown & Weinstein, 2008). Een veilige leraar-leerlingrelatie is een goede voorspeller van sterkere motivatie voor het leren, meer deelname aan activiteiten in de klas en betere schoolprestaties (Bergin & Bergin, 2009; Gregory & Thompson, 2010). Leerlingen die agressief gedrag vertonen, kunnen negatieve relaties ontwikkelen met hun leraren en klasgenootjes waardoor ze minder gemotiveerd zijn zich in te zetten voor hun schoolwerk (Miles & Stipek, 2006).

Adequaat leergedrag is de basis voor goede leerprestaties. Leergedrag kan beïnvloed worden en wanneer leraren hun leerlingen actief betrekken, zullen ze zich meer inzetten en zullen hun prestaties verbeteren. Verder onderzoek is nodig om meer helderheid te krijgen over het te meten gedrag en op welke manier dit het beste in kaart gebracht kan worden, zodat het te ontwikkelen leergedragvolgsysteem als middel gebruikt kan worden de leerling te helpen beter te presteren. Adequaat leergedrag zal dan verstrekkende gevolgen hebben voor de leerling zelf, de sfeer in de klas en op termijn ook voor de maatschappij.

References

- Algozzine, B., Wang, C., & Violette, A. S. (2011). Reexamining the Relationship Between Academic Achievement and Social Behavior. *Journal of Positive Behavior Interventions*, 3-16.
- Bergin, C., & Bergin, D. (2009). Attachment in the classroom. *Educational Psychology Review*, 141-170.
- Blom, M., Oudhof, J., Bijl, R., & Bakker, B. (2005). *Verdacht van criminaliteit: allochtonen en autochtonen nader bekeken*. Opgeroepen op July 30, 2011, van Centraal Bureau van de Statistiek.
- 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(3), 408-421.
- Chafouleas, S. M., Briesch, A. M., Riley-Tillman, T. C., Christ, T. J., Black, A. C., & Kilgus, S. P. (2010a). An investigation of the generalizability and dependability of direct behavior rating single 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., Riley-Tillman, T. C., & Christ, T. J. (2009). Direct behavior rating (DBR): an emerging method for assessing social behavior within a tiered intervention system. *Assessment for Effective Intervention*, 34(4), 195-200.
- Chafouleas, S. M., Volpe, R. J., Gresham, F. M., & Cook, C. R. (2010b). School-based behavioral assessment within problem-solving models: Current status and future directions. *School Psychology Review*, 343-349.
- Deno, S. L. (1985). Curriculum-based measurement: the emerging alternative. *Exceptional Children*, 52(3), 219-232.
- Deno, S. L. (2003). Developments in Curriculum-Based Measurement. *The Journal of Special Education*, 184-192.
- Deno, S. L. (2005). Problem-Solving Assessment. In R. Brown-Chidsey, *Assessment for intervention: a problem solving approach* (pp. 10-40). New York: Guilford.
- Douglas, K. (2009). Sharpening Our Focus in Measuring Classroom Instruction. *Educational Researcher*, 38(7), 518-521.
- Espin, C. A., & Yell, M. L. (1994). Critical Indicators of Effective Teaching for Preservice Teachers: Relationship Between Teaching Behaviors and Ratings of Effectiveness. *Teacher Education and Special Education*, 154-169.

- Espin, C., Wallace, T., Campbell, H., Lembke, E. S., Long, J. D., & Ticha, R. (2008). Curriculum-based measurement in writing. Predicting the success of high-school students on state standards tests. *Council for Exceptional Children*, 174-193.
- Espin, C., & Verhallen, M. (2011, February). *Progress monitoring for academic behavior*. Poster session presented at the meeting of Pacific Coast Research Conference, Coronado, California.
- Evans, S. W., & Owens, J. S. (2010). Behavioral Assessment Within Problem-Solving Models: Finding Relevance and Expanding Feasibility. *School Psychology Review*, 39(3), 427-430.
- Feinberg, A. B., & Shapiro, E. S. (2003). Accuracy of teacher judgments in predicting oral reading fluency. *School Psychology Quarterly*, 18(1), 52-65.
- Fleming, C. B., Harachi, T. W., Cortes, R. C., Abbott, R. D., & Catalano, R. F. (2004). Level and change in reading scores and attention problems during elementary school as predictors of problem behavior in middle school. *Journal of Emotional and Behavioral Disorders*, 12(3), 130-144.
- 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(1), 59-109.
- Fuchs, L. S., Fuchs, D., & Hamlett, C. L. (1993). Technological Advances Linking the Assessment of Students' Academic Proficiency to Instructional Planning. *Journal of Special Education Technology*, 12, 49-62.
- Furlong, M. J., & Christenson, S. L. (2008). Engaging students at schools and with learning: A relevant construct for all students. *Psychology in the Schools*, 45(5), 365-368.
- Greenwood, C. R., & Terry, B. (1994). Confirming a Performance-Based Instructional Model. *School Psychology Review*, 23(4), 652-669.
- Gregory, A., & Thompson, A. R. (2010). African American high school students and variability in behavior across classrooms. *Journal of Community Psychology*, 386-402.
- Gregory, A., & Weinstein, R. S. (2008). The discipline gap and African Americans: Defiance or cooperation in the high school classroom. *Journal of School Psychology*, 46, 455-475.
- Gresham, F. M., Cook, C. R., Collins, T., Dart, E., Resetshwane, K., Truelson, E., et al. (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*, 364-379.
- Hintze, J. M., & Matthews, W. J. (2004). The generalizability of systematic direct observations across time and setting: A preliminary investigation of the psychometrics of behavioral observation. *School Psychology Review*, 33(2), 258-270.

