

Toetsen en normeringen: enkele misverstanden opgehelderd

Jacqueline Visser

is marktgroepmanager.
E-mail: klantenservice@cito.nl

In dit artikel bespreken we vier misverstanden die heersen rondom de update van de normen van de LVS-toetsen van Cito aan het begin van het schooljaar 2013-2014.

De update van de normen van de LVS-toetsen van Cito aan het begin van het schooljaar 2013-2014 heeft nogal wat stof doen opwaaien in onderwijsland. Uit de vele vragen en reacties die we kregen bleek dat er verschillende misverstanden leven rondom toetsen en normeringen. In dit artikel bespreken we vier van deze misverstanden. Alvorens we dat doen, leggen we uit wat normen zijn en waarom ze soms aangepast moeten worden.

Wat is een norm?

Als een leerling een toets heeft gemaakt, dan kan de leerkracht bepalen hoeveel opgaven een leerling goed heeft gemaakt. Het aantal goed op een toets wordt veelal de ruwe toetscore genoemd. Stel nu dat Lynn, een leerling uit groep 5, 24 opgaven goed heeft van de in totaal 30 rekenopgaven van een toets. Wat betekent dat dan? Vragen die u daarbij kunt stellen, zijn: hoe goed kan Lynn rekenen? Rekent ze goed genoeg om door te kunnen in de methode? Is haar re-

kenvaardigheid voldoende om straks naar het havo te kunnen? Om betekenis te geven aan toetscores, worden de meeste toetsen genormeerd. Daarbij onderscheiden we twee soorten normeringen: relatieve en absolute. Met *relatieve normen* kun je de toetscore van een leerling vergelijken met de toetscores van andere leerlingen uit dezelfde doelgroep. In het geval van Lynn kun je op basis van een relatieve norm bepalen of zij met 24 goed de toets beter, even goed of slechter heeft gemaakt dan de leerlingen uit de doelgroep 'groep 5 van Nederland'. Vrijwel alle LVS-toetsen van Cito zijn relatief genormeerd, met de bekende vaardigheidsniveaus A t/m E en I t/m V.

Deze niveaus zijn gemaakt door de toetscores in de normgroep te verdelen in gelijke delen. Zo zit in niveau I de 20% hoogst scorende leerlingen en in niveau V de 20% laagste scorende (zie tabel). Deze niveaus geven de leerkracht een bredere benchmark voor de interpretatie van de score

Om betekenis te geven aan toetscores, worden de meeste toetsen genormeerd

van Lynn dan de eigen groep. Het kan best zo zijn dat Lynn met 24 goed de toets in de eigen groep het slechtst heeft gemaakt, maar dat haar score vergeleken met alle leerlingen van groep 5 in Nederland boven het gemiddelde ligt.

Met *absolute normen* vergelijk je de toetscore van een leerling met een beheersingsstandaard. Die standaard wordt gebruikt om te kunnen bepalen welke leerlingen over voldoende kennis en vaardigheden beschikken om bijvoorbeeld een bepaalde studie te volgen. De meest bekende beheersingsstandaard is de score (cesuur) die de grens vormt tussen voldoende/geslaagd en onvoldoende/gezakt bij de examens. De referentieniveaus voor de Nederlandse taal en rekenen zijn eveneens beheersingsstandaarden. Voor de LVS-toetsen Rekenen-Wiskunde en Begrijpend lezen zullen we begin 2015 aangeven bij welke vaardigheidsscores de referentieniveaus 1F en 1S/2F horen.

Waarom moeten normen soms aangepast worden?

Het is een bekend gegeven dat normen aan slijtage onderhevig zijn. Wat betreft leervorderingen-toetsen, zoals de LVS-toetsen, zijn er signalen dat de bijbehorende normen mogelijk sneller verschuiven dan we tot nu toe hebben aangenomen. In de literatuur worden diverse mogelijke oorzaken genoemd:
) veranderingen in de leerlingpopulatie waardoor het (gemiddelde) niveau verandert;

I - V		A - E	
20% hoogst scorende leerlingen	I 20%	A 25%	25% hoogst scorende leerlingen
20% boven het landelijk gemiddelde	II 20%	B 25%	25% ruim boven tot net boven het landelijk gemiddelde
20% landelijk gemiddelde	III 20%	C 25%	25% net tot ruim onder het landelijk gemiddelde
20% onder het landelijk gemiddelde	IV 20%	D 15%	15% ruim onder het landelijk gemiddelde
20% laagst scorende leerlingen	V 20%	E 10%	10% laagst scorende leerlingen

› veranderingen in het onderwijsaanbod: sommige (onderdelen van) vakgebieden worden (weer) belangrijker, andere verschuiven naar de achtergrond;

› het gebruik van volgsystemen zelf: leerkrachten zijn zich bewuster van hiaten bij (groepen) leerlingen en passen daarop hun aanpak aan. Ze remediëren hun probleemleerlingen beter en hun onderwijskundig-didactische aanpak wordt effectiever;

› het langzamerhand bekend worden van de inhoud van de toetsen waardoor het onderwijsaanbod bewust of onbewust wordt aangepast op die kennis;

› toename van het belang van de resultaten, zowel voor de leerling als voor de leerkracht.

