



university of
 groningen

faculty of behavioural
 and social sciences

Oei, ik zie geen groei!

Verschillen in leerkrachtverwachtingen over problematische groei bij sterke versus zwakke leerlingen.

Masterthesis

Judith van Aalsum

S3687392

Masteropleiding orthopedagogiek, Faculteit GMW, Rijksuniversiteit Groningen

Begeleider en eerste beoordelaar: Dr. N. Frans

Tweede beoordelaar: Dr. T. Kretschmer

Juni 2023

Aantal woorden: 11.010

Abstract

Oh no, there's no growth!

Differences in teacher expectations on problematic growth for high and low achieving students.

Attention to individual growth is increasingly important, which leads to a shift in the Dutch education system. The new pupil monitoring system of Cito adapts to this by placing more emphasis on individual growth. However, there is little known about when teachers experience growth to be problematic and if teachers draw different conclusions based on the same data. Based on earlier studies on teacher expectations, it is likely that teacher expectations on growth are different for high and low achieving students. The aim of this study was to investigate the way in which teacher expectations on problematic growth differ for high and low achieving students. Therefore, a qualitative, cross-sectional study was performed using data from eight primary school teachers and teacher counselors (*intern begeleiders*). They were interviewed about their own experiences and their perceptions on four case vignettes. A thematic analysis was performed. Teachers reported that their perception on problematic growth depends on a set of norms and the number of tests. Teachers perceive a test score to be more problematic if the score has been decreasing over an extended period of time. Several factors play a part in this, for example, teachers expected the problematic growth of high achieving students more often to be temporary. They explained the problematic growth as a lack of right lesson content, whereas the problematic growth of low achieving students was mostly due to bad behavior. For further studies on this subject it is recommended to apply more variation in the case vignettes. For example, by adding a high achieving student with negative student characteristics to the case vignettes.

Keywords: Cito, growth, problematic growth, teacher expectations, high and low achieving students

Samenvatting

Aandacht voor individuele groei lijkt een steeds belangrijkere rol te krijgen. Cito speelt hierop in door in hun nieuwe leerlingvolgsysteem meer aandacht te hebben voor individuele groei. Echter, er is nog weinig bekend over wanneer leerkrachten groei als stagnerend ervaren en of leerkrachten verschillende conclusies trekken op basis van dezelfde groeidata. Uit eerdere onderzoeken naar leerkrachtverwachtingen is het aannemelijk dat leerkrachtverwachtingen over groei verschillend zijn voor sterke en zwakke leerlingen. In deze masterthesis is onderzocht op welke manier de verwachtingen van leerkrachten over problematische groei verschillen bij sterke en zwakke leerlingen. Om de onderzoeksvragen te beantwoorden is een kwalitatief, cross-sectioneel onderzoek uitgevoerd, waarbij acht leerkrachten en intern begeleiders van basisscholen zijn geïnterviewd over eigen ervaringen. Daarnaast werden hun percepties op vier casussen bevroegd. Om de onderzoeksvragen te beantwoorden werd een thematische analyse uitgevoerd. Uit de interviews kwam naar voren dat leerkrachten problematische groei bepalen op basis van gestelde normen en het aantal toetsmomenten. Leerkrachten ervaren een toetsscore eerder als problematisch als de score gedurende meerdere toetsmomenten daalt. Verschillende factoren spelen hierbij een rol, bijvoorbeeld bij sterke leerlingen wordt eerder verwacht dat problematische groei tijdelijk is. De verklaring voor problematische groei was bij sterke leerlingen vooral het lesaanbod, terwijl dit bij zwakke leerlingen veelal het gedrag van de leerling was. Voor vervolgonderzoek wordt aangeraden om meer variatie aan te brengen aan de casussen, bijvoorbeeld door een sterke leerling met negatieve leerlingkenmerken toe te voegen.

Kernwoorden: Cito, groei, problematische groei, leerkrachtverwachtingen, sterke en zwakke leerlingen

In Nederland zijn basisscholen verplicht een Leerling- en Onderwijsvolgsysteem (LOVS) te hebben (Rijksoverheid, z.d.; Van der Kleij & Eggen, 2013). Van alle scholen gebruikt 85% het LOVS van Cito en daarmee is het één van de meest gebruikte leerlingvolgsystemen in Nederland (Cito, z.d.-b). Cito ontwikkelt gestandaardiseerde toetsen die genormeerd worden op basis van gegevens uit de landelijke populatie (Hollenberg et al., 2014). De resultaten van deze toetsen worden onder andere uitgedrukt in een vaardigheidsniveau en een vaardigheidsscore. Hoewel het vaardigheidsniveau in de praktijk een veelgebruikte maat is om het niveau van leerlingen te bepalen, kan je hiermee moeilijk signaleren hoeveel een leerling gegroeid is. Aangezien de focus in het basisonderwijs steeds meer op groei komt te liggen, wordt de vaardigheidsscore steeds belangrijker. Hiermee kan je namelijk resultaten tussen leerlingen en tussen toetsen over de tijd vergelijken. Dit is onder meer zinvol, omdat stagnerende groei bij individuele leerlingen hiermee gesignaleerd kan worden. Echter, er is nog weinig bekend over wanneer leerkrachten groei als stagnerend ervaren en of leerkrachten verschillende conclusies trekken op basis van dezelfde groeidata.

De vaardigheidsscores van de Cito LOVS-toetsen geven de prestaties van een leerling op elke toets binnen een vakgebied (bijv. rekenen) weer op één onderliggende vaardigheidsschaal (Cito, 2019; Vlug, 1997). Deze vaardigheidsschaal is gestandaardiseerd waardoor de individuele groei van een leerling binnen een bepaald vakgebied gevolgd kan worden over meerdere jaren. Cito maakt hiervoor gebruik van IRT-modellen, waarbij wordt uitgegaan dat de kans dat een leerling een opgave (item) goed heeft, wordt bepaald door de vaardigheid van een leerling en de moeilijkheidsgraad van een opgave (Verhelst, 1993). De toetsresultaten van verschillende toetsen kunnen met elkaar worden vergeleken, omdat de vaardigheid van een leerling en de moeilijkheid van de items op één vaardigheidsschaal zijn vastgesteld door Cito. Zo kunnen de toetsresultaten van een individuele leerling op de toets *rekenen eind groep 5 (E5)* vergeleken worden met de resultaten van de toets *rekenen midden 6 (M6)*. Daarnaast kunnen individuele toetsresultaten vergeleken worden met leeftijdsgenoten. Als alle vaardigheidsscores op dezelfde vaardigheidsschaal worden geplaatst, is het mogelijk om voor iedere groep op ieder toets moment een landelijk gemiddelde te berekenen. Op basis van de landelijke populatiegemiddelden worden vervolgens de vaardigheidsscores ingedeeld in vijf brede percentielgroepen, de vaardigheidsniveaus (i.e. I-V¹). Iedere percentielgroep representeert 20% van de populatie. Een III-score betekent dat de leerling bij de 20% leerlingen hoort die gemiddeld scoren. Leerlingen met een I-score behoren tot de 20% hoogst scorende leerlingen en leerlingen met een V-score tot de 20% laagst scorende leerlingen (Cito, 2019).

¹ De vaardigheidsniveaus werden eerder ook wel ingedeeld in niveau A-E. Waarbij groep D en E samen de laagst scorende 25% van leerlingen omschrijft en groep C, B en A ieder 25% daarboven. Beide indelingen worden in de praktijk nog gebruikt.

Het is niet verrassend dat leerkrachten de vaardigheidsniveaus als basis nemen voor het oordelen over het niveau en de groei van een leerling, omdat het gebruik van vaardigheidsniveaus dermate geïntegreerd is in het onderwijs dat de verschillende niveaus zelfs als criteria dienen voor diagnostiek, bijvoorbeeld een dyslexieonderzoek (De Jong et al., 2016). Een leerling moet op drie opeenvolgende meetmomenten een V-score halen (bij de 10% laagst scorende leerlingen behoren) als vereiste om een dergelijk onderzoek aan te mogen vragen. Toch vindt er langzaam een verschuiving plaats binnen het onderwijs waarin groeiscoringen een belangrijkere rol lijken te krijgen. Cito ontwikkelt momenteel een nieuwe generatie toetsen die meer gericht zijn op de groei die een leerling doormaakt (Cito, z.d.-a). In het protocol Ernstige RekenWiskunde-problemen en Dyscalculie (ERWD) wordt niet van vaardigheidsniveaus gesproken, maar van de vooruitgang in vaardigheidsscores. Stagnerende vaardigheidsscores in het LOVS en onvoldoende vooruitgang kan aanduiden dat een leerling in aanmerking komt voor een dyscalculieonderzoek (Groenestijn et al., 2011). Individuele groei lijkt steeds relevanter te worden, waarbij de aandacht voor niveau lijkt te verzwakken. Het is van belang om te weten wanneer leerkrachten groei als stagnerend ervaren, omdat de perceptie van leerkrachten over stagnerende groei gevolgen kan hebben voor het handelen van de leerkracht in de klas. Leerkrachten die stagnerende groei signaleren, zullen hun instructie kunnen aanpassen aan de behoeften van de leerling en extra instructie bieden aan leerlingen die dat nodig hebben (Stecher, 2002).

Ondanks deze verschuiving worden vaardigheidsniveaus nog vaak als uitgangspunt genomen om stagnatie te signaleren. Toch zijn ze minder geschikt om stagnatie in groei vast te stellen. Doordat de vaardigheidsscores bij vaardigheidsniveaus zijn ingedeeld in brede percentielgroepen, kunnen de niveaus een verkeerd beeld geven over de groei van een leerling. Een leerling met een vaardigheidsscore die net op de grens van twee vaardigheidsniveaus ligt, kan met een enigszins onder gemiddelde groei een vaardigheidsniveau dalen. Dit kan onterecht het beeld geven dat een leerling stagneert of sterk terugloopt in vaardigheid. Op basis van de vaardigheidsscores zou een leerkracht zich wellicht geen zorgen maken, maar door een daling in vaardigheidsniveau kunnen dezelfde vaardigheidsscores al snel tot zorgen leiden. Daarnaast kan een leerling op twee opeenvolgende toetsen hetzelfde vaardigheidsniveau behouden, maar wel bovengemiddeld gegroeid zijn als gekeken wordt naar de vaardigheidsscore. Individuele groei wordt dan verhuld, waardoor mogelijk ten onrechte geconcludeerd kan worden dat een leerling niet groeit (Frans et al., 2020). Het is van belang dat er meer inzicht is in de rol die vaardigheidsniveaus en -scores spelen bij het beoordelen en interpreteren van Citoresultaten.

De verwachtingen van leerkrachten spelen een belangrijke rol bij het interpreteren van toetsscores en het bepalen van problematische groei. Leerkrachtverwachtingen worden gedefinieerd

als de conclusies die leerkrachten trekken over het potentieel dat hun leerlingen kunnen bereiken in de klas (Riley & Ungenleider, 2012). Het is belangrijk om te begrijpen waar leerkrachtverwachtingen over groei vandaan komen, aangezien leerkrachtverwachtingen sterk verbonden zijn met de prestaties van leerlingen (Brault et al., 2014; Timmermans et al., 2016; Wang et al., 2021). Er zijn aanwijzingen dat verschillende leerkrachtkenmerken de verwachtingen van leerkrachten over de academische prestaties van leerlingen vormen. Uit het onderzoek van Aydin en Ok (2022) blijkt dat de zelfeffectiviteit van leerkrachten van invloed is op hun verwachtingen over niveau en groei. Zelfeffectiviteit gaat over het vertrouwen dat leerkrachten hebben in hun eigen bekwaamheid om succesvol hun vak uit te voeren (Heslin & Klehe, 2006). Geslacht, jaren ervaring als leerkracht en de genoten opleiding zijn factoren die van invloed zijn op de zelfeffectiviteit van leerkrachten (Cambay & Kaznaç, 2021) en daarmee ook op de verwachtingen die leerkrachten hebben. Onderzoek van Wang et al. (2018) liet bijvoorbeeld zien dat meer ervaren leerkrachten vaak lagere verwachtingen hebben van hun leerlingen. Er zou dus verondersteld kunnen worden dat geslacht, jaren ervaring en genoten opleiding van invloed zijn op de verwachtingen van leerkrachten over de prestaties van leerlingen.

Naast leerkrachtkenmerken zijn er ook leerlingkenmerken, zoals eerdere resultaten en motivatie, die de verwachtingen van leerkrachten over niveau beïnvloeden. Aangezien de percepties over niveau en groei in de praktijk dicht bij elkaar liggen, zijn de factoren die van invloed zijn op de leerkrachtverwachtingen over niveau mogelijk ook van invloed op de leerkrachtverwachtingen over groei. Eerdere prestaties en behaalde resultaten hebben een grote invloed op de leerkrachtverwachtingen over niveau (Riley & Ungerleider, 2012). In het onderzoek van Timmermans et al. (2015) werd 80% van de leerkrachtverwachtingen op leerlingniveau verklaard door eerdere prestaties. Dit betekent dat eerdere prestaties voor een groot deel bepalen wat de leerkracht verwacht over het niveau van de leerlingen. Daarnaast spelen ook de motivatie en werkhouding van een leerling mee in de verwachtingen van leerkrachten (Sneyers et al., 2019; Timmermans et al., 2016). Een positieve werkhouding, zoals geconcentreerd kunnen werken en motivatie in de klas, zorgt voor hogere verwachtingen bij leerkrachten (Timmermans et al., 2016). Naast werkhouding blijkt dat leerkrachten hogere verwachtingen hebben van meisjes ten opzichte van jongens (Timmermans et al., 2016; Wang & Cai, 2016; Wang et al., 2021). Er zijn verschillende bevindingen wat betreft sociaal gedrag. Uit het onderzoek van Timmermans et al. (2016) blijkt dat leerkrachten lagere verwachtingen hebben van leerlingen die meer positief sociaal gedrag vertonen in de klas. Uit ander onderzoek blijkt echter het omgekeerde (Hecht & Greenfield, 2002). Hoewel deze leerlingkenmerken vooral gericht zijn op de verwachtingen over niveau, spelen ze mogelijk net zo zeer een rol bij leerkrachtverwachtingen over groei.

Ondanks dat er nog weinig onderzoek is gedaan naar leerkrachtverwachtingen over groei, liet het onderzoek van Frans et al. (2020) zien dat leerkrachten een gemiddelde groei in vaardigheidsniveau eerder als problematisch ervaren bij laag presterende leerlingen. Twee opeenvolgende V-scores werden door alle leerkrachten in dit onderzoek als zeer problematisch gezien en gaven, onterecht, het idee dat een leerling onvoldoende groeit (Frans et al., 2020). Het is echter nog onduidelijk of leerkrachten dit ook ervaren als wordt gekeken naar vaardigheidsscores, aangezien dat onderzoek niet direct gericht was op de interpretatie van groei en niet heeft gekeken naar de vaardigheidsscores. Het is wel aannemelijk dat de verwachtingen van leerkrachten over groei verschillend zijn voor sterke en zwakke leerlingen. Leerlingen die bovengemiddeld presteren ervaren namelijk over het algemeen hogere verwachtingen van de leerkracht dan zwakke leerlingen (Weinstein, 2022). Ze krijgen vaak meer aandacht, meer aanzien en meer positieve reacties op hun inspanningen dan lager presterende leerlingen in dezelfde klas (Wang & Cai, 2016; Weinstein, 2022). Zwakke leerlingen krijgen meer structuur geboden en meer leerkracht gestuurde behandeling dan sterke leerlingen (Weinstein, 2022). Een mogelijke verklaring hiervoor kan zijn dat leerkrachten meestal groepsonderwijs geven, met veel verschillende niveaus binnen één klas. Dit, en de beperkte tijd, zorgt ervoor dat leerkrachten zich in eerste instantie richten op dat iedereen het niveau van de klas kan bijhouden. Een andere mogelijke verklaring hiervoor kan zijn dat hoge scores gerelateerd zijn aan goed ontwikkelde executieve functies, zoals volgehouden aandacht, taakgerichtheid en reactie-inhibitie (Martínez-Vicente et al., 2019). Deze executieve functies dragen bij aan een positieve leerhouding waardoor leerkrachten hogere verwachtingen hebben van sterke leerlingen. Kortom, verschillende leerlingkenmerken hebben invloed op de leerkrachtverwachtingen met betrekking tot niveau. Het is belangrijk om te weten welke leerlingkenmerken mogelijk een rol spelen met betrekking tot de leerkrachtverwachtingen over (problematische) groei, aangezien groei steeds belangrijker lijkt te worden.