- Horner, R. H., Chard, D. J., Boland, J. B., & Good III, R. H. (2006). The use of reading and behavior screening measures to predict nonresponse to school-wide positive behavior support: A longitudinal analysis. *School Psychology Review, 35* (2), 275-291.
- Johnson, W., McGue, M., & Iacono, W. G. (2005). Disruptive behavior and school grades: Genetic and environmental relations in 11-year-olds. *Journal of Educational Psychology, 97*(3), 391-405.
- Kazdin, A. E. (2005). *Parent management training: treatment for oppositional, aggressive, and antisocial behavior in children and adolescents*. Oxford: University Press.
- Leflot, G., Van Lier, P. A., Onghena, P., & Colpin, H. (2010). The role of teacher behavior management in the development of disruptive behaviors: An intervention study with the good behavior game. *Journal of Abnormal Child Psychology, 38*, 869-882.
- Markussen, E., Froseth, M. W., & Sandberg, N. (2011). Reaching for the unreachable: identifying factors predicting early school leaving and non-completion in Norwegian upper secondary education. *Scandinavian Journal of Educational Research, 55*(3), 225-253.
- McIntosh, K. (2005, June). *Academic, behavioral, and functional predictors of chronic problem behavior in elementary grades*. Opgeroepen op July 30, 2011, van ProQuest® Dissertations & Theses: <http://www.proquest.umi.com>
- McKown, C., & Weinstein, R. S. (2008). Teacher expectations, classroom context, and the achievement gap. *Journal of School Psychology, 46*, 235-261.
- Miles, S. B., & Stipek, D. (2006). Contemporaneous and longitudinal associations between social behavior and literacy achievement in a sample of low-income elementary school children. *Child Development, 103*-117.
- Ministerie van Onderwijs, Cultuur en Wetenschap. (2011a, February 2). *Daling voortijdig schoolverlaters, recentelijk en toekomstig*. Opgeroepen op July 26, 2011, van Rijksoverheid: <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2011/02/17/daling-voortijdig-schoolverlaters.html>
- Ministerie van Onderwijs, Cultuur en Wetenschap. (2011b, May 23). *Kamerbrief Actieplannen*. Opgeroepen op August 4, 2011, van Rijksoverheid: <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2011/05/23/brief-actieplannen.html>
- Ministerie van Onderwijs, Cultuur en Wetenschap. (2011c, January 31). *Naar Passend Onderwijs*. Opgeroepen op July 26, 2011, van Rijksoverheid: <http://www.rijksoverheid.nl/documenten-en-publicaties/kamerstukken/2011/01/31/naar-passend-onderwijs.html>