Wat dat laatste betreft: in het onderwijs spelen toetsresultaten een steeds grotere rol in de interne en externe verantwoording die scholen afleggen, waardoor bij scholen de druk toeneemt om hun onderwijsprestaties te verbeteren. Dat betekent dat het belang van leervorderingstoetsen toe neemt: voor de leerling, de leerkracht en de school. En dat heeft weer tot gevolg dat leerlingen beter hun best doen om de toets te maken, en de leerkracht en de school hun onderwijs meer richten op het betreffende deel van het curriculum.

In ons eigen onderzoek zien we grofweg twee trends in de resultaten op LVS-toetsen¹:

1. een verschuiving vrij snel na de uitgave van een toets, en
2. een geleidelijke verschuiving in de loop van de tijd.

Deze trends zijn niet voor alle leergebieden en groepen hetzelfde. Waarom dat zo is, weten we niet. Wel denken we dat de volgende factoren daarbij een rol spelen:

1. mate waarin leerlingen de toets tijdens een normeringsonderzoek ervaren als een échte afname die meetelt of als een proef: leerlingen zullen méér gemotiveerd zijn voor een echte afname;

2. de mate waarin een vaardigheid gemakkelijk leerbaar is en/of toetsinhouden (on)bewust gemakkelijk geoefend kunnen worden;

3. het belang dat aan een toets of vaardigheid gehecht wordt in het kader van opbrengstgericht werken en het afleggen van verantwoording: hoe belangrijker, des te meer aandacht de school aan een vaardigheid zal besteden.

De eerste factor heeft met name effect op het verschil tussen de normeringsgegevens en het eerste jaar van afname. De twee andere factoren kunnen beide gesignaleerde trends veroorzaken.

Enkele misverstanden nader verkend

Cito heeft de normen verhoogd in opdracht van de inspectie

De afgelopen tijd hebben we verschillende motieven gehoord over waarom Cito de normen van de LVS-toetsen heeft verhoogd. 'In opdracht van de inspectie' is er één van, net als 'zodat de scholen de nieuwe toetsen gaan kopen'. In het verlengde vragen sommigen zich af op basis waarvan we de normen hebben aangepast: is dat 'zomaar gebeurd?' Laten we heel helder zijn: we hebben de normen van onze toetsen niet zomaar aangepast.

Al langer kregen we signalen, ook van scholen en besturen, dat de leerlingen de LVS-toetsen beter maakten dan gedacht. Anders gezegd: het percentage leerlingen dat bovengemiddeld scoorde lag duidelijk hoger dan 50%. Ons onderzoek op basis van gegevens van duizenden scholen die onze toetsen gebruiken, bevestigde dat beeld: leerlingen maken vandaag de dag de toetsen beter dan de leerlingen in de normeringsonderzoeken van jaren geleden.

We hadden het daar natuurlijk bij kunnen laten: het is toch plezierig voor iedereen dat de leerlingen de toetsen zoveel beter maken? Waarom hebben we de normen dan toch aangepast en de scholen 'gestraft voor betere

resultaten'? Omdat de oude normen leiden tot een verkeerde interpretatie van de gegevens. Zoals in de inleiding gezegd: A t/m E en I t/m V zijn *relatieve* normen die aangeven hoe goed een leerling het doet ten opzichte van een referentiegroep. Op basis van de oude normen zou u ten onrechte kunnen concluderen dat uw school en uw leerlingen het prima doen ten opzichte van de andere basisscholen in Nederland. Door de nieuwe normen beschikt u weer over een juiste benchmark voor uw leerlingen en uw school. Dat voorkomt minder prettige verrassingen zoals tegenvallende scores op de Eindtoets. Bijkomend voordeel van de nieuwe normeringen is dat deze beter aansluit bij de criteria die de inspectie gebruikt bij het beoordelen van de toetsresultaten van de scholen.