Tot slot hebben ook gezinsomstandigheden en schoolkenmerken invloed op de leerkrachtverwachtingen over de prestaties van leerlingen. Deze kenmerken zouden kunnen worden gezien als indirecte leerlingkenmerken. Aangezien deze kenmerken van invloed zijn op de verwachtingen van leerkrachten over de prestaties van leerlingen, hebben ze mogelijk ook invloed op de verwachtingen over groei. Leerkrachten verwachten dat leerlingen een grotere kans op succes hebben als ouders thuis ondersteuning bieden, bijvoorbeeld bij huiswerk (Riley & Ungerleider, 2012; Sneyers et al., 2019; Timmermans et al., 2015). Ook de etnische achtergrond en sociaaleconomische status van de populatie van een school en de leerling hebben invloed. Leerkrachten hebben hogere verwachtingen van leerlingen uit welvarende gezinnen en zijn in negatieve zin bevooroordeeld als

het gaat over leerlingen uit gezinnen met een lage sociaaleconomische status (Brault et al., 2014; Timmermans et al., 2015).

Het doel van huidig onderzoek is onderzoeken hoe de verwachtingen van leerkrachten over groei verschillen tussen sterke en zwakke leerlingen. Hierbij ligt de focus op problematische groei. De onderzoeksvraag die hieruit voortvloeit is: “op wat voor manier verschillen de verwachtingen van leerkrachten over problematische groei bij sterke leerlingen ten opzichte van zwakke leerlingen?” Deze onderzoeksvraag wordt beantwoord met behulp van twee deelvragen:

1. Wanneer ervaren leerkrachten groei als problematisch?

2. Op wat voor manier zijn de verschillen in verwachtingen van leerkrachten over groei te verklaren vanuit leerlingkenmerken?

Voor deze thesis zijn drie verwachtingen geformuleerd. Er wordt verwacht dat 1) leerkrachten groei als problematisch beoordelen op basis van vaardigheidsniveaus, waarbij meermaals een laag vaardigheidsniveau als problematische groei wordt gezien. Daarnaast wordt verwacht dat 2) leerkrachten beperkte groei bij leerlingen met hoog vaardigheidsniveau minder snel als problematisch zien dan bij leerlingen met een laag vaardigheidsniveau. Als laatste wordt verwacht dat 3) het toevoegen van leerlingkenmerken (e.g. geslacht, motivatie, werkhouding) bijdraagt aan de verwachtingen van leerkrachten over sterke en zwakke leerlingen.

Methode

Design

Om de onderzoeksvragen te beantwoorden is een kwalitatief, cross-sectioneel onderzoek uitgevoerd. De verwachtingen van de participanten over de prestaties en groei van leerlingen zijn bevraagd middels interviews. Hierbij werd gekeken naar leerlingkenmerken die mogelijk van invloed waren op deze verwachting. Specifiek gericht op zwak en sterk presterende leerlingen.

Populatie en steekproef

De doelgroep van dit onderzoek waren leerkrachten en intern begeleiders die werkzaam waren in het reguliere basisonderwijs in Nederland. De participant moest in het bezit zijn van een diploma tot leerkracht en op het moment van het interview werkzaam zijn op een reguliere basisschool. Als laatste criterium moest de school gebruik maken van het LOVS van Cito. De participanten werden geworven middels een doelgerichte steekproef. De participanten werden geselecteerd om zoveel mogelijk variatie op specifieke kernmerken te creëren. Er waren aanwijzingen dat de volgende leerkrachtkenmerken van invloed konden zijn: geslacht, jaren ervaring en genoten opleiding (Aydin & Ok, 2022; Çambay & Kazanç, 2021; Luján, 2021). Hoewel omgevingskenmerken geen deel uitmaakten van dit onderzoek, waren er wel aanwijzingen dat de

populatie en context van een school van invloed waren op de verwachtingen van leerkrachten (Timmermans et al., 2015). Daarom zijn relevante kenmerken van de context van de scholen opgenomen in de resultaten. Er werd beoogd interviews af te nemen bij 8-12 participanten. Met dit aantal participanten wordt gemiddeld genomen saturatie bereikt bij kwalitatief onderzoek (Hennink et al., 2017).

Meetinstrument

Er werd een semigestructureerd interview afgenomen dat bestond uit drie onderdelen. De interviewleidraad is opgenomen in Bijlage 1. Allereerst werden ter bevestiging en inleiding van het interview een aantal basisvragen gesteld, zoals de opleiding(en) die de geïnterviewde genoten had, het aantal jaren ervaring als leerkracht en/of intern begeleider en wat basisinformatie over de locatie van de school.

Het tweede deel van het interview haakte in op de persoonlijke ervaring van de participant en de eigen klas. Ter ondersteuning werden de LOVS leerlingresultaten van de meest recente rekentoets van Cito uit de klas van de participant gebruikt. Intern begeleiders werden gevraagd een groep te kiezen. Er werd gevraagd om welke leerlingen de participant zich op dit moment zorgen maakte wat betreft toetsresultaten, groei in vaardigheidsscore of zorgen om andere aspecten en wat de reden van deze zorgen was. Er werd gepeild welke verwachtingen de participant van deze leerlingen had en waar die op gebaseerd waren. Participanten die geen eigen groepsresultaten mee hadden genomen, kregen algemene vragen over wanneer ze zich zorgen maakten om een leerling en welke factoren een rol speelden in het beoordelen van Citoscores. Hierbij werden ze aangemoedigd om over leerlingen in hun eigen klas na te denken.

Als laatste kreeg de participant vier fictieve, door de onderzoeker ontworpen, casussen te zien (zie Bijlage 2). Alle vier leerlingen uit de fictieve casussen hadden dezelfde, onder gemiddelde, groei in vaardigheidsscore, maar verschilden in niveau (hoog/laag) en achtergrondkenmerken (e.g. werkhouding en thuissituatie). De casussen zijn in vier mogelijke volgorden² gepresenteerd om mogelijke volgorde-effecten te voorkomen. De casussen die 'hoog' worden genoemd, zijn leerlingen met een hoog vaardigheidsniveau op rekenen (I- en II-scores). De casussen die 'laag' worden genoemd, zijn leerlingen met een laag vaardigheidsniveau op rekenen (IV- en V-scores). Allereerst werden er twee leerlingen na elkaar gepresenteerd waarbij één een hoog vaardigheidsniveau en de ander een laag vaardigheidsniveau had. Participanten kregen enkel een lijndiagram met vaardigheidsscores van de leerlingen te zien in een leerlingrapport van het LOVS 3.0 van Cito.

² 1) hoog-laag-hoog-laag, 2) hoog-laag-laag-hoog, 3) laag-hoog-hoog-laag, en 4) laag-hoog-laag-hoog

Daarna kregen de participanten twee leerlingen na elkaar gepresenteerd waarbij de vaardigheidsscore aangevuld werd met aanvullende leerling informatie over relevante leerlingkenmerken die mogelijk van invloed zijn op de verwachtingen van leerkrachten, zoals motivatie, werkhouding en geslacht (Sneyers et al., 2019; Timmermans et al., 2016). Eén leerling had een hoog vaardigheidsniveau en positieve leerlingkenmerken waardoor leerkrachten mogelijk hoge verwachtingen hadden. De andere leerling had een laag vaardigheidsniveau en negatieve leerlingkenmerken waardoor leerkrachten mogelijk lage verwachtingen hadden. De participant werd per rapport gevraagd wat de participant kon zeggen over het niveau van de leerling en de groei die de leerling had doorgemaakt. Ook werd gevraagd welke verwachtingen de participant had over de prestaties van de leerling.

Het interview werd afgesloten met de concrete vraag wanneer de participant groei als problematisch ervaarde. Daarna werd gevraagd of de participant alles heeft kunnen noemen en of de participant nog iets kwijt wilde over Cito.

Procedure

De leerkrachten en intern begeleiders werden benaderd via e-mail, telefoon of direct contact. Er werd binnen het eigen netwerk gezocht naar participanten met bepaalde kenmerken om zoveel mogelijk variatie op specifieke kenmerken te waarborgen. De participanten werden middels een informatiebrief geïnformeerd over het doel van het onderzoek (zie Bijlage 3). Hierin werd de participant ook geïnformeerd over haar rechten en de vertrouwelijkheid van de gegevens. Voorafgaand aan het interview werd schriftelijk actieve toestemming gegeven voor het maken van een geluidsopname en het gebruik van de data voor dit onderzoek (zie Bijlage 4). De participant werd op de hoogte gebracht dat hij/zij op ieder moment gedurende het onderzoek het recht had om het interview en/of de opname te stoppen, zonder hiervoor verantwoording af te hoeven leggen. In geval van beëindiging, kreeg de participant de keuze om de opname die tot dan toe gemaakt was te laten gebruiken of verwijderen. Voor het interview waren leerlingresultaten uit de klas van de participant nodig. De participant werd hierover voorafgaand aan het interview op de hoogte gesteld zodat de participant deze gegevens geanonimiseerd kon meenemen naar het interview. De leerlingresultaten zijn in het beheer van de participanten gebleven. Na afloop werd de participant geïnformeerd over het onderzoek en verder contact, waarna de opname werd gestopt.

De individuele interviews werden afgenomen in februari t/m april en duurden gemiddeld 34 minuten (range 23 tot 44 minuten). Eén interview vond plaats bij de participant thuis, één interview werd online afgenomen en de overige interviews vonden plaats op de school van de participant. Tijdens de interviews werd Nederlands of Fries gesproken, afhankelijk van de wensen van de participant. De Fries-gesproken interviews zijn in het Fries getranscribeerd. Verdere analyses zijn in het Nederlands gedaan. Namen of andere persoonlijke informatie die werden genoemd tijdens de

interviews en herleidbaar waren, zijn geanonimiseerd verwerkt in de transcripten zodat de privacy gewaarborgd kon worden. Direct na afloop van het interview werd de geluidsopname opgeslagen in een beveiligde omgeving van de Rijksuniversiteit Groningen, conform de richtlijnen van de Algemene Verordening Gegevensbescherming (AVG) en het Nieuwenhuis Datamanagement Protocol en verwijderd van de opnameapparatuur. Na afronding van het volledige onderzoek werden de geluidsopnames volledig verwijderd. De transcripten worden voor een periode van 10 jaar bewaard op de Y-schijf. Voor dit onderzoek is toestemming gegeven door De Ethische Commissie PedOn van de Faculteit Gedrags- en Maatschappijwetenschappen van de Rijksuniversiteit Groningen.

Analyse

Om de onderzoeksvragen te beantwoorden werd een thematische analyse uitgevoerd. Hierbij werd gezocht naar uitspraken over vaardigheidsscores, vaardigheidsniveaus en overige uitspraken waaruit problematische groei bleek. Er is ook gezocht naar uitspraken over de verwachtingen van leerkrachten over leerlingkenmerken zoals gedrag, resultaten en leerhouding. Als laatste werden relevante uitspraken over verwachtingen van leerkrachten over problematische groei van sterke en zwakke leerlingen meegenomen in de analyse. Om de opgestelde hypothesen te evalueren is actief gezocht naar uitspraken die gerelateerd zijn aan de onderzoeksvraag over problematische groei en de eerder opgestelde hypothesen ondersteunen of ondermijnen.

De interviews werden door de onderzoeker getranscribeerd en aandachtig nagelezen waarbij werd geprobeerd betekenis aan de inhoud te geven en reeds werd gezocht naar patronen (Braun & Clarke, 2006). In een tweede lezing werden de losse fragmenten die betrekking hadden op de onderzoeksvraag door de onderzoeker open gecodeerd. Het coderen is uitgevoerd met het programma ATLAS.ti 23. Na het open coderen van drie transcripten werden codes samengevoegd die inhoudelijk hetzelfde waren. De transcripten die volgden werden gecodeerd met bestaande codes en er werden nieuwe codes toegevoegd. Hierna volgde een tweede ronde coderen waarbij opnieuw inhoudelijk dezelfde codes werden samengevoegd. Vervolgens heeft de onderzoeker axiaal gecodeerd. Codes die inhoudelijk met elkaar verbonden waren, werden hierbij geclusterd tot potentiële thema's en sub thema's. Vervolgens heeft de onderzoeker per thema de bijbehorende fragmenten grondig gelezen en bepaald of er een samenhangend patroon was tussen de fragmenten. Thema's en sub thema's zijn op basis hiervan waar nodig heringedeeld. De thema's en sub thema's werden tussentijds besproken met een andere onderzoeker en de thesisbegeleider en waar nodig aangepast.

Na afloop van de analyse werd een member-check gedaan door participanten een korte interpretatie van de interviews te mailen (Bijlage 5). Hierbij werd gevraagd of ze zich herkennen in

de genoemde thema's en interpretaties. Hiermee werden de bevindingen van het onderzoek gevalideerd. Tevens is de betrouwbaarheid van het onderzoek beoordeeld met behulp van de interbeoordelaarsbetrouwbaarheid. Hiervoor werd gebruik gemaakt van Cohen's Kappa. Nadat het codeschema volledig was, werd een andere onderzoeker gevraagd om stukken uit interviews te coderen met behulp van het codeschema en werd er naar overeenkomsten gezocht. De interbeoordelaarsbetrouwbaarheid was matig ($\kappa = .26$) (Landis & Koch, 1977). De discrepanties zijn met de andere onderzoeker besproken. Op basis hiervan zijn twee codes hernoemd ter verduidelijking.

Resultaten

Kenmerken participanten en interviews

Er zijn uiteindelijk acht interviews ($N=8$) afgenomen. In Tabel 1 staan de kenmerken van de participanten zoals geslacht, opleiding en jaren ervaring. De steekproef bestond uit 3 mannen (37.5%) en 5 vrouwen (62.5%), waarvan 5 participanten de reguliere PABO hebben gevolgd (62.5%), 2 participanten de academische PABO (25%) en 1 participant de verkorte PABO (12.5%). Er waren 3 startende leerkrachten (37.5%, 0-3 jaar ervaring), 2 gevorderde leerkrachten (25%, 4-15 jaar ervaring) en 3 ervaren leerkrachten (37.5%, >15 jaar ervaring). Participant D werkt momenteel als intern begeleider en heeft ook meer dan 10 jaar als leerkracht gewerkt. Participanten G en H hebben op dit moment geen eigen klas maar werken ambulantly als leerkracht op één (G) of meerdere (H) scholen. Daarnaast hebben zij verschillende functies op basisscholen (gehad), waarvan de meest recente taken zijn opgenomen in Tabel 1.

Tevens wordt de context van de scholen van de participanten genoemd. Twee participanten (25%) geven les op een school in een stad, zes participanten (75%) geven les op een school in een dorp. In Tabel 1 valt te zien dat drie scholen een onder gemiddelde populatie hebben, drie scholen een gemiddeld presterende leerlingpopulatie en twee scholen een populatie hebben met een hoog uitstroomperspectief (Scholen op de kaart, z.d.). Dit houdt in dat veel leerlingen een hoog advies voor vervolgonderwijs krijgen. Leerkracht C en D werken op dezelfde school. Ook leerkracht E en F werken op dezelfde school. In beide gevallen was er sprake van één leerkracht die al minimaal 15 jaar op die school werkte en één leerkracht die pas sinds 2 maanden op de desbetreffende school werkte.

