

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/19813> holds various files of this Leiden University dissertation.

**Author:** Stevenson, Claire Elisabeth

**Title:** Puzzling with potential : dynamic testing of analogical reasoning in children

**Issue Date:** 2012-09-13

# Summary in Dutch

## (Samenvatting)

Onderzoek naar de cognitieve ontwikkeling van kinderen laat zien dat grote verschillen niet alleen optreden in wat kinderen al kunnen maar ook in *hoe* ze leren. Dynamisch testen is een methode om cognitieve vaardigheden in ontwikkeling – zoals bijvoorbeeld het redeneervermogen – te meten. Het gaat bij dynamisch testen niet alleen om wat een kind al weet, maar vooral om zijn of haar vermogen om te leren (Elliott, 2003; Sternberg & Grigorenko, 2002). Dit doel onderscheidt dynamische tests van conventionele, statische tests, zoals intelligentietests. Ondanks het feit dat statische tests veelvuldig gebruikt worden als er vragen zijn over de schoolprestaties van een kind, zijn ze bekritiseerd omdat ze vooral de huidige cognitieve vaardigheden en niet zozeer het potentieel van een kind in kaart brengen. Met dynamische tests kan informatie over het cognitief potentieel en instructiebehoefte van een kind verkregen worden en dit kan belangrijk zijn voor keuzes over onderwijs (Bosma & Resing, 2012). Dynamisch testen onderscheidt zich van conventionele testsituaties omdat er training wordt gegeven in aanvulling op een of meer statische testmomenten. Zo kan worden nagegaan hoe en in hoeverre het kind leert gedurende het hele traject van voormeting tot en met nameting (Elliott

et al., 2010).

De doelstelling van dit promotieonderzoek was om een dynamische test voor basisschoolleerlingen te ontwikkelen die het leervermogen van een kind op het gebied van analogisch redeneren in kaart brengt. Analogietaken zijn gekozen omdat deze, vanwege de complexe wijze waarop zulke taken opgelost dienen te worden, vaak gebruikt worden in intelligentietests en omdat het analogisch redeneervermogen relevant is voor het schoolse leren (Goswami, 1992). Een eerste streven was inzicht te krijgen in de factoren die een rol spelen in de grote variatie in leervermogen van kinderen. In de verschillende studies in dit proefschrift zijn twee factoren onderzocht: (1) vorm van training – zowel type instructies als type opgaven en (2) persoonskenmerken zoals etniciteit en werkgeheugen. Een tweede streven was te bepalen of leervermogen gemeten met deze test schoolse prestaties voorspelt.

In hoofdstuk 1 zijn de algemene uitgangspunten van de ontwikkelde dynamische test en de factoren die mogelijk van invloed zijn op de prestatie op deze test besproken. Dynamisch testen werd in dit proefschrift opgevat als een methode gericht op het in kaart brengen van het cognitief potentieel en het leerproces van een kind tijdens een testafname. De in dit proefschrift gebruikte dynamische test, AnimaLogica, bestond uit een voortoets, gevolgd door een korte training en een natoets. De *voortoets* geeft een indicatie van het *huidige analogisch redeneervermogen* – een meting waarbij geen hulp of feedback wordt geboden (Resing, 1997). De voortoets wordt gevolgd door twee *trainingen* waarin het kind volgens de ‘graduated prompts’-methode getraind wordt. ‘Graduated prompting’ is een stapsgewijze trainingsmethode waarbij volgens een hiërarchisch principe zo weinig mogelijk hulp wordt geboden om het kind zo zelfstandig mogelijk de taak te laten oplossen (bijv. Campione & Brown, 1987; Resing, 1993; Resing & Elliott, 2011). Eerst werden algemene, metacognitieve instructies gegeven die het plannen stimuleerden of de

---

aandacht op de taak richtten. Daarna werd specifiekere hulp gegeven waardoor het kind steeds meer inzicht kreeg in hoe de taak opgelost diende te worden. Als deze stappen er nog niet toe leidden dat het kind de juiste oplossing vond, dan maakte de trainer de opgave samen met het kind. Het achterliggende idee is dat het kind alleen hulp krijgt als dat nodig is, zodat de instructiebehoefte van het kind gemeten kan worden. De *hoeveelheid* benodigde instructie tijdens de training geeft een indicatie van het leervermogen van het kind. De *typen* instructies die tot zelfstandige oplossingen hebben geleid geven een indicatie van welke instructies mogelijk ook op school effectief zouden kunnen zijn. De trainingssessies worden gevolgd door een *natoets*, die net zoals de voortoets zonder hulp of feedback werd afgenomen. De natoets geeft het *potentieel vermogen* aan – wat een kind mogelijk zou kunnen met geïndividualiseerde instructie. Het analyseren van de *uitleg* van het kind en welke *strategieën* hij of zij heeft toegepast geeft informatie over het leerproces – oftewel hoe het kind geleerd heeft tijdens de dynamische test (bijv. Resing et al., 2009). Het vermogen om de geleerde kennis spontaan op een nieuw probleem toe te passen, zogeheten *transfer*, geeft aan in hoeverre het kind na een korte interventie begrijpt wat analogisch redeneren is (bijv. Campione et al., 1985; Ferrara et al., 1986; Resing, 1990).