De LVS-toetsen brengen de verschillen tussen leerlingen in beeld om hen te kunnen indelen in 20% beste t/m 20% zwakste

Het doel van de LVS-toetsen van Cito is niet het sorteren van leerlingen in de niveaus A-E of I-V. Het primaire doel van de toetsen is het in kaart brengen van de vaardigheid van leerlingen, zodat leerkrachten hun ontwikkeling kunnen volgen in de tijd en het onderwijs op hoofdlijnen kunnen afstemmen op de vaardigheden van hun leerlingen. Vanwege dit doel is de inhoud van de LVS-toetsen breder dan het onderwijsaanbod in een bepaald vak. Immers, hoe je het ook wendt of keert: de betere leerlingen kennen en kunnen vaak al meer dan wat in de methode staat. Als een school ook voor deze leerlingen uitdagend onderwijs wil bieden, is het niet voldoende als een leerkracht weet dat 'ze alles al kunnen wat de afgelopen zes weken in de methode aan bod is gekomen'. De kans is groot dat de betere leerlingen veel méér kunnen. Vandaar dat in de toetsen van Cito ook opgaven zitten die

moelijker zijn dan hetgeen in het onderwijs is aangeboden. Waar komt het misverstand nu vandaan dat de LVS-toetsen alleen maar bedoeld zijn om leerlingen in categorieën te plaatsen? Veel scholen en leerlingadministratiesystemen beperken hun rapportages over toetsen tot het benoemen in welk vaardigheidsniveau leerlingen zitten, bijvoorbeeld niveau I (= de 20% beste leerlingen) of E (= de 10% zwakste leerlingen). In zo'n geval lijkt het inderdaad of het ordenen van leerlingen van sterkste naar zwakste, het enige is wat er toe doet. Wij benadrukken in onze rapportages juist het gebruik van vaardigheidsscores om daadwerkelijk naar de ontwikkeling van een kind in de tijd te kijken en daar als leerkracht je plan op te trekken. En dan zie je bijvoorbeeld dat die zwakke leerling wel steeds vaardiger wordt, of dat die sterke leerling nauwelijks beter is geworden in een bepaald vak.

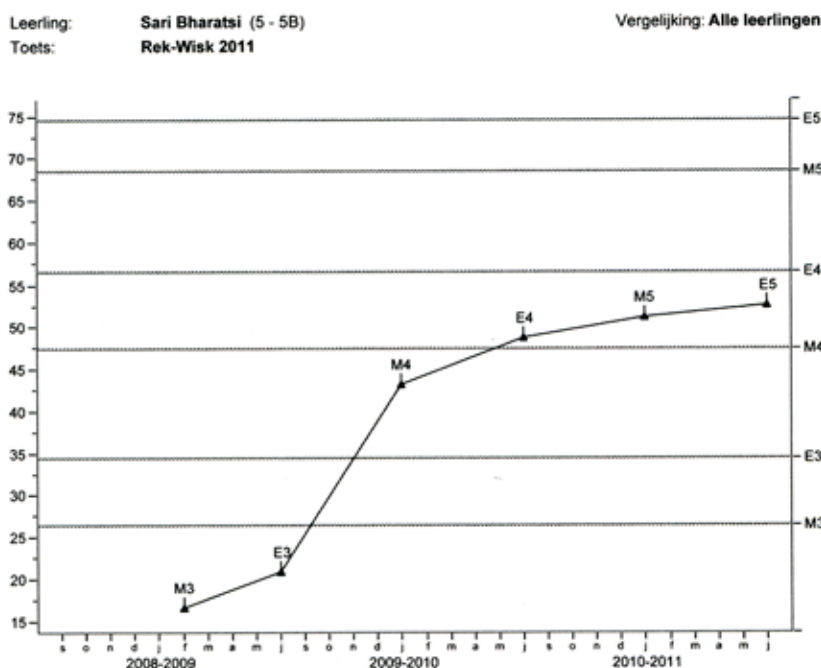
Met de LVS-toetsen breng je niet de opbrengsten van je onderwijs in kaart, maar ben je vanaf groep 3 al bezig met het uitstroomprofiel naar het VO
Het klopt inderdaad dat je met de

LVS-toetsen niet sec de opbrengsten van je onderwijs in kaart brengt. Primaire doel van de LVS-toetsen is het in beeld brengen van de vaardigheid en de groei daarvan bij leerlingen, zowel van de zwakkere als van de sterkere leerlingen. Om op basis daarvan het onderwijsaanbod af te stemmen op de vaardigheden van de leerlingen. Het is zeker niet de bedoeling om al op basis van toetsresultaten in groep 3 te gaan werken richting een uitstroomprofiel in het VO. Leerlingen ontwikkelen zich niet lineair² en er kan nog veel gebeuren in de zes jaren tussen groep 3 en groep 8 waardoor leerlingen ineens meer kunnen dan van te voren gedacht. Gelukkig maar, want daar is goed onderwijs immers voor bedoeld!