Als laatste zijn kenmerken van de interviews, zoals de gesproken taal, meegenomen resultaten en de volgorde van de casussen opgenomen in Tabel 1. Tijdens drie interviews (37.5%) werd Fries gesproken. Drie participanten (37.5%) hadden geen resultaten meegenomen waardoor dat onderdeel met alternatieve vragen is bevestigd. Eén interview (F) is niet getranscribeerd in verband met een

Tabel 1*Kenmerken participanten.*

Participant	Geslacht	Jaren ervaring	Genoten opleiding	Functie	Context school	Gesproken taal tijdens interview	Groepsresultaten meegenomen	Volgorde casussen*
A	Vrouw	0-3	Academische PABO	Leerkracht bovenbouw	Stadsschool met een hoog uitstroomperspectief	Nederlands	Ja	A
B	Man	0-3	Academische PABO	Leerkracht bovenbouw	Dorpsschool met een gemiddeld presterende populatie	Nederlands	Nee	B
C	Man	0-3	PABO	Leerkracht bovenbouw	Dorpsschool met een onder gemiddeld presterende populatie. Leerkracht werkt pas 2 maanden op deze school.	Fries	Ja	C
D	Vrouw	15+	Verkorte PABO	Intern begeleider	Dorpsschool met een onder gemiddeld presterende populatie.	Fries	Ja	C
E	Vrouw	4-15	PABO	Leerkracht midden- en bovenbouw	Dorpsschool met een gemiddelde populatie. Leerkracht werkt pas 2 maanden op deze school.	Nederlands	Ja	D
F	Man	15+	PABO	Leerkracht midden- en bovenbouw	Dorpsschool met een gemiddelde populatie	Nederlands (Geen opname van interview)	Ja	D
G	Vrouw	15+	PABO	Ambulante leerkracht en hoogbegaafdheids-specialist	Dorpsschool met een hoog uitstroomperspectief	Nederlands	Nee	A
H	Vrouw	4-15	PABO	Ambulante leerkracht en organisator talentonderwijs	Stadsschool in achterstandswijk	Fries	Nee	B

Noot. * A=hoog-laag-hoog-laag, B=hoog-laag-laag-hoog, C=laag-hoog-hoog-laag, D=laag-hoog-laag-hoog

mislukte opname. Hierdoor is gebruik gemaakt van een uitgebreide samenvatting van het interview. De samenvatting is 45 minuten na het interview geschreven door de interviewer. Hiervoor werden de aantekeningen gebruikt die tijdens het interview zijn gemaakt en de uitgebreide aantekeningen die 10 minuten na het interview door de interviewer zijn geschreven.

Categorieën en thema's

Om de onderzoeksvraag te beantwoorden zijn er verschillende thema's gevormd in drie verschillende categorieën, namelijk 1) problematische groei in het algemeen, 2) problematische groei bij sterke leerlingen en 3) problematische groei bij zwakke leerlingen. De categorieën en bijbehorende thema's zijn samengevat en toegelicht in een codeschema (zie Tabel 2). In Bijlage 6 is het codeboek opgenomen.

De eerste categorie gaat over problematische groei bij leerlingen in het algemeen. Hieronder vielen uitspraken over normen voor problematische groei die leerkrachten hanteren, het aantal meetmomenten voordat achterblijvende groei problematisch is, vaardigheidsniveaus in het algemeen en andere resultaten dan het laatste, individuele Citoresultaat. Dit zijn bijvoorbeeld resultaten van eerdere Citotoetsen, Cito groepsresultaten of resultaten van methodetoetsen. Andere thema's die onder deze categorie vielen, waren de eigen ervaringen van de leerkracht over problematische groei, de vergelijking tussen het Citoresultaat en de leerpotentie van leerlingen en het gedrag dat leerlingen in de klas en tijdens de Citotoets laten zien dat van invloed is op de resultaten.

De tweede categorie is toegespitst op problematische groei van sterke leerlingen. Hierbij vielen uitspraken over waarom participanten zich wel of geen zorgen maken om de sterke leerlingen, verwachtingen over de leerling op basis van niveau, gedrag dat leerkrachten verwachten of zien bij sterke leerlingen, en de verwachtingen van leerkrachten over de toekomstige groei van sterke leerlingen. Ook uitspraken over verklaringen voor problematische groei bij sterke leerlingen vielen onder deze categorie.

De laatste categorie betreft de problematische groei van zwakke leerlingen. Hieronder vielen uitspraken over het oordeel van leerkrachten over problematische groei bij zwakke leerlingen, verwachtingen van leerkrachten over de toekomstige groei en de behoeften die zwakke leerlingen hebben om groei in de toekomst te bevorderen. Daarnaast vielen in deze categorie uitspraken over verklaringen voor problematische groei bij zwakke leerlingen.

Tabel 2*Codeschema.*

Categorie	Thema	Omschrijving thema
Verwachtingen over problematische groei bij leerlingen in het algemeen	Normen voor problematische groei	De verschillende normen die participanten hanteren om problematische groei vast te stellen.
	Het aantal meetmomenten	Verwachtingen over problematische groei op basis van het aantal meetmomenten.
	Vaardigheidsniveau van de leerling	Verwachtingen over problematische groei bepaald door het vaardigheidsniveau van een leerling
	Andere resultaten dan het laatste Citoresultaat	Verwachtingen over problematische groei bepaald door andere resultaten dan het laatste Citoresultaat, zoals methoderesultaten en groepsresultaten.
	Eigen ervaring van de leerkracht	Verwachtingen over problematische groei bepaald door de eigen ervaring van leerkrachten
	Problematische groei afhankelijk van leerpotentie	Verwachtingen over problematische groei gebaseerd op de leerpotentie van een leerling
	Score te verklaren door gedrag van kind	Het gedrag van een leerling dat invloed heeft op de Citoresultaten
Verwachtingen over problematische groei van sterke leerlingen	Waarom maken leerkrachten zich geen zorgen?	Redenen waarom leerkrachten zich geen zorgen maken over problematische groei
	Verwachtingen op basis van niveau	Verwachtingen over sterke leerlingen die gebaseerd zijn op het vaardigheidsniveau
	Verwacht score weer omhoog	Verwachtingen over de toekomstige groei van sterke leerlingen
	Verwachtingen over gedrag dat groei stimuleert	Verwachtingen over gedrag van sterke leerlingen waardoor ze groei laten zien op Citoresultaten
	Verklaringen voor problematische groei	Verklaringen, zoals aanbod, waardoor sterke leerlingen problematische groei laten zien
Verwachtingen over problematische groei van zwakke leerlingen	Oordeel over een lage groei bij een zwakke leerling	Verschillende oordelen en zorgen over de problematische groei van een zwakke leerling
	Verwachtingen over toekomstige groei	Verwachtingen over groei die zwakke leerlingen in de toekomst mogelijk wel of niet laten zien
	Meer winst te behalen door beter aansluiten op behoeften van leerling	De behoeften van zwakke leerlingen die de groei mogelijk stimuleren
	Verklaringen voor problematische groei	Verklaringen, zoals gedrag, waardoor zwakke leerlingen problematische groei laten zien

Leerkrachtverwachtingen over problematische groei bij leerlingen in het algemeen

Normen voor problematische groei

Participanten gaven aan dat ze de norm voor problematische groei op verschillende manieren bepalen. Eén participant gaf aan dat zij daarvoor de checklist voor gemiddelde vaardigheidsgroei gebruikt. Dit is een lijst vanuit Cito die aangeeft hoeveel vaardigheidspunten een leerling gemiddeld moet groeien. Problematische groei wordt vervolgens bepaald als een leerling minder in vaardigheid groeit dan staat aangegeven op de checklist. Daarnaast noemde één participant dat problematische groei wordt bepaald door de normen die gesteld zijn door de school of het schoolbestuur. Dit kan ook samenhangen met de populatie van de school. Participant H werkt op verschillende scholen met verschillende populaties. Ze gaf daarom aan dat scholen met een allochtone leerlingpopulatie wellicht andere normen stellen dan scholen met een autochtone leerlingpopulatie.

Het hangt namelijk ook heel erg van de populatie af wat je van een kind mag verwachten aan groei. Je kan je bijvoorbeeld bij een begrijpend lezen toets voorstellen dat een kind op een gegeven moment stagneert omdat het thuis een andere taal spreekt. Je woordenschat zal dan minder groot zijn dan iemand die thuis Nederlands spreekt. (Participant H, vertaald uit het Fries)

Participanten hanteren verschillende concrete normen voor problematische groei. Alle participanten ($N=8$) zijn het eens dat groei problematisch is als een leerling in vaardigheidsscore achteruitgaat. *“Gaat hij [red. de vaardigheidsscore] naar beneden, dan gaan bij mij wel alle alarmbellen af, zo van: hey, dit is wel echt iets.”* (Participant H, vertaald uit het Fries). Daarnaast werd door twee participanten aangehaald dat groei problematisch is als de vaardigheidsscore gelijk blijft, omdat een leerling dan niet groeit. De helft van de participanten gaven aan dat een vaardigheidsgroei van minder dan vier punten op de Cito rekenen-wiskunde ook problematisch is. *“Zoals hier, is het maar een groei van een vaardigheidspunt van één. Ja dat is natuurlijk heel erg weinig.”* (Participant G). Twee participanten (25%) nemen ook het leerrendement³ mee en noemden dat een leerrendement lager dan 80% problematisch is.

Het aantal meetmomenten

De verwachtingen van participanten over problematische groei worden ook gebaseerd op het aantal meetmomenten. Hoewel twee participanten van mening zijn dat er na één moment geringe groei sprake is van een problematische groei, zijn vier participanten (50%) het eens dat er minimaal

³ Het leerrendement vertaalt de vaardigheidsscore naar een didactische leeftijd die gemiddeld hoort bij dat niveau (didactische leeftijdsequivalent) en deelt dit door de werkelijke didactische leeftijd (het aantal maanden genoten onderwijs vanaf groep 3) van de leerling (Anstadt, 2022).

twee momenten van geringe groei (na elkaar) moet zijn voordat het problematisch wordt. Participant G zei hierover het volgende:

“Daar [red. problematisch vinden] zou je dan toch twee of drie [toetsmomenten] echt wel voor nodig [hebben], dat je denkt van ja (...) want anders zou het betekenen dat je als kind altijd maar moet pieken en dat kan gewoon niet. En Cito blijft ook gewoon een momentopname” (Participant G)

Nog twee participanten sloten aan bij de overtuiging dat je na één keer geringe groei tussen twee toetsmomenten nog niet kan concluderen dat een groei problematisch is, omdat Cito een momentopname is.

Ik denk op het moment dat een kind één keer weinig groeit in vaardigheidsscore, dat je niet gelijk het over problematische groei hebt. Want ik vind Cito echt een momentopname. Op het moment dat een leerling eigenlijk in een paar keer, eigenlijk dus minstens twee keer achter elkaar, heel weinig groei laat zien, dat is denk ik het moment dat je je al wel zorgen moet gaan maken. En ja, wanneer dat dan werkelijk problematisch is? Ja eigenlijk vind ik na twee keer al, dan heb je het wel echt over een stagnering. (Participant A)

Vaardigheidsniveau van de leerling

Bij zes participanten (75%) worden vaardigheidsniveaus expliciet genoemd als het gaat over problematische groei. Hierbij vinden ze het allemaal belangrijk dat een leerling binnen zijn eigen niveau groeit, ongeacht wat dat niveau is. Participant A lichtte dit toe met een voorbeeld uit haar eigen klas:

Ik vind het niet per se heel zorgelijk. Bijvoorbeeld een leerling, die scoort een IV-score en nog een IV-score en weer een IV-score. Die gaat wel van 232, naar 248, naar 259. Die laat wel elke Cito weer groei zien. (Participant A)

Het dalen in niveau werd dan ook problematisch genoemd door twee participanten. *“Als een leerling bij wijze van een II-score heeft (...) maar hij eindigt met een III-score, dat is niet oke”* (Participant D, vertaald uit het Fries). Twee andere participanten vonden een verschil van twee vaardigheidsniveaus pas echt opvallend. *“Dat valt mij wel op, als je van een I-score, dus de hoogst scorende 20% bent, en je zakt ineens naar een III-score, zit je ineens in de middelste 20%. Die vind ik wel opvallen”* (Participant A).

Andere resultaten dan het laatste Citoresultaat

De participanten gaven allemaal aan andere resultaten dan het laatste Citoresultaat mee te nemen in hun overweging of groei problematisch is. Eén van de participanten benoemde bijvoorbeeld dat het problematisch is als een leerling goed scoort op methodelessen maar niet goed scoort op de Citotoets. Zes andere participanten (75%) gaven ook aan de resultaten van

methodelessen en methodetoetsen mee te nemen. Daarnaast vergelijken drie participanten (37.5%) het laatste Citoresultaat met eerdere Citoresultaten van dezelfde leerling. Het vergelijken met eerdere Citoscores kan ertoe leiden dat een participant groei eerder of juist minder snel problematisch vindt. *“Maar als je dan kijkt naar de voorgaande scores, als je niet alleen vergelijkt met de laatste score (...) dan zie je wel dat de scores eigenlijk over langere termijn niet per se heel opvallend zijn”* (Participant A). Ook worden de Citoresultaten van verschillende vakken met elkaar vergeleken. *“Leerlingen die bijvoorbeeld rekenen en spelling wel heel hoog hebben, maar bijvoorbeeld begrijpend lezen wat lager is, dat je kijkt waar kan dat aan liggen?”* (Participant G). Hoog scoren, op één vakgebied na, is dus opvallend en mogelijk problematisch volgens deze participant. De helft van de participanten gaven aan zich zorgen te maken als een leerling geringe groei laat zien bij meerdere vakken. *“Eén toets een keer verzieken, dat kan gebeuren. Maar als je ziet dat alle toetsen gewoon tegenvallen dan gaan we ons wel zorgen maken.”* (Participant B). Als laatste kunnen ook groepsresultaten van belang zijn bij het beoordelen van individuele Citoscores. Participant D vertelde dat als meerdere leerlingen uit dezelfde groep weinig vaardigheidsgroei op dezelfde Citotoets laten zien, dat de participant zich geen zorgen zou maken om de leerling maar juist over de leerkracht en het geboden onderwijs.

Problematische groei afhankelijk van leerpotentie

Twee participanten (van dezelfde school) gebruikten een potentietest om te kijken wat een leerling in potentie zou moeten kunnen. Een potentietest toetst wat van een leerling in potentie, op basis van IQ, verwacht mag worden (Boomtest onderwijs, z.d.). De Citoresultaten van een leerling zijn soms lager dan de cognitieve mogelijkheden van een leerling. De potentie van een leerling wijkt dan af van de Citoscore, wat als problematisch kan worden gezien. Participant C lichtte dit toe met een voorbeeld uit zijn eigen klas:

“In vergelijking weer deze toets [red. Citotoets] met de potentietest. Kwam deze [leerling] dus eigenlijk veel lager uit op de Citotoets dan op de potentietest, dan wat er eigenlijk in haar zit. En dat is wel een zorg dingetje waar je ziet deze [leerling] stagneert dus eigenlijk op Cito best wel.”
(Participant C, vertaald uit het Fries)

Vijf andere participanten betrekken ook de leerpotentie van een leerling bij het beoordelen van problematische groei. Zij maken hiervoor geen gebruik van een potentietest, maar maken zich wel zorgen als een leerling lager scoort dan ze verwachten. Dit gebeurt bijvoorbeeld als de groei niet overeenkomt met de groei van voorgaande Citotoetsen. *“Als de groei niet overeenkomt met de rest van de lijn (...) dat vind ik dan opvallend. Als het niet in lijn is met de vorige jaren.”* (Participant E).