Twee facetten die het regelmatig gebruik van dynamische testen bemoeilijken zijn in AnimaLogica meegenomen: (1) de duur van een dynamische test en (2) de manier waarop verandering in prestatie wordt gemeten (Grigorenko & Sternberg, 1998). De afname van AnimaLogica is aanzienlijk korter dan bij eerdere dynamische tests het geval was – ongeveer 80 minuten – hetgeen overeenkomt met de duur van andere cognitieve tests. Dit is o.a. bereikt door het verkorten van de trainingstijd en gebruik te maken van een taak die gemakkelijk op de computer afgenomen kon worden (zie Stevenson et al., 2011 voor een bespreking van papieren versus computerafname). De psychometrische kwaliteit van dynamische tests is vaak

onduidelijk of wordt als onvoldoende beschouwd. De reden hiervoor is dat de klassieke wijze van het meten van de *mate van verandering* – door simpelweg het aantal goede oplossingen op de voortoets en de natoets te vergelijken – door psychometrici als onbetrouwbaar wordt beschouwd (De Bock, 1976). Een bijkomend probleem is dat een verschil van bijvoorbeeld vier juiste antwoorden een andere waarde kan hebben voor een kind dat oorspronkelijk twaalf opgaven goed had of een kind dat maar één opgave van de twintig goed had (Embretson, 1991b). Welk kind heeft meer geleerd? Door de problemen met betrouwbaarheid kan de mate van verandering op basis van ruwe scores beter niet gebruikt worden in het dynamisch testonderzoek (Resing, Elliott, & Grigorenko, 2012). Toch kan de mate van verandering mogelijk waardevolle informatie opleveren over het leervermogen als deze op een andere wijze – met behulp van item-respons theorie – wordt berekend (Embretson & Reise, 2000). Het hoofddoel van de (nog gaande) ontwikkeling van AnimaLogica was de instructiebehoefte en het potentieel van een kind te meten terwijl rekening werd gehouden met psychometrische standaarden en een korte, simpele afname.

Kinderen vertonen grote verschillen in zowel instructiebehoefte als mate van verandering in hun prestaties op een dynamische test (bijv. Resing et al., 2009). Met dynamisch testen wordt getracht deze verschillen te meten. Het doel van dynamisch testen is dus niet om blijvende verandering aan te brengen, maar om het leerpotentieel en het leerproces in kaart te brengen (Resing, Elliott, & Grigorenko, 2012). Het gemeten leervermogen wordt echter beïnvloed door factoren als de *vorm* van training en ook door *kenmerken van het kind*.

In hoofdstuk 2 is aandacht besteed aan de rol die de vorm van de opgaven speelt bij het verkrijgen van inzichten in het leervermogen van een kind. Kinderen uit groep 2 kregen ofwel ‘graduated prompts’-training met meerkeuzevragen, ofwel ‘graduated prompts’-training met open vragen waarbij het antwoord

---

geconstrueerd moest worden, ofwel geen training. De twee groepen die 'graduated prompts'-training kregen lieten na training meer progressie in analogisch redeneren zien dan de controlegroep. Dit schetst het beeld dat de 'graduated prompts'-training gemiddeld gezien een effectieve manier is om het analogisch redeneren van vijf- en zesjarigen te stimuleren. Er was geen verschil in de mate van vooruitgang van voortoets naar natoets tussen de twee trainingsgroepen, maar de 'antwoordconstructiegroep' kon gemiddeld gezien wel betere uitleg geven van hun antwoorden en lieten een ander strategiegebruik zien dan de kinderen in de 'meerkeuzegroep'. Als antwoord op de hoofdvraag of trainen met meerkeuze- dan wel met antwoordconstructieopgaven het meest geschikt zou zijn voor dynamisch testonderzoek werd geconcludeerd dat antwoordconstructie een specifiek inzicht geeft in het redeneerproces van de kinderen. Daarom werd gekozen voor antwoordconstructieopgaven in de dynamische test afgenomen in het vervolgonderzoek.