Soms horen we de opmerking dat je met methodetoetsen wél de opbrengsten van het onderwijs in beeld kunt brengen. De misvatting die daaronder ligt, is het idee dat je door precies die stof te toetsen die aan bod is gekomen in de methode, meet wat je als onderwijs hebt bereikt. In werkelijkheid is het echter niet zo makkelijk om te bepalen wat de exacte opbrengst van het onderwijs is, immers:

) als je alleen dat toetst wat een kind net heeft geleerd, weet je niet of sprake is van transfer van het geleerde en of een leerling een vaardigheid ook op langere termijn nog beheerst;
) kinderen leren niet alleen op school en ze leren ook niet alleen dat wat de leerkracht ze aanbiedt; diverse 'schoolse' vaardigheden pikken ze bijvoorbeeld ook thuis op.
De derde stelling stipt overigens wel een concreet gevaar aan. Als scholen voor VO strakke eisen gaan stellen aan de relatieve scores (A-E; I-V) van leerlingen op LVS-toetsen voor de plaatsing op een bepaald onderwijstype, dan is het risico dat LVS-toetsen alleen maar als sorteermachine gaan fungeren. Met allerlei mogelijke uitwassen die ook rondom de Eindtoets zijn ontstaan. Denk dan aan trainingsbureautjes die leerlingen laten oefenen op de LVS-toetsen. Dat zou heel erg zonde zijn, omdat daarmee de waarde van de toetsresultaten voor het onderwijs op de basisschool afneemt.

De LVS-toetsen zijn veel strenger dan de methodetoetsen
Vaak horen we de verzuchting 'op een methodetoets doet een kind het goed, op de Cito-toetsen scoort hij onvoldoende'. Het klopt inderdaad dat kinderen die vrijwel alles goed hebben op een methodegebonden toets, niet per definitie ook alle opgaven van de LVS-toetsen van Cito goed hebben. Dit hangt samen met het doel dat verschilt tussen beide type toetsen. Methodetoetsen zijn bedoeld om na te gaan wat een leerling heeft geleerd en onthouden van de voorgaande lessen. De meeste leerlingen zullen de opgaven van deze toetsen goed maken. Als een leerling nog te veel fouten maakt, zal een leerkracht instructie inzetten gericht op het halen van de betreffende lesdoelen. Het doel van de methodeonafhankelijke toetsen van Cito gaat verder dan het be-



palen 'of leerlingen de stof die ze in de afgelopen lessen hebben gemaakt een week later nog steeds kennen'. De inhoud van de toetsen is zodanig dat de vaardigheid van een leerling in de volle breedte in kaart wordt gebracht. Getoetst wordt of leerlingen dat wat ze hebben geleerd, ook kunnen toepassen in opgaven die niet van te voren getraind of geoefend zijn. Dat maakt het mogelijk om te bepalen of leerlingen vaardiger worden, hoe vaardig ze zijn vergeleken met leeftijdsgenoten en welke leerlingen iets extra's aankunnen.

De misvatting die hierbij naar boven komt, is dat als een leerling

op een LVS-toets onder het gemiddelde scoort, dit een onvoldoende score zou zijn. Dit soort termen gebruiken wij zelf helemaal niet. Maar er zijn administratiesystemen en scholen die onze niveaus A t/m E hebben vertaald naar 'goed t/m onvoldoende'. Met als uiterste consequentie dat alle leerlingen in Nederland die onder het landelijk gemiddelde scoren, een onvoldoende zouden behalen, hetgeen uiteraard heel vreemd is en naar onze mening ook onjuist.

Tot slot

In dit artikel heb ik geprobeerd meer duidelijkheid te geven over

de normeringen en specifiek de normen van de LVS-toetsen van Cito. Wilt u meer weten over toetsen? Via de website van Cito kunt u de publicatie Toetsen op school Primair Onderwijs downloaden³.

Noten

1 Zie Jos Keuning, Maartje Hilde en Anke Weekers (2014). Begrijpend leesprestaties onderzocht – Een analyse op basis van Cito dataretour. In: Tijdschrift voor Orthopedagogiek, nummer 53, p. 2-13.

2 Jos Keuning en Jacqueline Visser (2013). De (on)nauwkeurigheid van een leerrendementsverwachting: Ervaringen met de LVS-toetsen Rekenen-Wiskunde. In: Tijdschrift voor Orthopedagogiek, 52, p. 164-176.

3 Judith Hollenberg en Marleen van der Lubbe (2011) Toetsen op school Primair onderwijs. Arnhem: Cito.