Eigen ervaring van de leerkracht

Een andere factor die meespeelt in de verwachtingen van drie participanten (37.5%) is de eigen ervaring. Participant G zei bijvoorbeeld bij het bekijken van één van de casussen *“een klein dipje in midden 6. Dat zien we wel vaker, dat in groep 6 het [red. de vaardigheidsgroei] wat omlaaggaat. Dan ga je meer van het inzicht verwachten.”* (Participant G). Op basis van eigen ervaring weten sommige leerkrachten dat resultaten op een bepaalde Cito-toets vaker tegenvallen, maar ook vaak weer omhooggaan. Zo'n trend kan dan een verklaring zijn voor tegenvallende resultaten, waardoor een leerkracht zich minder snel zorgen maakt, aangezien het ieder jaar terugkomt. *“Dat zie je soms wel, vooral in groep 5 met de rekentoets. De meesten gaan in het midden van groep 5 allemaal onderuit. En ik weet nu dat dat iets is wat uit zichzelf wel weer bijtrekt”* (Participant H, vertaald uit het Fries).

Gedrag

Alle participanten nemen kind kenmerken en het gedrag van een leerling mee in hun verwachtingen over problematische groei. Participant H gaf bij de casus zonder aanvullende leerling informatie aan dat haar zorgen afhangen van het gedrag van de leerling. *“Of ik me zorgen maak hangt echt helemaal af van de leerling. Dat kan ik niet uit deze grafiek halen”* (Participant H, vertaald uit het Fries). Ze noemde dat als een leerling bijvoorbeeld zijn best doet, oplet tijdens instructies en zich goed inzet, dat ze de leerling zou complimenteren over zijn resultaten. Echter, als dezelfde leerling geen aandacht heeft voor de les en geen gebruik maakt van de mogelijkheden, dat ze zich dan wel zorgen zou maken. Participanten gaven ook aan dat een leerling kan stagneren door persoonlijke ontwikkeling of een diagnose, zoals ADHD. Daarnaast kijken vijf participanten (67.5%) naar hoe de leerling in de klas is. Eén participant noemde dat een leerling tijdens de les niet aangeeft dat de lesstof moeilijk is. Een andere participant viel het op dat als een leerling niet oplet tijdens de les, dit zichtbaar is in tegenvallende resultaten. Ook werkhouding (25%), inzet (37.5%) en zelfvertrouwen en concentratie (25%) hebben invloed op de resultaten volgens de participanten. *“Als een kind gemotiveerd is om te leren, dan kan hij sowieso veel meer. Dan kan je er veel meer uithalen”* (Participant E). Naast algemeen gedrag in de klas, kan ook de werkhouding tijdens de Cito-toets afname een verklaring zijn voor de score, volgens vier participanten (50%). *“Ze [red. de leerlingen] kunnen veel beter en ze doen het dan niet. En dat is wel het stukje nauwkeurigheid (...). Dat is wel de grote verklaring bij veel tegenvallende resultaten.”* (Participant B). Ook blokkeren of een slechte inzet tijdens de Citoafname kan een verklaring zijn voor een problematische groei. *“...leerlingen waarvan wij denken: hey maar jij kan veel beter. En waar ligt het dan aan? En vaak is het dan een kwestie van werkhouding geweest tijdens de Cito-toets.”* (Participant B)

Leerkrachtverwachtingen over problematische groei bij sterke leerlingen

Drie participanten (37.5%) gaven aan dat zij zich geen zorgen maakten om de sterke leerling uit de casus zonder aanvullende leerling informatie. De redenen hiervoor waren “*omdat er [red. de vaardigheidsscore] een stijgende lijn in zit*” (Participant G) en “*omdat hij [red. de leerling] een heel constant beeld laat zien*” (Participant H, vertaald uit het Fries). Daarnaast gaf één participant, na het horen van aanvullende leerling informatie, aan dat minder groei niet erg was, omdat de Cito een momentopname is. Eén participant maakte zich geen zorgen, tenzij de daling in scores ook in de reguliere methodelessen te zien was. Dan zou de participant eerder adviseren om extra ondersteuning aan te bieden aan de leerling.

Op basis van niveau werd door één participant genoemd dat de participant minder snel de toets zou analyseren omdat de leerling een I-score heeft gehaald. “*Ik zou bij Robin [red. zwakke leerling] eerder dan bij Jip [red. sterke leerling] nu even een analyse gaan doen (...) omdat Jip nog gewoon goed in de I zit.*” (Participant G). Twee andere participanten gaven ook aan dat zij zich geen zorgen maakten, omdat de leerling een I-score heeft. In de casus met aanvullende leerling informatie was de leerling in het laatste toets moment een niveau gezakt. Drie participanten (37.5%) gaven aan dat ze dit niet erg vonden, met name omdat één keer iets minder groei wel vaker gebeurt. Een andere participant gaf bij de casus zonder aanvullende leerling informatie aan dat dalen in niveau niet erg was, omdat de leerling geen grote pieken en dalen in de vaardigheidsscores zitten, maar de daling in niveau bij de leerling met aanvullende leerling informatie vond deze participant wel erg, omdat de vaardigheidsgroei eerder gelijkmatig was en tijdens het laatste toetsmoment achteruitgegaan is. De participant zou dan ook gaan analyseren waar de lage vaardigheidsgroei vandaan komt, om te zorgen dat de daling niet verder doorzet bij toekomstige toetsmomenten.

Alle participanten signaleerden de lage vaardigheidsgroei bij de sterke leerlingen. Drie participanten (37.5%) gaven aan dat ze verwachten dat de scores wel weer omhooggingen. Eén participant benoemde dat de score weer omhoog zou gaan, mits de leerling extra ondersteuning krijgt. Een andere participant gaf aan dat de leerling gemotiveerd moest worden, waarna de score weer omhoog zou gaan. Bij de casus met aanvullende leerling informatie speelt volgens zes participanten (75%) ook specifiek gedrag mee in hun verwachtingen over groei. Vier participanten (50%) benoemden de motivatie van de leerling als stimulerende factor voor groei. Daarnaast werd ook een goede werkhouding genoemd. “*Maar als ik zie dat zij tijdens de les goed meedoet en dat zij tijdens de Cito zelf ook gewoon een redelijke taakhouding houdt, dan zou ik niet weten waar die daling vandaan zou moeten komen.*” (Participant B). Onder een goede werkhouding valt ook een leerling die het graag goed wilt doen, zijn of haar vinger opsteekt en om hulp durft te vragen.

Als laatste gaven participanten verschillende verklaringen voor de problematische groei van de sterke leerlingen in de casussen en van hun eigen sterke leerlingen. Vijf participanten (62.5%) gaven aan dat een sterke leerling uitdaging nodig heeft. Een gebrek aan uitdaging kan een mogelijke verklaring zijn voor een problematische groei.

Ik denk dat het hier fout is gegaan, dat hij hier de verrijking niet heeft gehad. Hij is hier gewoon in het lesaanbod van de klas meegegaan en daardoor heeft hij de verrijking niet gekregen waardoor hij die sprong niet kan maken. (Participant C, vertaald uit het Fries)

Daarnaast worden ook andere factoren genoemd die met het aanbod te maken hebben, zoals de lesstof gaat te snel (25%) of wordt moeilijker (12.5%). Eén participant gaf aan dat basisvaardigheden een onderliggend probleem kunnen zijn voor de problematische groei. Een andere participant benoemde dat de lesstof wellicht onvoldoende geautomatiseerd is, ofwel, de lesstof is onvoldoende ingeoeffend waardoor het lang duurt voordat een leerling kennis of vaardigheden kan ophalen.

Andere dingen die nog gezegd waren als mogelijke verklaring voor problematische groei, zijn dat de leerling mogelijk aan het plafond van zijn kunnen zit, waardoor er geen groei meer plaatsvindt. *“Dit is een hele sterke leerling, maar die zit al op haar einde denk ik. Dus dan zie je die groei niet meer.”* (Participant D, vertaald uit het Fries). Verder werd genoemd dat sterke leerlingen de toetsvragen van Cito wellicht anders interpreteren, waardoor ze de vragen fout maken, wat invloed heeft op hun groei. Als laatste werden ook executieve functies als mogelijke verklaring voor problematische groei van sterke leerlingen genoemd.

Leerkrachtverwachtingen over problematische groei bij zwakke leerlingen

De participanten hadden verschillende meningen over de lage vaardigheidsgroei van zwakke leerlingen. Drie participanten (37.5%) maakten zich zorgen om de groei van de zwakke leerling, waarvan twee participanten aangaven dat ze verwachtten dat de leerling in potentie hoger zou kunnen scoren. *“Als ik puur kijk naar zijn scores dan denk ik ook er zit meer in. Want hij scoort toch vaak wel op die grens tussen IV en V, terwijl hij nu wel echt een lage V-score laat zien.”* (Participant A). De andere participant maakte zich vooral zorgen dat de leerling een niveau gezakt was. Twee andere participanten gaven juist aan dat ze vonden dat de leerling mooi groeide binnen zijn eigen niveau. Ze verwachtten dat de inzet van de leerling van de casus zonder aanvullende leerling informatie goed was geweest, wat volgens hun zichtbaar was in de resultaten. *“Robin doet volgens mij knetterhard zijn best voor rekenen (...) Hij laat een hele mooie stijgende lijn nog steeds zien. Wel in die IV en V-range, maar hij laat wel groei zien.”* (Participant G). Deze participant deed bij de casus met aanvullende leerling informatie een soortgelijke uitspraak. De participant benoemde dat als de randvoorwaarden, zoals werkhouding en motivatie, goed zijn, dat dit niveau van de leerling

het hoogst haalbare was. Over de toekomstige groei van de zwakke leerlingen deed slechts één participant uitspraken. De participant verwachtte weer groei te zien bij het volgende Citomoment, maar zeker geen III-score. Daarnaast moest de zwakke leerling wel in de gaten worden gehouden want verder zakken in vaardigheidsscore was niet wenselijk.

Vijf participanten (62.5%) gaven aan dat ze dachten dat er meer winst in resultaten te behalen is als er beter wordt aangesloten op de behoeften van de zwakke leerling. De meeste genoemde behoefte bij de zwakke leerling met aanvullende leerling informatie was dat hij wellicht meer gebaat zou zijn bij het inkorten van de hoeveelheid opdrachten. *“Dat je bijvoorbeeld behapbare stukjes [aanbiedt] omdat hij het ook lastig vindt om binnen de tijd dingen af te krijgen. Doe dan iets minder maar zorg dan dat die dingen goed zijn.”* (Leerkracht E). Ook zou de leerkracht meer betekenis kunnen geven aan rekenen en de leerling meer motiveren, om op die manier beter aan te sluiten. Andere zaken die zwakke leerlingen volgens vier verschillende participanten nodig hebben, zijn: concreet materiaal, visuele ondersteuning, het kunnen oefenen van moeilijke lesstof en de leerling heeft het nodig om meegenomen te worden in een verlengde instructie.

Als laatste werden verschillende mogelijke verklaringen genoemd voor de problematische groei bij zwakke leerlingen. Vrijwel alle verklaringen werden gegeven op basis van de casus met aanvullende leerling informatie. Vooral gedrag blijkt een belangrijke invloed te hebben op de resultaten, volgens zeven participanten (87.5%). Gedrag dat genoemd werd, was bijvoorbeeld motivatie. *“Ik denk dat het [red. weinig groei] misschien een motivatiedingetje is, dat hij niet zo goed door heeft waar hij het allemaal voor doet.”* (Participant H, vertaald uit het Fries). Twee andere participanten gaven ook aan dat als de zwakke leerling gedemotiveerd was, dit invloed had op de resultaten. Ook een slechte inzet en werkhouding, zoals het snel en slordig werken, werden als verklaring voor problematische groei gegeven. Ander gedrag dat volgens de participanten van invloed is op de resultaten is weinig concentratie, aandacht en aanleg.

Naast verklaringen vanuit gedrag, werden ook andere verklaringen genoemd door participanten, die mogelijk van invloed zijn op problematische groei bij zwakke leerlingen. Ook deze verklaringen werden genoemd op basis van de casus met aanvullende leerling informatie. Twee participanten benoemden dat de lengte van de Citotoetsen een hiaat kon zijn voor de leerling.

“Tijdens zo ’n Cito is de hele klas vaak toch wel redelijk taakgericht. Maar inderdaad, dat zou wel lastig zijn als hij echt helemaal op zichzelf aangewezen is om anderhalf uur lang stil te zitten, zo ’n toets te maken. Zou een hiaat kunnen zijn.” (Participant B)

Een andere participant trok in twijfel of de zwakke leerling toetsbaar is, mede door de gedragsproblemen die bij de casus werden besproken. Andere verklaringen die door drie verschillende participanten werden genoemd, zijn *“hij begint het inderdaad lastiger te krijgen”*

(Participant B), *“Je gaat checken is het misschien ook de talige kant die belemmert? Zijn de sommen bijvoorbeeld verder prima maar is het vooral de talige dingen?”* (Participant E) en tot slot *“vooral zijn executieve functies ontbreken”* (Participant C, vertaald uit het Fries).

Discussie

In deze thesis is onderzoek gedaan naar de verschillen in leerkrachtverwachtingen over problematische groei bij sterke en zwakke leerlingen. Hiervoor is onderzocht 1) wanneer leerkrachten groei als problematisch ervaren en 2) op wat voor manier de verschillen in verwachtingen van leerkrachten over groei te verklaren zijn vanuit leerlingkenmerken. In dit onderzoek gaven leerkrachten aan dat ze het dalen, gelijk blijven of gering groeien in vaardigheidsscore als problematisch ervaren. Daarnaast noemde de helft van de participanten dat stagnaties in groei altijd over meerdere toetsmomenten worden vastgesteld. Ten tweede ervaren leerkrachten beperkte groei bij leerlingen met een hoog vaardigheidsniveau minder snel als problematisch dan bij leerlingen met een laag vaardigheidsniveau. Leerkrachten leken bij sterke leerlingen eerder te verwachten dat problematische groei tijdelijk was. Zo gaven leerkrachten aan dat vooral het aanbod invloed heeft gehad op een mindere vaardigheidsgroei. Bij zwakke leerlingen werd eerder verwacht dat gedrag van invloed is geweest op de mindere vaardigheidsgroei en verwachtten de leerkrachten dat ze beter moeten aansluiten op de behoeften van een zwakke leerling.

Een verklaring voor problematische groei die door leerkrachten uit dit onderzoek werd genoemd was onder andere dat de Cito LOVS toetsen een momentopname laten zien. Deze veronderstelling lijkt gegrond als gekeken wordt naar de bevindingen van het onderzoek van Frans et al. (2020), waar 60% van hun steekproef op één moment een stagnatie tussen twee toetsmomenten liet zien. Zij suggereerden dan ook dat één moment van stagnatie bij veel leerlingen van tijdelijke aard is. Een leerling moest dus meerdere momenten van geringe groei laten zien voordat een leerkracht dit als problematisch ervaarde. Hiermee hielden leerkrachten rekening met mogelijke vertekening door meetfouten in de Cito-toetsen (van Loon et al., 2015). Toevallige meetfouten kunnen onterecht het beeld geven dat een leerling onvoldoende is gegroeid. Indien onvoldoende groei zich bij een tweede toetsmoment herhaalt, is er sprake van meer consistentie, stabiliteit en precisie van het toetsresultaat van Cito. De kans dat er dan daadwerkelijk sprake is van stagnerende groei is groter, waardoor leerkrachten zich na twee meetmomenten met onvoldoende groei wel zorgen maakten.

