



Inspectie van het Onderwijs
*Ministerie van Onderwijs, Cultuur en
Wetenschap*

FACTSHEET LEESVAARDIGHEID EINDE (S)BO 2020 - 2021

Technische Toelichting

Juni 2022

INHOUD

Inleiding en leeswijzer 3

- 1 Algemene toelichting Peil.Leesvaardigheid 4**
- 1.1 De steekproef 4
- 1.2 Periode van afname 4
- 1.3 Van kerndoelen en referentieniveaus naar instrument. 4
- 1.4 Het instrumentarium 6
- 1.4.1 De leestoets 6
- 1.4.2 De vragenlijsten 7

- 2 De leerlingresultaten uit de factsheet 8**
- 2.1 Beheersing referentieniveaus leesvaardigheid 8
- 2.2 Relatie met achtergrondkenmerken 8
- 2.3 Trend over tijd 10

- 3 Referenties 12**

- 4 Bijlage 1. 14**

Inleiding en leeswijzer

De resultaten beschreven in de factsheet 'Leesvaardigheid einde (s)bo 2020-2021' zijn afkomstig uit het gelijknamige peilingsonderzoek naar leesvaardigheid aan het einde van (speciaal) basisonderwijs ((s)bo). Deze peiling is onder regie van de Inspectie van het Onderwijs uitgevoerd door een consortium bestaande uit het Expertisecentrum Nederlands, KBA Nijmegen, Cito en de Radboud Universiteit Nijmegen. Met de peilingsonderzoeken brengt de inspectie een leerdomein in de volle breedte in kaart. Met de uitkomsten van de onderwijspeilingen weten we wat kinderen leren op school, hoe deze resultaten zich in de tijd ontwikkelen en hoe deze zich verhouden tot het onderwijsaanbod. De uitkomsten vormen input voor een brede dialoog over de inhoud, kwaliteit en het niveau van het onderwijs op de betreffende leerdomeinen.

De factsheet vormt een voorloper op het volledige rapport Peil.Leesvaardigheid dat in december 2022 wordt gepubliceerd en richt zich specifiek op de prestaties van leerlingen aan het einde van het s(b)o op het gebied van begrijpend lezen, hoe deze prestaties zich hebben ontwikkeld over tijd en op de belangrijkste samenhangen van de leesprestaties met leerlingkenmerken. Het doel van de factsheet is om scholen al voor de zomervakantie van 2022 te informeren over hoe de prestaties op dit leerdomein ervoor staan en hen aangrijpingspunten mee te geven om het lezen en de leesvaardigheid van leerlingen verder te stimuleren.

Leeswijzer

In deze technische toelichting geven we eerst een algemene toelichting op de steekproef voor dit peilingsonderzoek, op de constructie van de gebruikte instrumenten en op de wijze van dataverzameling. In het tweede deel verantwoorden we de specifieke uitkomsten die in de factsheet centraal staan.

1 Algemene toelichting Peil.Leesvaardigheid

In dit hoofdstuk wordt een algemene toelichting gegeven op hoe het peilingsonderzoek Leesvaardigheid einde (s)bo is opgezet. Daarbij worden enkele aandachtspunten meegegeven die van belang zijn voor de duiding van de leerlingresultaten.

1.1 De steekproef

De gegevens die we in deze factsheet presenteren, zijn in het voorjaar van 2021 verzameld op een representatieve steekproef¹ van 110 basisscholen (180 klassen) en 41 sbo-scholen (86 klassen). In totaal deden 3318 bo-leerlingen en 1000 sbo-leerlingen mee aan het onderzoek. Op 105 basisscholen werd de dataverzameling voor het huidige peilingsonderzoek gecombineerd met de dataverzameling voor Progress in International Literacy Study (PIRLS) 2021, het internationale onderzoek naar begrijpend lezen in groep 6 van de basisschool².

1.2 Periode van afname

Op het moment van dataverzameling, hadden scholen net de tweede sluiting ten gevolge van de maatregelen tegen de verspreiding van het coronavirus achter de rug. Deze maatregelen hadden een aantal gevolgen voor de werving en de praktische uitvoering van het onderzoek. Zo stond de scholenwerving oorspronkelijk gepland voor het voorjaar van 2020. Er is echter besloten pas na de zomervakantie van 2020 te starten met de scholenwerving, omdat scholen voor de zomervakantie nog (te) druk waren met de herstart na de eerste sluiting ten gevolge van de coronapandemie. De vooraf opgestelde wervings- en afnameprocedures zijn niet veranderd, maar het onderzoek vond wel plaats in een bijzondere periode. Bij de interpretatie van de resultaten moet de uitzonderlijke situatie waarin de resultaten verzameld zijn dan ook in ogenschouw worden genomen. Hoewel de invloed van de coronacrisis op de leesprestaties niet rechtstreeks uit de resultaten van dit peilingsonderzoek is af te leiden, toont monitoringsonderzoek aan dat de leergroei met betrekking tot begrijpend lezen in de afnameperiode relatief laag was, maar zich nu weer heeft hersteld (Haelermans et al., 2022).