In hoofdstuk 3 werd onderzocht of de ontwikkelde dynamische test geschikt is voor zowel autochtone als allochtone leerlingen. Bij traditioneel afgenomen intelligentietests zijn kinderen van de dominante cultuur over het algemeen in het voordeel (bijv. Van de Vijver, 2002). Dit kan bijvoorbeeld komen door verschillen in taalvaardigheid of verschillen in ervaring met soortgelijke opgaven of testsituaties. Deze problemen kunnen ertoe leiden dat er een vertekend beeld ontstaat van de huidige vermogens en het leerpotentieel van etnische minderheden vergeleken met die van hun autochtone leeftijdsgenoten (Sternberg et al., 2002). Item-respons theorie werd toegepast voor het meten van vooruitgang en werd rekening gehouden met onder andere de persoonsfactor werkgeheugen. In dit onderzoek waren autochtone en allochtone leerlingen verdeeld over drie groepen: 'graduated prompts'-training, zelfstandig oefenen met de opgaven of geen training (controle). Er werden geen verschillen gevonden tussen autochtone en allochtonen leerlingen in de

‘graduated prompts’-groep in mate van vooruitgang, strategiegebruik, behoefte aan instructie of uitleg van hun oplossingen tijdens de training. Hieruit werd geconcludeerd dat de dynamische test ingezet kan worden voor het meten van leervermogen bij cultureel-diverse schoolpopulaties. Werkgeheugen bleek niet tussen beide leerlinggroepen te verschillen, en was bij beide groepen gerelateerd aan het analogisch redeneervermogen.

In hoofdstuk 4 werd de samenhang tussen twee vormen van werkgeheugen, het verbale en visuo-spatiele werkgeheugen, en de prestaties op de dynamische test onderzocht. De focus hierbij lag op transfer – oftewel het spontaan kunnen toepassen van hetgeen tijdens de trainingen is geleerd op andere, gerelateerde opgaven (Jacobs & Vandeventer, 1971). Twee groepen leerlingen, verdeeld over een ‘graduated prompts’-trainingsgroep en een controle groep dat oefende met dezelfde opgaven, participeerden in het onderzoek. Bij de voormeting en nameting werden naast de analogieën met dierenfiguren ook twee andere redeneertaken, plus een ‘reversal’ taak afgenomen, waarbij het kind opgaven dient te ontwerpen voor de trainer en uitleg moet geven hoe de taken opgelost kunnen worden (Bosma & Resing, 2006). De kinderen die beter presteerden op de voormeting bleken over het algemeen een efficiënter werkgeheugen te hebben. De *mate* van vooruitgang bleek echter geen verband te houden met werkgeheugen. De prestaties op de transfertaken bij de nameting waren enigszins gerelateerd aan de prestaties op de voormeting. Bij transfer speelde het werkgeheugen wederom geen rol. Redeneervermogen en werkgeheugen zijn twee constructen die vaak gemeten worden als een schoolpsycholoog inzicht wil krijgen in de cognitieve capaciteiten van een kind. Deze constructen bleken weinig samenhang te vertonen met maten voor leervermogen en transfer. Dit betekent dat leervermogen en transfer belangrijk zouden kunnen zijn bij het in kaart brengen van het cognitief potentieel van een kind.

---

In hoofdstuk 5 werd dieper ingegaan op de meting van de mate van vooruitgang tussen de voormeting en nameting en de samenhang hiervan met het werkgeheugen. Het meten van verandering wordt door psychometrici onbetrouwbaar geacht wanneer er sprake is van verschillen in percentage goed tussen de voor- en nameting (Lord, 1963). Item-respons theorie biedt mogelijkheden om de mate van vooruitgang op betrouwbare wijze te meten (Embretson & Reise, 2000). Item-respons theorie werd in de studie in dit hoofdstuk gebruikt om niet alleen mate van vooruitgang te meten maar ook binnen eenzelfde statistisch model de verschillen in leervermogen tussen kinderen te verklaren aan de hand van enerzijds het type training dat werd gegeven plus anderzijds persoonskenmerken zoals werkgeheugen. Basisschoolleerlingen uit groepen twee, drie en vier waren verdeeld over twee trainingcondities: 'graduated prompts' en feedback. Bij de feedbacktraining kreeg het kind net zoals in de 'graduated prompts' training vijf kansen om het goede antwoord te construeren. In tegenstelling tot de 'graduated prompts' getrainde kinderen kregen ze geen instructies over hoe ze dat moesten doen, maar kregen ze alleen te horen of hun antwoord goed of fout was. De mate van vooruitgang van de kinderen in de feedbackconditie bleek minder sterk te zijn dan die van de kinderen in de 'graduated prompts'-groep. In beide gevallen was het werkgeheugen geen verklarende factor voor de mate van vooruitgang. De prestaties op de voortoets hingen samen met leeftijd, maar leeftijd was geen verklarende factor van de individuele verschillen in vooruitgang. Er was enig verband te zien tussen de prestaties van een kind op de voormeting en zijn mate van vooruitgang bij de natoets. Dit gaf echter geen volledig beeld van het leervermogen. Wel bleek dat kinderen die hoge scores behaalden op landelijke rekentoetsen ook beter presteerden op de voormeting en ook meer vooruitgang lieten zien tijdens het dynamisch testen. Dit ondersteunt eerdere conclusies dat de mate van vooruitgang mogelijk een belangrijk construct vormt bij het meten van het cognitief potentieel van een kind



(bijv. Embretson & Prenovorst, 2000).