Tegen verwachting in werd niet gevonden dat leerkrachten uit dit onderzoek meerdere keren een laag vaardigheidsniveau als problematische groei zagen. Dit is tegenstrijdig met resultaten uit het onderzoek van Frans et al. (2020), waarbij leerkrachten twee opeenvolgende V-scores als

onvoldoende groei en zeer problematisch ervaren. Echter, in dit onderzoek werden vaardigheidsniveaus wel meegenomen door leerkrachten bij het beoordelen van problematische groei. Leerkrachten gaven aan dat het dalen van één of zelfs twee niveaus problematisch is. Opvallend was dat deze uitspraken vooral werden gedaan als leerkrachten over leerlingen uit hun eigen klas spraken. Bij de casus van de sterke leerling met aanvullende leerling informatie werd genoemd dat de daling in niveau (van I naar II) niet erg was. Dit komt overeen met het onderzoek van Frans (2019), waarin leerkrachten een daling van een I naar een II niet als problematisch zagen. Hoewel het dalen van één niveau door sommige leerkrachten te verklaren was, bijvoorbeeld doordat een leerling op de grens van twee vaardigheidsniveaus scoort, duidde een daling van twee vaardigheidsniveaus wel op een forse daling in groei volgens de leerkrachten uit dit onderzoek. Een andere opvallende bevinding was dat twee leerkrachten als enige helemaal geen vaardigheidsniveaus noemden bij het bespreken van de eigen leerlingen en de casussen. Deze twee leerkrachten werkten reeds actief met *Leerling in Beeld*, wat een mogelijke verklaring kan zijn voor deze bevinding, aangezien *Leerling in Beeld* zich meer focust op groei in plaats van niveau.

Ondanks dat alle leerkrachten in dit onderzoek geringe groei in vaardigheidsscore signaleerden bij de casussen, maakten ze zich bij zwakke leerlingen eerder zorgen dan bij sterke leerlingen. Een reden die ze hiervoor aangaven was een laag vaardigheidsniveau en gedrag dat mogelijk van invloed was geweest op de toetsresultaten. Bij sterke leerlingen leidde de geringe groei in de meeste gevallen niet tot zorgen. Een verklaring die leerkrachten onder andere gaven was dat de leerling “hoog scoort”. Een andere verklaring die werd gegeven bij sterke leerlingen, is het bereiken van het leerplafond. Hiermee doelden leerkrachten op het plafondeffect, wat optreedt als een toets te gemakkelijk is voor een leerling (Roberts, 1978). De resultaten op de toets van een leerling laat dan niet de daadwerkelijke vaardigheden van de leerling zien, waardoor de vaardigheidsscores mogelijk minder betrouwbaar zijn (Hollenberg et al., 2014; Roberts, 1978). Een ander statistisch fenomeen dat mogelijk verbonden is aan een dalende groei bij een sterke leerling is ‘regressie naar het gemiddelde’ (Barnett et al., 2005). Dit fenomeen komt voor bij scores die ver van de gemiddelde score afliggen, bijvoorbeeld een hele hoge score. Extreme scores zijn eerder geneigd om een daling te laten zien, omdat de meetfouten gemiddeld genomen groter zijn bij zulke scores. Dit wekt de illusie bij een volgend toetsmoment dat de score daalt. Een leerkracht kan zich dan zorgen maken om de vaardigheidsgroei van een leerling, terwijl de daling in vaardigheidsscore in werkelijkheid door meetfouten komt (Barnett et al., 2005).

De verwachting was dat het toevoegen van leerlingkenmerken (e.g. methoderesultaten, werkhouding, geslacht) bij zou dragen aan de verwachtingen van leerkrachten over sterke en zwakke leerlingen. Bij het bespreken van leerlingen uit de klas van de participanten werden met name

eerdere prestaties en behaalde resultaten genoemd. Dit komt overeen met eerdere bevindingen over de grote invloed van eerdere resultaten op leerkrachtverwachtingen met betrekking tot niveau (Riley & Ungerleider, 2012; Timmermans et al., 2015). Eerder behaalde resultaten, zoals resultaten van methodetoetsen, lijken ook in dit onderzoek een grote invloed te hebben. Bij het bespreken van de casussen met aanvullende leerling informatie leken andere resultaten minder een rol te spelen. Mogelijk komt dit doordat leerkrachten gestuurd werden om ook over andere factoren, zoals gedrag, na te denken.

Een ander leerlingkenmerk dat veelal naar voren kwam was leerlinggedrag. Leerlinggedrag vooral werd besproken bij de casussen met aanvullende leerling informatie, mogelijk doordat participanten door de aanvullende leerling informatie aan het denken werden gezet over de factor gedrag. Gedrag leek bij zwakke leerlingen een grotere rol te spelen dan bij sterke leerlingen. Deze bevinding komt overeen met resultaten uit het onderzoek van Becherer et al. (2021). Hieruit bleek dat externaliserend gedrag via taakgericht gedrag invloed heeft op de academische prestaties van leerlingen. Echter, een belangrijke kanttekening hierbij is dat in dit onderzoek bij zwakke leerlingen vooral negatief gedrag werd toegelicht, terwijl bij sterke leerlingen positief gedrag werd benadrukt. Leerkrachten in dit onderzoek noemden dat goede motivatie en werkhouding er bij de sterke leerling voor zorgden dat de leerkracht zich geen zorgen maakte om de problematische groei. Bij de zwakke leerling werd het negatieve gedrag (e.g. demotivatie, weinig concentratie en een slechte werkhouding) juist vaak als verklaring genoemd voor de problematische groei. Er is echter niet onderzocht of leerkrachten zich ook zorgen zouden maken om een zwakke leerling met positieve gedragskenmerken en een sterke leerling met negatieve gedragskenmerken. Hier ligt een belangrijke, mogelijke verklaring voor het gevonden verschil en tevens een limitatie in het design van dit onderzoek. Het hebben van positieve of negatieve gedragskenmerken hoeft niet altijd een relatie te hebben met hoge of lage resultaten. Een leerling met hoge motivatie en een goede werkhouding kan ook lage resultaten hebben. Eén participant (H) besprak deze situatie tijdens één van de casussen. Zij gaf aan dat als de zwakke leerling zich altijd inzet en een goede werkhouding heeft, dat ze zich geen zorgen zou maken. Zou de leerling met dezelfde score zicht slecht inzetten, dan zou ze zich wel zorgen maken.

Een leerlingkenmerk dat niet expliciet genoemd werd door leerkrachten in dit onderzoek, was geslacht. Hoewel meerdere onderzoeken hebben gevonden dat geslacht geen invloed heeft op academische prestaties, kan het wel zijn dat geslacht invloed heeft gehad op de percepties van de leerkrachten in dit onderzoek (Becherer et al., 2021; Van der Ende et al., 2016). Bij de casussen met aanvullende leerlingkenmerken werd door de leerkrachten eerder de naam van de leerling genoemd dan bij de casussen zonder leerlingkenmerken. Er kan hierdoor mogelijk sprake zijn geweest van

‘genderbias’ waarbij de leerkrachtverwachtingen worden bepaald op basis van hoe leerkrachten verwachten dat jongens en meisjes zich gedragen in de klas. Uit onderzoek blijkt dat hard werken en een hogere kwaliteit van werk eerder wordt toegeschreven aan meisjes dan aan jongens (Timmermans et al., 2016). Toch hoeft het noemen van de naam van de leerling niet te betekenen dat de verwachtingen van leerkrachten beïnvloed zijn door geslacht. Mogelijk vonden sommige leerkrachten het simpelweg gemakkelijker om over ‘Emma’ te spreken dan over ‘deze leerling’, of waren ze hier eerder toe geneigd, doordat de interviewer de namen en het geslacht expliciet noemde bij het bespreken van de aanvullende leerling informatie.

Dit onderzoek had sterke kanten en een aantal limitaties. In dit onderzoeksdesign is gebruik gemaakt van interviews, zodat de motivaties en achterliggende gedachten van leerkrachten over groei uitgebreid konden worden bevraagd. Dit heeft bijgedragen aan de diepgang van het onderzoek. Echter, hierdoor zijn de resultaten zeer vatbaar voor sturing en interpretatie van de onderzoeker. Daarom is tijdens de interviews door de interviewer geprobeerd een balans te zoeken tussen interactie en objectiviteit. De interviewer heeft zich open opgesteld, geprobeerd empathisch te luisteren en zich te identificeren met de participanten. Doordat een interviewer betrokkenheid toont en de dialoog wil aangaan, zal een participant zich meer comfortabel voelen en de antwoorden op vragen beantwoorden door persoonlijke verhalen en voorbeelden (Ellis & Berger, 2001). Aan de andere kant heeft de interviewer bij de casussen juist geprobeerd om een onzijdige houding aan te nemen, aangezien de interviewer de casussen zelf heeft ontworpen. Hiermee wilde de interviewer vermijden dat de achterliggende implicaties van de casussen van invloed waren op de antwoorden van de participanten. Ondanks dit streven, kan volledige objectiviteit van de interviewer niet gegarandeerd worden. Een veelvoorkomend fenomeen bij interviews is dubbele subjectiviteit (Matthews et al., 1989). Dit houdt in dat de interviewer en de participant elkaar beïnvloeden. In de analyse is daarom ook naar objectiviteit gestreefd door open te coderen en hierbij dicht bij de oorspronkelijke uitspraken van participanten te blijven.

Zoals hierboven is genoemd zijn de casussen door de onderzoeker ontworpen. Hierbij zijn de casussen zo ontworpen dat ze zoveel mogelijk lijken op de leerlingrapporten waar leerkrachten bekend mee zijn. Hiermee wordt een realistisch beeld verkregen hoe leerkrachten uit het onderzoek een dergelijk leerlingrapport analyseren. Ook werden de casussen afgewisseld tussen de participanten, om volgorde-effecten te voorkomen. Dit heeft positief bijgedragen aan het onderzoek. Inhoudelijk is geprobeerd om de casussen overeenkomstig, maar niet identiek te laten zijn. De overeenkomst was nodig om casussen met elkaar te vergelijken. Alle casussen hadden een gemiddelde groei tot het laatste toetsmoment, waarbij de vaardigheidsgroei maximaal vier punten was. Om verschillende percepties van participanten uit te lokken, zijn geen identieke situaties

gecreëerd. Hoewel sommige uitspraken bevestigd konden worden aan de hand van deze casussen, zou het toevoegen van meer gevarieerde casussen kunnen bijdragen aan een breder beeld van de percepties van leerkrachten. Op basis van de resultaten van dit onderzoek zouden de casussen aangepast kunnen worden, zodat ze meer variëren op een aantal belangrijke bevindingen. Allereerst zouden de casussen kunnen variëren op het aantal momenten geringe groei. Hier was al minimaal sprake van bij casus Luuk, waarbij de laatste twee momenten vrijwel geen groei lieten zien. Toch had een langere periode van geringe groei explicieter in casussen naar voren kunnen komen. Ten tweede ontbreekt in dit design een sterke leerling met negatieve leerlingkenmerken en een zwakke leerling met positieve leerlingkenmerken. Door deze variatie toe te voegen, had de invloed van vaardigheidsniveaus en gedrag op de verwachtingen van leerkrachten uitgebreider kunnen worden bevraagd.

Er is gestreefd naar maximale variatie in de steekproef door te variëren op de kenmerken geslacht, opleiding en ervaring. Hiermee is geprobeerd een zo veelzijdig mogelijk beeld van opvattingen te krijgen. Hoewel de kenmerken zeer gevarieerd aanwezig waren, is er geen bewijs gevonden dat er een systematisch verschil is met betrekking tot deze kenmerken. Een factor die hierbij mogelijk een rol speelt, is dat de participanten uit het eigen netwerk van de onderzoeker komen. Hierdoor komen de meeste participanten overwegend uit de noordelijke provincies. Uit de interviews bleek dat de context van de school (e.g. uitstroomperspectief, dorp/stad) mogelijk wel invloed heeft gehad op de uitspraken van de participanten. Zo gaf één participant aan dat leerlingen met een lage III-score tot de zorggroep horen, aangezien de school een hoog uitstroomperspectief heeft. Een andere participant, die op scholen in verschillende wijken werkt, gaf aan dat de etnische achtergrond en thuissituatie van een leerling bepalen of een score problematisch is. Twee andere participanten noemden dat ze allereerst naar leerkrachtgedrag kijken bij het beoordelen van scores, doordat zij hier recent een negatieve ervaring mee hadden gehad. Schoolcontext leek dus medebepalend te zijn in dit onderzoek voor de percepties van leerkrachten op problematische groei. Gezien deze bevindingen is een mogelijke belemmering dat er twee keer een situatie was waarin twee participanten op dezelfde school werkzaam waren. Echter, doordat in beide gevallen één van de participanten slechts twee maanden werkzaam was op de desbetreffende school, wordt verwacht dat de gedeelde context weinig tot geen invloed heeft gehad op de percepties van deze leerkrachten.

Als laatste zijn er limitaties met betrekking tot de betrouwbaarheid en validiteit. De betrouwbaarheid van het onderzoek was matig. Een mogelijke verklaring hiervoor was dat er een relatief groot aantal codes waren, waardoor Cohen's Kappa over het algemeen lager uitvalt. Daarnaast zijn relatief kleine stukken van de transcripten gecodeerd door een tweede onderzoeker. Hierdoor kan een klein verschil in codering, grote invloed op de betrouwbaarheid hebben. Er heeft

een overleg met de tweede onderzoeker plaatsgevonden om discrepanties te bespreken. Op basis van dit overleg zijn twee codes hernoemd. Om het onderzoek te valideren is naar alle participanten ($N=8$) een member-check verstuurd met een korte samenvatting van de resultaten (zie Bijlage 6). Drie participanten (37.5%) hebben hierop een reactie gegeven. Hoewel deze participanten in hun reacties aangaven dat ze zich herkenden in de resultaten, kan niet gezegd worden of dit ook het geval was voor de participanten die geen reactie hebben gegeven. Tot slot is één opname van een interview mislukt, waardoor er geen transcript van dit interview is, maar slechts een samenvatting. De samenvatting is na een korte tijd geschreven door de interviewer. Hierdoor wordt verwacht dat de inhoud van veel uitspraken overeenkomen met de realiteit. Echter, doordat verondersteld werd dat het interview werd opgenomen, zijn minimale aantekeningen gemaakt tijdens het interview. Daardoor kunnen uitspraken afwijken van de oorspronkelijke implicatie en kunnen belangrijke uitspraken niet zijn opgenomen in de samenvatting. Het design liet het niet toe om het interview nogmaals uit te voeren, omdat de percepties van de participant mogelijk beïnvloed zijn door de casussen en de nabespreking van het onderzoek.

Vanuit dit onderzoek kunnen verschillende aanbevelingen worden gedaan voor vervolgonderzoek. Allereerst wordt aangeraden om meer variatie aan te brengen in de aangeboden casussen. Bijvoorbeeld het toevoegen van casussen met meerdere momenten van geringe groei. Daarnaast zou het interessant zijn om een casus van een sterke leerling met negatieve leerlingkenmerken en een zwakke leerling met positieve leerlingkenmerken toe te voegen om de invloed van gedrag op de verwachtingen van leerkrachten omtrent groei verder te onderzoeken.

Bovendien zou voor een vervolgonderzoek meer variatie aangebracht kunnen worden in de steekproef. In dit onderzoek leek schoolcontext van invloed te zijn op de percepties van participanten. Daarom wordt aanbevolen om in de steekproef te variëren op schoolcontext, bijvoorbeeld een grote en kleine school, of verschillende typen onderwijs.

Kortom, in deze masterthesis is onderzocht op welke manier leerkrachten verschillen in verwachtingen over problematische groei bij sterke versus zwakke leerlingen. Hieruit bleek dat leerkrachten eerder verwachten dat problematische groei van sterke leerlingen van tijdelijke aard is. Ook wordt de oorzaak van problematische groei vooral bij het aanbod gezocht. Bij zwakke leerlingen is gedrag van grote invloed als verklaring voor problematische groei. Het is van belang dat leerkrachten zich bewust worden van deze verschillen in verwachtingen, aangezien de verwachtingen van leerkrachten van invloed zijn op de prestaties van leerlingen (Brault et al., 2014; Timmermans et al., 2016; Wang et al., 2021).