- Ministerie van Onderwijs, Cultuur en Wetenschap. (2011d, April). *Schooluitval voorkomen in Nederland*. Opgeroepen op August 1, 2011, van Aanval op schooluitval: <http://www.aanvalopschooluitval.nl>
- Morgan, P. L., Farkas, G., Tufis, P. A., & Sperling, R. A. (2008). Are reading and behavior problems risk factors for each other? *Journal of Learning Disabilities*, 417-436.
- 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.
- Ramsay, M. C., Reynolds, C. R., & Kamphaus, R. W. (2002). *Essentials of Behavioral Assessment*. New York: John Wiley.
- Riley-Tillman, T. C., Chafouleas, S. M., Sassu, K. A., Chanese, J. M., & Glazer, A. D. (2008). Examining agreement between Direct Behavior Ratings (DBRs) and systematic direct observation data for on-task and disruptive behavior. *Journal of Positive Behavior Interventions*, 10, 136–143.
- Riley-Tillman, T. C., Kalberer, S. M., & Chafouleas, S. M. (2005). Selecting the right tool for the job: A review of behavior monitoring tools used to assess student response-to-intervention. *The California School Psychologist*, 81-91.
- Rijksinstituut voor Volksgezondheid en Milieu. (2006, September 14). *Nationale Atlas Volksgezondheid*. Opgeroepen op October 1, 2011, van RIVM: http://www.rivm.nl/vtv/object_map/o1225n21844.html
- Roeser, R.W., van der Wolf, K., & Strobel, K., R. (2001). On the relation between social-emotional and school functioning during early adolescence. Preliminary finding from Dutch and American samples. *Journal of School Psychology*, 39, 111-139.
- Shapiro, E. S. (2004). *Academic skills problems*. New York: The Guilford Press.
- Shapiro, E. S., & Clemens, N. H. (2005). Conducting systematic direct classroom observations to define school-related problems. In R. Brown-Chidsey, *Assessment for intervention: a problem solving approach* (pp. 175-199). New York: Guilford.
- Shapiro, E., & Kratochwill, T. (Red.). (2000). *Behavioral Assessment in Schools : Theory, Research, and Clinical Foundations* (2d ed.). New York: The Guilford Press.
- Tobin, T. J., & Sugai, G. M. (1999). Using sixth-grade records to predict school violence, chronic discipline problems, and high school outcomes. *Journal of Emotional and Behavioral Disorders*, 7(1), 40-53.
- Verhulst, F.C., Ende, J. van der & Koot, H.M. (1997). *Handleiding voor de Teachers's Report Form (TRF)*. Rotterdam: afd. Kinder- en jeugdpsychiatrie, Sophia Kinderzieken-huis/AZR/ EUR.

- Vile Junod, R. E., DuPaul, G. J., Jitendra, A. K., Volpe, R. J., & Cleary, C. S. (2005). Classroom observations of students with and without ADHD: Differences across types of engagement. *Journal of School Psychology, 2006*(44), 87-104.
- Visscher, A., & Ehren, M. (2011, July 13). *Passend Onderwijs*. Opgeroepen op August 4, 2011, van Rijksoverheid: <http://www.rijksoverheid.nl/onderwerpen/passend-onderwijs/documenten-en-publicaties/rapporten/2011/07/13/de-eenvoud-en-complexiteit-van-opbrengstgericht-werken.html>
- Volpe, R. J., DiPerna, J. C., Hintze, J. M., & Shapiro, E. S. (2005). Observing students in classroom settings: A review of seven coding schemes. *School Psychology Review, 24*(4), 454-474.
- 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, 27*(4).

Appendix

Definities van de gedragingen inclusief voorbeelden:

- ALG werd gescoord wanneer de leerling mondeling, schriftelijk of motorisch reageert op de vraag van de docent. (bv. voorlezen, vragen beantwoorden, schrijven, wijzen of rekenen). De leerling is actief aan het leren.
- AT werd gescoord de leerling gericht met de taak bezig is die opgelegd is door de leraar. De ogen van de leerling zijn op de taak gericht (bv. in zichzelf lezen, luisteren, kijken, etc.).
- NAT werd gescoord wanneer de leerling niet op de taak let die opgelegd is door de leraar. De ogen van de leerling zijn niet gericht op de taak (bv. in de richting van de leraar als de leraar aan het instrueren is; naar de geschikte materialen als de leerling aan academisch werk bezig is).
- SG werd gescoord wanneer de leerling de leeromgeving in de klas verstoort voor zichzelf of dat voor andere leerlingen. Storend gedrag kan één van de volgende gedragingen zijn:
 - o Lawaai: alle geluiden die door de leerling worden gemaakt, die ofwel de leraar of een andere leerling afleidt. Het geluid kan vocaal worden gegenereerd (kletsen, onbegrijpelijke geluiden) of non-verbaal (tikken met een potlood, klikkende vingers, etc.).
 - o Niet op zijn plaats: als de leerling niet op de toegestane plaats aanwezig is.
 - o Lichamelijk contact: zich bemoeiend met de activiteiten van een andere leerling, door het aanraken van hem/haar of zijn/haar eigendom. Schoppen, slaan, duwen, scheuren, breken, beetpakken, worden gecategoriseerd als fysiek contact.