De voorspellende waarde van de mate van vooruitgang gemeten met AnimaLogica op schoolprestaties werd onderzocht in hoofdstuk 6. Conventionele tests, zoals een intelligentietest, hebben enige voorspellende waarde ten aanzien van toekomstige schoolprestaties (Sternberg et al., 2001). Dynamische testuitkomsten lijken van toegevoegde waarde te zijn ten aanzien van deze voorspelling (Caffrey et al., 2008). Het is echter niet duidelijk welk aspect van de metingen het meeste bijdraagt aan de voorspelling: de instructiebehoefte tijdens de training, de mate van vooruitgang of het transfervermogen. Dit onderzoek bouwt voort op eerdere onderzoek (bijv. Beckmann, 2006; Resing, 1993), maar voegde ook drie aspecten toe: (1) de opzet was longitudinaal, (2) de testgroep bestond uit reguliere basisschoolkinderen, en (3) de voorspelling was op nationaal genormeerde toetsen toegepast. Kinderen uit groep drie van de basisschool werden dynamisch getest met de 'graduated prompts'-methode. Van elk kind zijn de prestaties voor rekenen en lezen, afkomstig uit de gegevens van het leerlingvolgsysteem, verzameld op drie momenten: 3 weken voor het dynamisch testen, 5 maanden na het dynamisch testen en 1 jaar later. In deze studie zijn conventionele en dynamische testgegevens vergeleken bij het voorspellen van de scores op rekenen en lezen. De prestatie op de voortoets werd beschouwd als een conventionele meting van het analogisch redeneren. De mate van vooruitgang, instructiebehoefte tijdens training en prestaties op de 'reversal' transfertaak waren de dynamische metingen. De mate van vooruitgang bleek de beste voorspeller van de scores van de kinderen in zowel rekenen als lezen. Deze vondst gaf verdere ondersteuning voor de hypothese dat dynamische testgegevens van toegevoegde waarde kunnen zijn bij het in kaart brengen van het cognitief potentieel van een kind.

Ten slotte werd in hoofdstuk 7 geconcludeerd dat hoewel *huidige prestaties* op analogietaken beïnvloed worden door persoonskenmerken zoals leeftijd,

---

werkgeheugen en etniciteit, deze factoren het *leervermogen* gemeten met een dynamische test niet verklaren. Het type training dat gegeven wordt heeft echter wel invloed op de mate van vooruitgang van de voormeting naar de nameting. Uit het onderzoek gepresenteerd in dit proefschrift blijkt dat 'graduated prompting' tot grotere vooruitgang leidt dan feedbacktraining of zelfstandig oefenen. Ook de vorm van de opgaven speelt een rol waarbij antwoord-constructie of zelfs opgavencreatie, zoals in de 'reversal' taak, een meer volledig beeld van het leerpotentieel geven.

De algemene conclusie in dit proefschrift is dat uitkomsten op een dynamische test met een 'graduated prompts'-training van toegevoegde waarde kunnen zijn wanneer onderwijzers zich afvragen wat het leerpotentieel van een kind is. Het geeft mogelijk ook een eerlijker beeld van het leervermogen van allochtone leerlingen die op conventionele tests in het nadeel zijn. Er blijft echter grote variabiliteit in de prestaties en vooruitgang van kinderen op analogisch redeneertaken. De individuele verschillen in leervermogen gemeten met deze dynamische test zijn niet eenduidig of gemakkelijk te verklaren uit persoonskenmerken of trainingsvorm, maar ze geven wel informatie die toekomstige schoolprestaties in rekenen en lezen kan helpen voorspellen. Een dynamische test zou ingezet kunnen worden om na te gaan welke kinderen meer potentieel hebben dan wat op dit moment uit de schoolprestaties blijkt. Ook zou een dynamische test vroegtijdig kunnen signaleren welke kinderen dreigen achter te lopen. Hierbij zou instructiebehoefte en transfervermogen waardevolle informatie kunnen bieden zodat onderwijzers een passende interventie kunnen ontwikkelen dat een kind helpt zijn of haar cognitief potentieel optimaal te benutten.