Referenties

- Aydin, O., & Ok, A. (2022). A Systematic Review on Teacher's Expectations and Classroom Behaviors. *International Journal of Curriculum and Instructional Studies*, 12(1), 247–274. <https://doi.org/10.31704/ijocis.2022.011>
- Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, 34(1), 215–220. <https://doi.org/10.1093/ije/dyh299>
- Becherer, J., Köller, O., & Zimmermann, F. (2021). Externalizing behaviour, task-focused behaviour, and academic achievement: An indirect relation? *British Journal of Educational Psychology*, 91(1), 27–45. <https://doi.org/10.1111/bjep.12347>
- Brault, M. C., Janosz, M., & Archambault, I. (2014). Effects of school composition and school climate on teacher expectations of students: A multilevel analysis. *Teaching and Teacher Education*, 44, 148–159. <https://doi.org/10.1016/j.tate.2014.08.008>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. <https://doi.org/10.1191/1478088706qp063oa>
- Çambay, Ö., & Kazanç, S. (2021). Self-Efficacy Perceptions of Science Teachers for Assessment Literacy. *International Journal of Education and Literacy Studies*, 9(2), 142. <https://doi.org/10.7575/aiac.ijels.v.9n.2p.142>
- Cito. (2019). *Toetsscore, vaardigheidsscore... en dan?* Cito. Geraadpleegd op 19 november 2022, van <https://www.cito.nl/-/media/files/ve-en-po/cito-flyer-toetsscore-vaardigheidsscore-en-dan.pdf>
- Cito. (z.d.-a). *Maak kennis met Leerling in beeld. Groei centraal.* Geraadpleegd op 4 januari 2023, van <https://www.cito.nl/onderwijs/primair-onderwijs/lvs-leerling-in-beeld/groei-centraal>
- Cito. (z.d.-b). *Over Cito. Schooladvies: Cito levert een stukje van de puzzel.* Geraadpleegd op 9 november 2022, van <https://www.cito.nl/onderwijs/primair-onderwijs/schooladvies-cito-levert-stukje-puzzel>
- De Jong, P. F., de Bree, E. H., Henneman, K., Kleijnen, R., Loykens, E. H. M., Rolak, M., Struiksmā, A. J. C., Verhoeven, L., & Wijnen, F. N. K. (2016). *Dyslexie: diagnostiek en behandeling* (5^e herziene versie). Brochure van de Stichting Dyslexie Nederland
- Ellis, C. & Berger, L. (2001). Their Story/My Story/Our Story: Including the Researcher's Experience in Interview Research. In J.F. Gubrium & J.A. Holstein (Eds.). *Handbook of Interview Research*. <https://doi.org/10.4135/9781412973588>
- Frans, N. (2019). *A captivating snapshot of standardized testing in early childhood : on the stability and utility of the Cito preschool/kindergarten tests.*

- Frans, N., Minnaert, A., Post, W., & Oenema-Mostert, I. (2020). Ondergemiddeld is niet gelijk problematisch. Signalering met de Cito-kleutertoetsen. *Tijdschrift Voor Orthopedagogiek*, 2020(2), 20–27.
- Groenestijn, M. van, Borghouts, C., Janssen, Christien., Bodin, Connie., & Laeven, Agnes. (2011). *Protocol ernstige rekenwiskunde-problemen en dyscalculie: BAO, SBO, SO*. Koninklijke Van Gorcum.
- Hecht, S. A., & Greenfield, D. B. (2002). Explaining the predictive accuracy of teacher judgments of their students' reading achievement: The role of gender, classroom behavior, and emergent literacy skills in a longitudinal sample of children exposed to poverty. *Reading and Writing: An Interdisciplinary Journal*, 15, 789–809.
- Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough? *Qualitative Health Research*, 27(4), 591–608. <https://doi.org/10.1177/1049732316665344>
- Heslin, P. A., & Klehe, U. C. (2006). Self-efficacy. In S.G. Rogelberg (Ed.), *Encyclopedia of Industrial/Organizational Psychology* (Vol. 2, pp. 705–708). Sage.
<http://ssrn.com/abstract=1150858>
- Hollenberg, J., van Boxtel, H., & Keuning, J. (2014). Werken met leerlingen met hoge vaardigheidsscores in het Cito Volgsysteem primair en speciaal onderwijs. *Tijdschrift Voor Orthopedagogiek*, 53, 347–356.
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. *Biometrics*, 33(1), 159–174. <https://doi.org/10.2307/2529310>
- Luján, E. L. (2021). The beliefs of primary school teachers: A comparative analysis. *International Journal of Instruction*, 14(3), 223–240. <https://doi.org/10.29333/iji.2021.14313a>
- Martínez-Vicente, M., Suárez, J. M., & Valiente-Barroso, C. (2019). Executive functioning and learning in primary school students. *Electronic Journal of Research in Educational Psychology*, 17(1), 55–80. <https://www.researchgate.net/publication/332130588>
- Matthews, S. H., Lewis, J., & Meredith, B. (1989). Daughters who care: daughters caring for mothers at home. *Contemporary Sociology*, 18(6), 954–954. <https://doi.org/10.2307/2074237>
- Rijksoverheid. (z.d.). *Hoe legt de basisschool de prestaties van mijn kind vast?* Geraadpleegd op 9 november 2022, van <https://www.rijksoverheid.nl/onderwerpen/basisonderwijs/vraag-en-antwoord/hoe-legt-de-basischool-de-prestaties-van-mijn-kind-vast>

- Riley, T., & Ungerleider, C. (2012). Self-fulfilling Prophecy: How Teachers' Attributions, Expectations, and Stereotypes Influence the Learning Opportunities Afforded Aboriginal Students. *Canadian Journal of Education*, 35(2), 303–333.
<https://doi.org/10.2307/canajeducrevucan.35.2.303>
- Roberts, S.J. (1978). *Test Floor and Ceiling Effects. ESEA Title I. Evaluation and reporting System* [Report]. RMC Research Corp., Mountain View.
- Scholen op de kaart (z.d.). *Vind en vergelijk scholen in de buurt*. Geraadpleegd op 11 juni 2023, van <https://scholenopdekaart.nl/>
- Sneyers, E., Vanhoof, J., & Mahieu, P. (2019). The Impact of Primary School Teachers' Expectations of Pupils, Parents and Teachers on Teacher Track Recommendations. *Educational Studies - AESA*, 55(3), 327–345.
<https://doi.org/10.1080/00131946.2018.1562925>
- Stecher, B. M. (2002). Consequences of Large-scale, High-stakes Testing on School and Classroom Practice. In L. S. Hamilton, B. M. Stecher, & S. P. Klein (Eds.), *Making Sense of Test-Based Accountability in Education* (pp. 79–100). RAND Corporation.
<https://doi.org/10.7249/mr1554edu.11>
- Timmermans, A. C., de Boer, H., & van der Werf, M. P. C. (2016). An investigation of the relationship between teachers' expectations and teachers' perceptions of student attributes. *Social Psychology of Education*, 19(2), 217–240. <https://doi.org/10.1007/s11218-015-9326-6>
- Timmermans, A. C., Kuyper, H., & van der Werf, G. (2015). Accurate, inaccurate, or biased teacher expectations: Do Dutch teachers differ in their expectations at the end of primary education? *British Journal of Educational Psychology*, 85(4), 459–478.
<https://doi.org/10.1111/bjep.12087>
- Van der Kleij, F. M., & Eggen, T. J. H. M. (2013). Interpretation of the score reports from the Computer Program LOVS by teachers, internal support teachers and principals. *Studies in Educational Evaluation*, 39(3), 144–152. <https://doi.org/10.1016/j.stueduc.2013.04.002>
- Van Loon, D., van der Meulen, B. F., & Minnaert, A. E. M. G. (2015). *Effectonderzoek in gedragswetenschappen. Een introductie*. Garant.
- Verhelst, N.D. (1993). Itemresponsstheorie. In T.J.H.M. Eggen & P.F. Sanders (red.). *Psychometrie in de Praktijk* (pp. 83-183). Cito. Geraadpleegd op 3 juni 2023, van <https://www.cito.nl/kennis-en-innovatie/onderzoek/psychometrie-in-de-praktijk-boek>
- Vlug, K. F. M. (1997). Because every pupil counts: the success of the pupil monitoring system in The Netherlands. *Education and Information Technologies*, 2, 287–306.

- Wang, M., & Cai, J. (2016). The Application of Pygmalion Effect in Classroom Education. *Advances in Social Science, Education and Humanities Research, ICADCE 2016*.
<https://doi.org/10.2991/icadce-16.2016.239>
- Wang, S., Meissel, K., & Rubie-Davies, C. M. (2021). Teacher expectation effects in Chinese junior high schools: Exploring links between teacher expectations and student achievement using a hierarchical linear modelling approach. *Social Psychology of Education, 24*(5), 1305–1333.
<https://doi.org/10.1007/s11218-021-09654-7>
- Weinstein, R. S., Marshall, H. H., Sharp, L., & Botkin, M. (1987). Pygmalion and the Student: Age and Classroom Differences in Children's Awareness of Teacher Expectations. *Child Development 58*(4), 1079-1093. <https://doi.org/10.2307/1130548>

Bijlagen

Bijlage 1	Interviewleidraad
Bijlage 2	Casussen
Bijlage 3	Informatiebrief
Bijlage 4	Geïnformeerde toestemming
Bijlage 5	Member-check
Bijlage 6	Codeboek

Bijlage 1

Interviewleidraad

Onderdeel	Vragen/Onderwerpen
Introductie onderzoek	<ul style="list-style-type: none"> - Bedanken, resultaten mee, duur (30-40 min), elk moment stoppen. - Doel: Ik wil inzicht krijgen in hoe leerkrachten/IB'ers Cito-toetsen analyseren en welke factoren hierbij worden meegenomen. - Geïnformeerde toestemming getekend (nog vragen?) + toestemming voor geluidsopname + waarom? - Opbouw van gesprek - Vragen vooraf? <p>OPNAME STARTEN</p>
Algemeen	<p><i>Man/vrouw Dorp/stad Volgorde A/B/C/D</i></p> <ol style="list-style-type: none"> 1. Hoelang werk je al in het onderwijs? <ul style="list-style-type: none"> o Voor IB: heb je ook als leerkracht gewerkt? 2. Kun je wat vertellen over de opleiding(en) die je hebt gevolgd? 3. Hoe lang werk je al op deze school?
Eigen ervaring	<ol style="list-style-type: none"> 4. Kun je mij wat vertellen over je ervaringen met de toetsen van Cito? <u>Jaren ervaring, afnemen, analyseren</u> 5. Wat kan je vertellen over hoe Cito toetsen worden ingezet op jouw school? <u>Afname, analyseren, betrokkenen, acties n.a.v. analyse</u> <p>Je hebt de meest recente toetsresultaten van je/een klas meegenomen, zou je die erbij willen pakken?</p> <ol style="list-style-type: none"> 6. Welke leerlingen uit je groep vallen jou op? <ul style="list-style-type: none"> o Wat maakt dat je deze leerlingen opvallend vindt? 7. Over wiens toetsresultaten maak je je zorgen? 8. Welke leerlingen stagneren in hun ontwikkeling, volgens jou? <ul style="list-style-type: none"> o Wat is de reden dat je denkt dat deze leerlingen stagneren? 9. Welke acties zou je ondernemen voor deze groep? <u>Zorgleerlingen, plusleerlingen</u> 10. Wat verwacht je bij de volgende toets afname van de leerlingen die we zojuist hebben besproken? (Maximaal 3 leerlingen bespreken) <u>Niveau, groei</u> <ul style="list-style-type: none"> - Ik wil als voorbeeld graag een Cito rekentoets nemen. Zodra deze is afgenomen ga je die analyseren. Waar kijk je allereerst naar? - Wat voor leerlingen vallen jou op? - Bij het analyseren maak je ook een groepsplan (neem ik aan). Hoe deel je je leerlingen hierbij in? (Risicogroep etc.) - Wanneer begin/maak je je zorgen te maken om een leerling? - Welke acties zou je ondernemen voor deze groepen?
Casus 1&2	<p>Ik ga je nu de rekenresultaten van vier fictieve leerlingen achter elkaar laten zien en bij iedere leerling heb ik een paar vragen.</p> <p><i>Bij beide casussen:</i></p> <ol style="list-style-type: none"> 11. Wat is je eerste indruk van deze leerling? 12. Maak je je zorgen om deze leerling? <ul style="list-style-type: none"> o Vanwaar maak je je wel/geen zorgen? <u>Groei</u> 13. Wat zou je de leerkracht van deze groep adviseren? <ul style="list-style-type: none"> o Wat maakt dat je dit advies geeft?
<p>Volgordes: A: Jip - Robin B: Jip - Robin C: Robin - Jip D: Robin - Jip</p>	

<p>Casus 3&4</p> <p>Volgordes: A: Emma - Luuk B: Luuk - Emma C: Emma - Luuk D: Luuk - Emma</p>	<p>Ik ga je nu de derde en vierde leerlingen laten zien en hierbij vertel ik ook wat extra informatie over de leerling.</p> <p><i>Bij beide casussen:</i></p> <p>14. Wat kan je zeggen over deze leerling? <u>Zorgen, groei, leerlingkenmerken, adviezen</u></p>
<p>Afsluiten</p>	<p>Misschien had je het al gemerkt aan de vraagstelling, maar ik was vooral benieuwd naar de hoe er wordt gekeken naar groei en de vaardigheidsscores. Waarschijnlijk heb je al meegekregen dat het LOVS van Cito vernieuwd is ('Leerling in Beeld').</p> <p>Het idee achter de nieuwe rapportages is dat de focus meer op groei komt te liggen. Een lastige vraag waar Stichting Cito een antwoord op zoekt is wanneer (een gebrek aan) groei voor leerkrachten als problematisch wordt ervaren.</p> <p>Daarom zou ik willen afsluiten met de laatste vraag:</p> <p>15. Op welke manier vind jij dat problematische groei herkenbaar is in de vaardigheidsscores van een leerling?</p> <ul style="list-style-type: none"> - Bedanken - Nog iets wat niet aan bod is gekomen? - Procedure member-check: mailtje met interpretatie binnen ongeveer 4 weken. Fijn als je een reactie geeft of je je herkent in de interpretatie. Zo niet, kort toelichten waarom.
<p><i>Optioneel: na opname</i></p>	<p>"Sinds ik de opname heb uitgezet heb je nog een aantal interessante dingen gezegd die bruikbaar zijn voor mijn onderzoek. Heb je er bezwaar tegen dat ik deze informatie ook gebruik voor het onderzoek?"</p>

Emma is een vrolijk meisje van 10 jaar oud. Ze is heel gemotiveerd om te leren en is meestal de eerste die haar vinger op steekt als ze het antwoord weet. Ze doet goed mee tijdens de les en luistert meteen als de leerkracht het stilleteken geeft. Ze vindt ieder vak interessant, maar vooral de wereldoriëntatie vakken interesseren haar. Op de methodetoetsen scoort Emma goed, ze werkt hard en nauwgezet. Haar vader heeft weinig tijd voor haar, hij is veel aan het werk als hartchirurg. Haar moeder werkt parttime als kleuterjuf en is meer thuis. Haar ouders helpen haar geregeld met de tafels oefenen en het voorbereiden van de spreekbeurt.

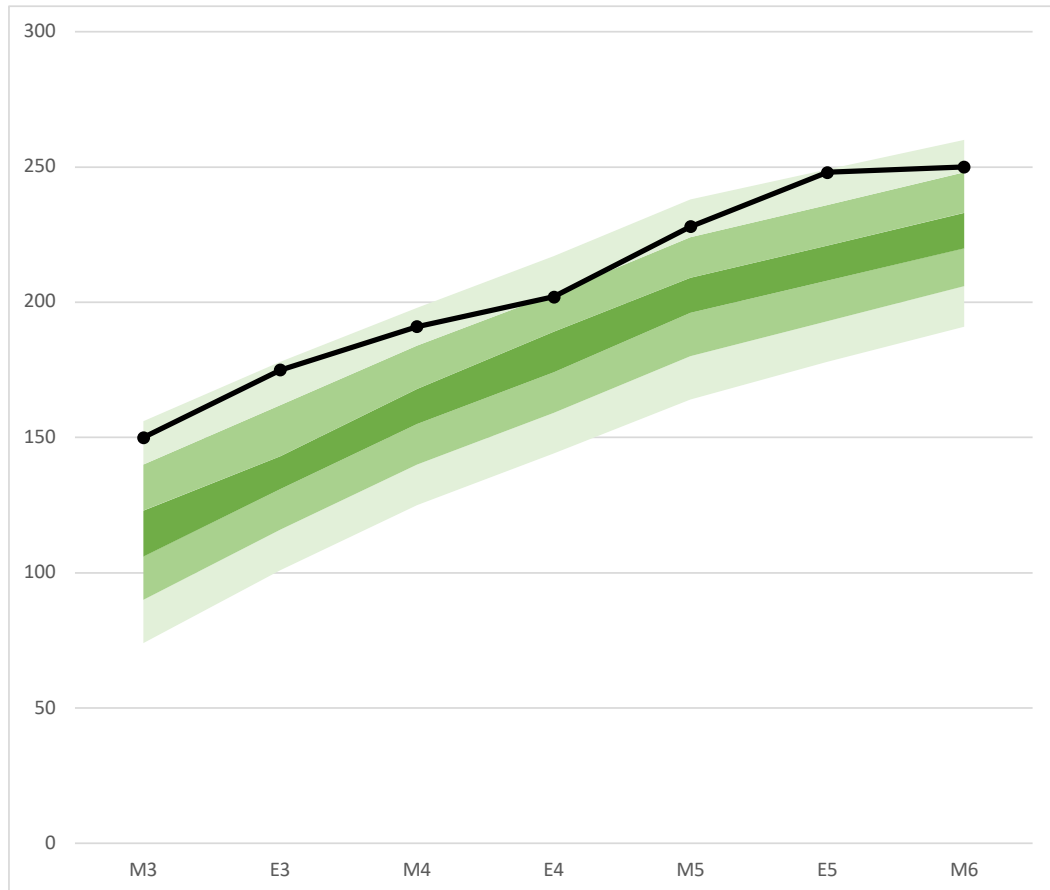
Luuk is een stoere, voetbal jongen van 10 jaar oud. Hij komt wat dromerig over, waardoor de leerkracht hem regelmatig even bij de les moet roepen. Luuk is een grote fan van wisbordjes, dat wil zeggen, hij tekent er graag op. Luuk zit vaak omgedraaid op zijn stoel. Hij vindt het lastig om de instructies van de leerkracht direct op te volgen. Het liefst speelt hij buiten in de pauze. Als lievelingsvak noemt Luuk gym. De methodetoetsen van Luuk zijn net voldoende. Hij vindt het lastig om alle opdrachten binnen de tijd af te krijgen en zijn schriften zijn vaak een rommeltje. Beide ouders werken fulltime, moeder is kapper en vader als automonteur in een garage.

Bijlage 2

Casussen

Leerlingrapport

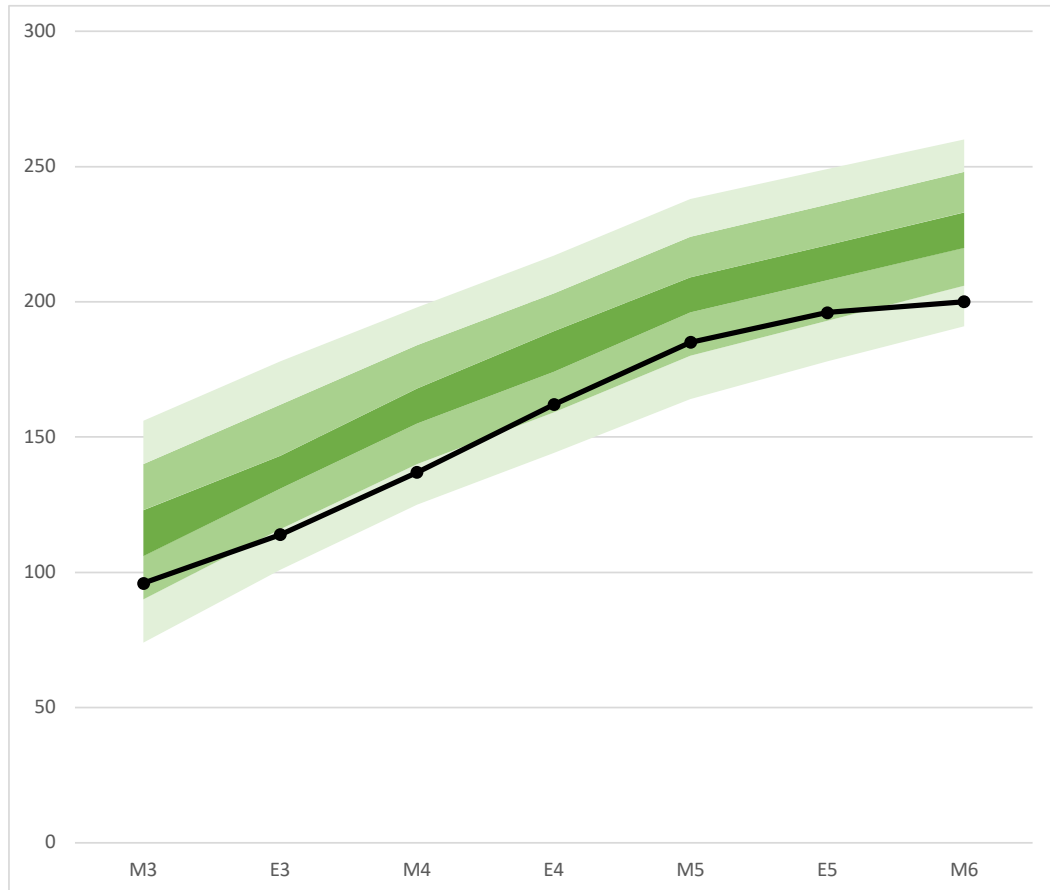
Leerling: Jip Janssen
Toets: Rek-Wisk 3.0



Afn. Datum	Jgr	Taak	Vaard.score	Niveau norm
30-01-20	3	M3	150	I
30-06-20	3	E3	175	I
30-01-21	4	M4	191	I
30-06-21	4	E4	202	II
30-01-22	5	M5	228	I
30-06-22	5	E5	248	I
30-01-23	6	M6	250	I

Leerlingrapport

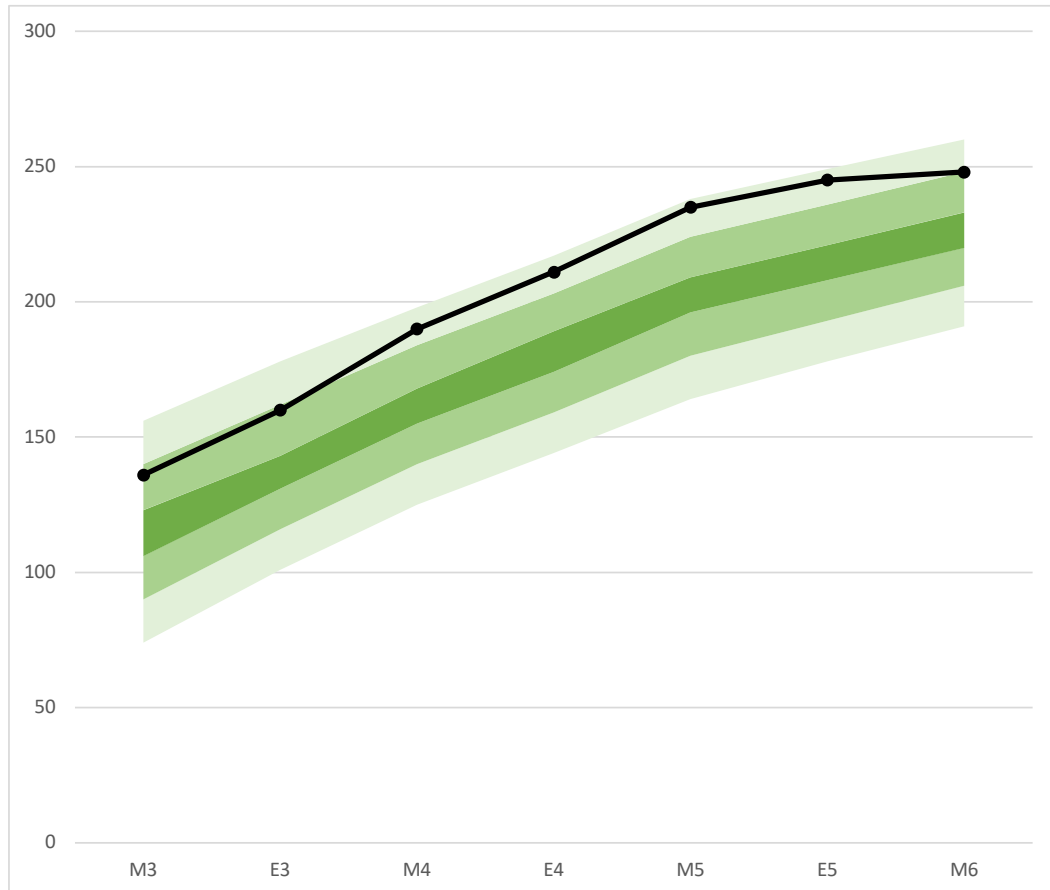
Leerling: Robin De Vries
Toets: Rek-Wisk 3.0



Afn. Datum	Jgr	Taak	Vaard.score	Niveau norm
30-01-20	3	M3	96	IV
30-06-20	3	E3	114	V
30-01-21	4	M4	137	V
30-06-21	4	E4	162	IV
30-01-22	5	M5	185	IV
30-06-22	5	E5	196	IV
30-01-23	6	M6	200	V

Leerlingrapport

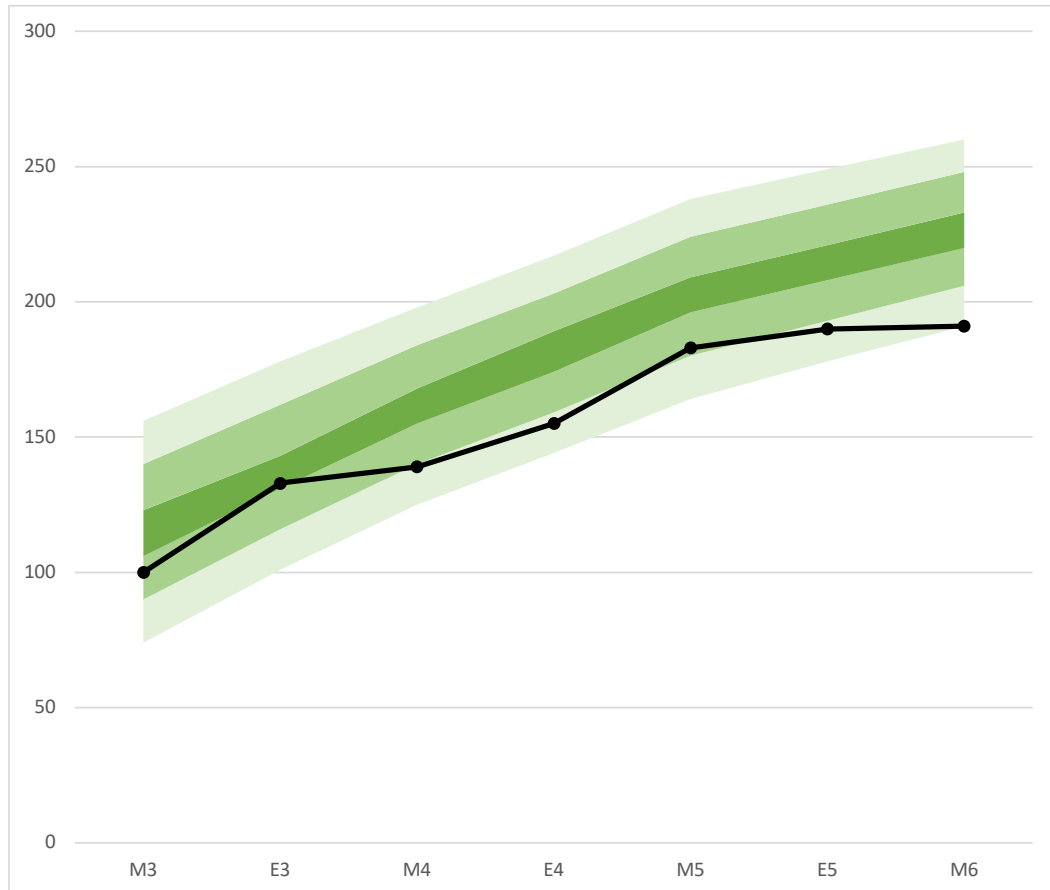
Leerling: Emma De Jong
Toets: Rek-Wisk 3.0



Afn. Datum	Jgr	Taak	Vaard.score	Niveau norm
30-01-20	3	M3	136	II
30-06-20	3	E3	160	II
30-01-21	4	M4	190	I
30-06-21	4	E4	211	I
30-01-22	5	M5	235	I
30-06-22	5	E5	245	I
30-01-23	6	M6	248	II

Leerlingrapport

Leerling: Luuk Boonstra
Toets: Rek-Wisk 3.0



Afn. Datum	Jgr	Taak	Vaard.score	Niveau norm
30-01-20	3	M3	100	IV
30-06-20	3	E3	133	III
30-01-21	4	M4	139	V
30-06-21	4	E4	155	V
30-01-22	5	M5	183	IV
30-06-22	5	E5	190	V
30-01-23	6	M6	191	V

Bijlage 3

Informatiebrief

Beste leerkracht of intern begeleider,

Ik benader u voor een masteronderzoek van de afdeling Pedagogiek en Onderwijswetenschappen van de Rijksuniversiteit Groningen. Met dit masteronderzoek hoop ik meer inzicht te krijgen in de manier waarop leerkrachten en intern begeleiders naar Cito resultaten kijken en handelen, een vraag die speelt vanuit stichting Cito. Ik ben op zoek naar leerkrachten en intern begeleiders die werkzaam zijn op een school die gebruik maakt van het LOVS van Cito.

In het kort

- Het onderzoek gaat over het analyseren van en handelen op basis van Cito resultaten.
- Dit wordt onderzocht met een eenmalig gesprek van ongeveer 30-40 minuten.
- Na afloop van het gesprek ontvangt u een mail waarop u eenmalig hoeft te reageren.
- Alle gegevens worden anoniem en vertrouwelijk behandeld.

Wat onderzoeken we?

In dit onderzoek wordt gekeken hoe leerkrachten en intern begeleiders naar Cito resultaten kijken. Hoewel Cito de toetsen weliswaar ontwikkelt, zijn leerkrachten en intern begeleiders bepalend voor het gebruik. Wanneer wordt een score in de praktijk bijvoorbeeld als problematisch gezien? En welke acties worden ondernomen op basis van de Cito resultaten? Het onderzoek zal meer inzicht geven in welke factoren leerkrachten meenemen in het analyseren van Cito toetsen.

Hoe onderzoeken wij dit?

Voor dit onderzoek worden gesprekken gevoerd met leerkrachten en intern begeleiders. Hiervoor zal ik eenmalig naar uw school komen om kort met u te spreken, op een moment dat u uitkomt. Het gesprek is volledig vrijblijvend en mag op ieder moment gestopt worden. Ik ben geïnteresseerd in uw persoonlijke ervaringen bij het analyseren van Cito toetsen en de mogelijke acties na de analyse. Daarnaast worden er vier korte casussen besproken. Direct na afloop van het gesprek is een korte nabespreking over het onderzoek. Er wordt een audio opname van het gesprek gemaakt, zodat ik ons gesprek terug kan luisteren.

Wat vragen we van u?

Als u met dit onderzoek mee wilt doen, zal er eenmalig een gesprek plaatsvinden, bij voorkeur op uw school. Mocht het wenselijk zijn dan kan het gesprek ook online worden gevoerd via Teams. Het gesprek duurt ongeveer 30-40 minuten. U hoeft zich niet voor te bereiden op het gesprek. Wel vraag ik of u het **Groepsoverzicht – toets** van Cito rekenen (bij voorkeur geanonimiseerd) wilt meenemen naar het gesprek. Als intern begeleider kiest u één groep uit waarvan u de gegevens meeneemt. Deze gegevens verlaten de school niet en persoonlijke informatie over leerlingen wordt niet gebruikt in het onderzoek.

Circa 1-4 weken na het gesprek ontvangt u een mail met een korte interpretatie van het gesprek. U wordt verzocht dit door te lezen en te bevestigen of u zich herkent in de interpretatie of dat uw antwoorden wellicht verkeerd zijn geïnterpreteerd.

Wat gebeurt er met de gegevens van u en uw leerlingen?

Ik ben niet geïnteresseerd in persoonlijke informatie van leerlingen, zoals namen en scores, enkel in uw interpretatie van de scores. Het groepsoverzicht blijft in uw beheer en verlaat de school niet. U wordt verzocht de namen eventueel af te plakken zodat ik niks kan herleiden. Uw naam en die van de school zullen niet worden genoemd in het onderzoek en worden niet gedeeld met eventuele derden. In publicaties zal geen informatie worden genoemd die herleidbaar is naar u of uw leerlingen. Er wordt vertrouwelijk met alle informatie van u en uw leerlingen omgegaan.

Alle verzamelde informatie zal op een beveiligde omgeving van de Rijksuniversiteit Groningen worden opgeslagen. Enkel de onderzoeker en de begeleider van het onderzoek zullen toegang hebben tot de informatie. De opnames van het gesprek worden na het onderzoek verwijderd. De uitgeschreven en geanonimiseerde gesprekken worden voor een periode van 10 jaar bewaard. Een beperkt aantal personen hebben toegang tot deze informatie om onderzoeksuitkomsten te controleren.

Mee doen?

Wanneer u mee wilt doen met dit onderzoek kunt u de bijgevoegde toestemmingsverklaring invullen en retour sturen naar het volgende e-mailadres: j.s.van.aalsum@student.rug.nl of meenemen naar het gesprek.

Deze brief wordt apart opgeslagen van de rest van de data, op een beveiligde locatie. Na afronding van het onderzoek wordt de toestemmingsverklaring vernietigd.

Wanneer ik uw toestemming heb ontvangen neem ik zo snel mogelijk contact met u op om verdere afspraken te maken.

Ik hoop u voldoende te hebben geïnformeerd en wil u alvast hartelijk bedanken voor uw medewerking!

Hartelijke groet,
Judith van Aalsum

Contact

Mocht u nog vragen hebben dan kunt u mij bereiken via onderstaande gegevens:

Judith van Aalsum
E-mail: j.s.van.aalsum@student.rug.nl
Telefoon: 06 12188231

Begeleiding: Dr. Niek Frans
E-mail: n.frans@rug.nl
Telefoon: 050 3637631

Bijlage 4

Geïnformeerde toestemming

Ik (naam leerkracht/intern begeleider)

.....

stem toe mee te doen aan een onderzoek naar het analyseren van Cito resultaten, dat uitgevoerd wordt door masterstudent Judith van Aalsum.

Ik bevestig de volgende zaken:

- Ik heb de informatiebrief gelezen en heb de inhoud ervan begrepen.
- Ik ben me ervan bewust dat deelname geheel vrijwillig is. Ik kan mijn medewerking op elk tijdstip stopzetten en de verkregen informatie laten verwijderen. Ik hoef niet te verklaren waarom ik stop. Stoppen heeft geen negatieve gevolgen.
- Ik stem toe dat er een geluidsopname wordt gemaakt van het gesprek. Ik ben me ervan bewust dat de geluidsopname beveiligd wordt bewaard en niet met derden wordt gedeeld. De geluidsopname wordt na het onderzoek verwijderd. Daarnaast ben ik ervan bewust dat eventuele herleidbare informatie niet wordt opgenomen in de resultaten van het onderzoek en publicatie.

Door het ondertekenen van deze brief verklaar ik dat ik wil deelnemen aan dit onderzoek en dat ik voldoende geïnformeerd ben over het doel en de uitvoering van het onderzoek.

Datum:

Handtekening leerkracht/intern begeleider:

.....

.....

Graag verzoek ik u hieronder uw e-mailadres te noteren zodat ik u kan bereiken na afloop van het gesprek voor een korte interpretatie van het gesprek. Mocht u een korte samenvatting van de resultaten van het gehele onderzoek ook graag ontvangen dan kunt u dat aangeven door onderstaande optie aan te vinken.

E-mailadres:

.....

Ik ontvang graag een korte samenvatting van de resultaten van het onderzoek.

Datum:

Handtekening onderzoeker J. van Aalsum:

.....

.....

Bijlage 5

Member-check

Algemene bevindingen over problematische groei in Citoscores

Om te bepalen wat problematisch is, gebruiken leerkrachten normen. De norm wordt gesteld door de school of er wordt een checklist van Cito gebruikt voor gemiddelde vaardigheidsgroei. De geïnterviewde leerkrachten vinden het problematisch als een leerling in vaardigheidsscore achteruitgaat, gelijk blijft of slechts een geringe groei laat zien. Ook een leerrendement van <80% wordt als problematisch ervaren. Daarnaast gaven leerkrachten aan dat de populatie van de school een rol speelt. Er gelden wellicht andere normen als een leerling bijvoorbeeld een andere taal spreekt. Uit de meeste interviews kwam naar voren dat stagnatie op lange termijn is. Sommige leerkrachten maken zich al zorgen bij één keer weinig groei maar de meeste leerkrachten maken zich pas zorgen als een leerling minimaal twee keer achter elkaar weinig groei laat zien.

Uit de interviews kwamen een aantal factoren die worden meegenomen in de overweging of groei problematisch is. Als een leerling binnen zijn eigen niveau groeit, vinden de meeste geïnterviewde leerkrachten het goed. Als het vaardigheidsniveau (I-V of A-E) daalt, vinden ze het problematisch. Ook groepsresultaten, methodelessen en toetsen, de resultaten van eerdere Cito toetsen, de Citoresultaten van andere vakgebieden en de potentie die een leerling in zich heeft, worden meegenomen door de geïnterviewde leerkrachten. Als de Citoscore afwijkt van deze andere resultaten, wordt groei eerder als problematisch gezien, tenzij het een trend is. Bijvoorbeeld omdat de leerkracht ieder jaar een dip in scores ziet op een bepaalde Cito-toets, en de scores uit ervaring wel weer omhooggaan.

Als laatste spelen kind kenmerken en het gedrag van leerlingen een belangrijke rol volgens de geïnterviewde leerkrachten. Er wordt door leerkrachten gekeken hoe een leerling in de klas is. Werkhouding, inzet, opletten, zelfvertrouwen en concentratie zijn allemaal van invloed op de resultaten. Ook de werkhouding tijdens de Cito-toets kan een verklaring zijn voor de score. Bijvoorbeeld als een leerling blokkeert, zich niet inzet of niet goed in zijn/haar vel zit.

Bevindingen over sterke leerlingen (I-score of A-score)

De meeste geïnterviewde leerkrachten maken zich geen zorgen om een sterke leerling met weinig vaardigheidsgroei als de leerling een constante groei laat zien. Een I-score maakt ook dat leerkrachten zich minder snel zorgen maken dan bij weinig vaardigheidsgroei van een leerling met een IV- of V-score. Uit de interviews kwam naar voren dat wordt verwacht dat de score van de sterke leerling wel weer omhooggaat. Leerkrachten gaven aan dat dit komt omdat sterke leerlingen gemotiveerd zijn en een goede werkhouding laten zien.

Verklaringen voor problematische groei bij sterke leerlingen liggen volgens leerkrachten vooral in het aanbod en de lesstof. Uit de interviews kwam naar voren dat een sterke leerling uitdaging nodig heeft om te groeien. Daarnaast kan de lesstof moeilijker worden, te snel gaan of zijn de basisvaardigheden onvoldoende geautomatiseerd. Andere dingen die in de interviews werden genoemd zijn dat sterke leerlingen mogelijk problemen hebben met executieve functies of de toets vragen anders geïnterpreteerd hebben. Tevens werd genoemd dat een sterke leerling weinig groei laat zien omdat zijn/haar plafond in het leren bereikt is.

Bevindingen over zwakke leerlingen (V-score of D-/E-score)

Leerkrachten gaven in de interviews aan zich zorgen te maken als een zwakke leerling in niveau gezakt is. Mogelijk scoort de zwakke leerling lager dan er in potentie in zit, maar als de randvoorwaarden (zoals inzet en werkhouding) goed zijn, dan is een lage score misschien het hoogst haalbare. De geïnterviewde leerkrachten verwachten voor de toekomst geen III-score van een leerling die een V-score heeft, maar wel dat er weer groei is. Hiervoor moet er beter worden aangesloten op de behoeften van de leerling. Mogelijke

manieren om dit te bereiken werden genoemd in de interviews, zoals: meer betekenis geven aan rekenen, verlengde instructie, moeilijke onderdelen oefenen, visuele ondersteuning en concreet materiaal. Ook hoeft een zwakke leerling niet alles af te hebben. Als laatste heeft de zwakke leerling het nodig om gemotiveerd te worden.

Uit de interviews bleek dat verklaringen voor problematische groei bij zwakke leerlingen vooral gedragsmatige factoren zijn. Weinig concentratie en aandacht, gedemotiveerd en een slechte werkhouding en inzet kunnen volgens leerkrachten verklaren waarom de zwakke leerling niet gegroeid is. Ook kan de talige kant van rekenen de zwakke leerling belemmeren of ontbreken executieve functies. Mogelijk vindt de zwakke leerling de lesstof lastiger of is de lengte van de Citotoets te lang. Bij gedragsproblemen zijn er zorgen bij de geïnterviewde leerkrachten of de zwakke leerling wel toetsbaar is.

Bijlage 6

Codeboek

Verwachtingen over problematische groei bij leerlingen in het algemeen

Normen voor problematische groei

Bepaling van norm

- Checklist voor gemiddelde vaardigheidsgroei
- Norm hangt af van de school

Norm voor problematisch groei

- In vaardigheidsscore achteruitgegaan is problematisch
- Problematisch als leerling niet groeit
- Groei van <4 punten op rekenen per halfjaar is problematisch
- Leerrendement <80% is problematisch

Groei afhankelijk van populatie

- Groei stagneert als je andere taal spreekt

Op basis van het aantal meetmomenten

- Problematisch bij één keer weinig groei

Stagneren is op lange termijn

- Problematisch bij (minimaal) twee keer achter elkaar weinig groei
 - o Kind kan niet altijd pieken
 - o Cito is een momentopname

Vaardigheidsniveau van de leerling

- Groeien binnen eigen niveau goed
- Problematisch als niveau daalt
- Verschil twee vaardigheidsniveaus opvallend

Andere resultaten dan het laatste Citoresultaat

- Groepsresultaat van belang bij beoordelen (van een individuele) score
- Altijd naar methodetoetsen en lessen kijken
 - o Zorgelijk als lessen goed zijn maar Citoresultaten niet
- Vergelijken met voorgaande Cito scores

Resultaten van meerdere/verschillende vakken

- Zorgen bij weinig groei in meerdere vakken
- Hoog scoren, op één vakgebied na, is opvallend

Eigen ervaring van de leerkracht

- Trend kan verklaring zijn voor groepsresultaten waardoor leerkracht zich minder snel zorgen maken
 - o Dipje in midden 6 gebeurt vaker want meer inzicht nodig
- Terugzakken in niveau gaat uit ervaring wel weer omhoog

Problematische groei afhankelijk van leerpotentie

- Verwachting over groei en niveau op basis van potentietest
- Afwijken van leerpotentie is problematisch
 - o Problematische groei als het niet overeenkomt met eigen lijn

Score te verklaren door gedrag van kind***Kind kenmerken spelen een rol bij leerkrachtverwachting***

- Stagnatie door persoonlijke ontwikkeling
- ADHD heeft invloed op resultaten

Kijken naar hoe leerling in de klas is

- Leerling geeft niet aan dat stof moeilijk is
- Niet opletten in les valt op
- Werkhouding en nauwkeurigheid vallen op bij tegenvallende resultaten
- Inzet bepalend voor resultaten
- Zelfvertrouwen en concentratie van invloed op resultaten

Werkhouding tijdens Cito als verklaring voor score

- Blokkeren bij Cito zichtbaar in resultaten
- Kijken naar hoe leerling in z'n vel zit
- Onserieuze houding valt op

Verwachtingen over problematische groei van sterke leerlingen

Waarom maken leerkrachten zich (geen) zorgen?

- Sterke leerling: geen zorgen vanwege stijgende lijn
- Sterke leerling: geen zorgen want constant beeld
- Sterke leerling: minder groei niet erg want momentopname
- Sterke leerling: daling ook in reguliere lessen, dan eerder zorgen/actie

Verwachtingen op basis van niveau

- Sterke leerling: geen zorgen want I-score
- Sterke leerling: dalen in niveau niet erg

Sterke leerling: verwacht score weer omhoog

- Sterke leerling: verwacht score weer omhoog door extra ondersteuning
- Sterke leerling: verwacht score weer omhoog door motivatie

Verwachtingen over gedrag dat groei stimuleert

- Sterke leerling: gemotiveerd
- Sterke leerling: goede werkhouding
 - o Sterke leerling: graag goed doen
 - o Sterke leerling: steekt vinger op en durft hulp te vragen

Verklaringen voor problematische groei

- Sterke leerling: executieve functies kan problematisch zijn
- Sterke leerling: geen groei want interpreteert anders
- Sterke leerling: kan aan plafond zitten

Sterke leerling: aanbod in de gaten houden

- Sterke leerling: basisvaardigheden onderliggend probleem
- Sterke leerling: uitdaging nodig
- Sterke leerling: lesstof wordt moeilijker
- Sterke leerling: lesstof gaat te snel
- Sterke leerling: lesstof is niet voldoende geautomatiseerd

Verwachtingen over problematische groei van zwakke leerlingen

Oordeel over een lage groei bij een zwakke leerling

- Zwakke leerling: leerling scoort lager dan er in potentie in zit
- Zwakke leerling: randvoorwaarden zijn goed dus dit is wat erin zit
 - o Zwakke leerling: inzet is goed want mooie groei binnen eigen niveau
- Zwakke leerling: zorgen om groei want niveau gezakt

Verwachtingen over toekomstige groei

- Zwakke leerling: verwacht geen III-score
- Zwakke leerling: verwacht weer groei te zien
- Zwakke leerling: in de gaten houden dat leerling niet verder zakt

Zwakke leerling: meer winst te behalen door beter aansluiten op behoeften van leerling

- Zwakke leerling: meer betekenis geven aan rekenen
- Zwakke leerling: meenemen in verlengde instructie
- Zwakke leerling: moeilijke onderdelen oefenen
- Zwakke leerling: motivatie nodig
- Zwakke leerling: niet alles af maar wel weten hoe het moet
- Zwakke leerling: visuele ondersteuning nodig
- Zwakke leerling: concreet materiaal stimuleert concentratie

Verklaringen voor problematische groei

- Zwakke leerling: talige kant bij rekenen belemmert
- Zwakke leerling: executieve functies ontbreken
- Zwakke leerling: leerling vindt de lesstof lastiger
- Zwakke leerling: lengte Cito-toets kan hiaat zijn
- Zwakke leerling: zorgen of leerling toetsbaar is

Zwakke leerling: gedrag van invloed op resultaten

- Zwakke leerling: weinig *concentratie*
- Zwakke leerling: *aandacht* en *aanleg* spelen een rol
- Zwakke leerling: *gedemotiveerd*
 - o Zwakke leerling: breuken kan hem niets schelen
- Zwakke leerling: slechte *werkhouding* en inzet
 - o Zwakke leerling: werkt snel en slordig waardoor denkstappen overslaan