1.3 Van kerndoelen en referentieniveaus naar instrument.

Centraal uitgangspunt voor het peilingsonderzoek vormen de kerndoelen lezen (Ministerie van OCW, 2006). In de kerndoelen komen de doelen voor lezen terug in de domeinen schriftelijk taalonderwijs en taalbeschouwing (waaronder strategieën). Het betreffen globaal geformuleerde doelen die voorschrijven waar het leesonderwijs zich op zou moeten richten. De Expertgroep Doorlopende Leerlijnen Taal en Rekenen heeft in 2009 in opdracht van het ministerie van Onderwijs, Cultuur en Wetenschap voor taal en rekenen referentieniveaus opgesteld die de beoogde opbrengsten van het onderwijs definiëren: het Referentiekader Taal en Rekenen. Deze referentieniveaus zijn in 2010 vastgesteld (Staatsblad van het Koninkrijk der Nederlanden, 2010). De referentieniveaus beschrijven wat leerlingen, van de

1 Omdat grote basisscholen enigszins oververtegenwoordigd waren in de steekproef, zijn er aanvullende analyses uitgevoerd om een mogelijke vertekening in de resultaten na te gaan. De conclusie uit deze analyses is, dat hoewel de onderzoeksgroep qua verdeling in schoolgrootte verschilt van de populatie, dit geen invloed heeft gehad op de leesprestaties. Daarom (en om complexiteit niet onnodig te verhogen) is weging van de scores niet noodzakelijk bevonden.

2 Publicatie van het onderzoeksrapport van PIRLS zal gelijktijdig met het publieksrapport Peil.Leesvaardigheid plaatsvinden.

basisschool tot aan het hoger onderwijs, moeten kennen en kunnen. De niveaus worden beschreven in twee 'kwaliteiten':

- de fundamentele kwaliteit (F)
- de streefkwaliteit (S)

Het eerste fundamentele beheersingsniveau (1F-niveau) en het eerste streefniveau (1S-niveau) zijn van toepassing op het einde van het (speciaal) basisonderwijs. Daarbij geldt dat het 1S niveau tevens het 1F-niveau omvat (OCW, 2009c; Staatsblad 2010-265). Niveau 1S wordt gezien als het algemeen maatschappelijk niveau dat minimaal nodig is om goed te kunnen participeren in de maatschappij. De inschatting was dat, op het moment van het formuleren van de referentieniveaus, aan het einde van het primair onderwijs 1F haalbaar was voor 75% van de leerlingen en 1S voor 50% van de leerlingen. Bij de invoering van de referentieniveaus is de ambitie gesteld dat respectievelijk 85% van de leerlingen het 1F-niveau zou gaan beheersen en 65% het 1S niveau (Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2008b).

In het referentiekader valt het leerdomein leesvaardigheid uiteen in het lezen van zakelijke teksten en het lezen van fictionele, narratieve en literaire teksten (vanaf hier en in de factsheet: fictieve teksten). Aan deze leestaken zijn kenmerken van de taakuitvoering (hierna: leesdomeinen) verbonden die als criteria gelden voor wat leerlingen op de verschillende niveaus moeten kennen en kunnen. Voor het lezen van fictieve teksten gaat het om begrijpen, interpreteren en evalueren. Voor het lezen van zakelijke teksten gaat het om techniek en woordenschat, begrijpen, interpreteren, evalueren, samenvatten en opzoeken.

Domeinbeschrijving en review

Voor het peilingsonderzoek "Leesvaardigheid einde (s)bo 2020-2021" is door SLO een domeinbeschrijving gemaakt (SLO, 2019). In deze domeinbeschrijving staan de wettelijke eisen voor leesvaardigheid, hoe deze terugkomen in uitwerkingen daarvan zoals de tussendoelen, leerlijnen en lesmethoden en welke resultaten eerdere peilingen op het gebied van leesvaardigheid opleverden. Ook is in de domeinbeschrijving, op basis van een veldraadpleging met domeinexperts, taalcoördinatoren, schoolleiders en leerkrachten, gespecificeerd welke inhoud minimaal onderdeel moesten zijn van het peilingsonderzoek. Naast de domeinbeschrijving is er ten behoeve van het peilingsonderzoek ook een review opgesteld (Houtveen et al., 2019). Deze review brengt in kaart wat er vanuit wetenschappelijk onderzoek bekend is over de samenhang tussen kenmerken van het onderwijsleerproces op het gebied van leesvaardigheid en de leesprestaties en leesmotivatie. Ook geeft deze review aanbevelingen over de kenmerken uit het onderwijsleerproces leesvaardigheid die ten minste in beeld moeten worden gebracht.

Voor het ontwikkelen van de instrumenten voor het meten van de leesvaardigheid zijn de referentieniveaus en de domeinbeschrijving als uitgangspunt genomen. Voor het ontwikkelen van de vragenlijsten voor leerlingen, leerkrachten en schoolleiders (zie paragraaf 1.4) is gebruik gemaakt van een conceptueel model dat was gebaseerd op de reviewstudie van Houtveen et al. (2019), recente wetenschappelijke inzichten en het meest recente peilingsonderzoek naar leesvaardigheid (Kuhlemeier e.a., 2014). Het conceptueel model beschreef algemene en domeinspecifieke kenmerken op leerling-, leerkracht- en schoolniveau die een rol spelen bij begrijpend lezen.

1.4 Het instrumentarium

1.4.1 De leestoets

Om de leesvaardigheid te kunnen meten is een leestoets ontwikkeld met in totaal 165 opgaven uit vier leesdomeinen: begrijpen, interpreteren, samenvatten³ en evalueren. De leesdomeinen techniek en woordenschat en opzoeken zijn niet meegenomen in dit peilingsonderzoek (conform domeinbeschrijving, zie SLO 2019). Wel zijn er in het instrumentarium binnen het domein begrijpen enkele gesloten en open vragen naar woordbetekenissen opgenomen. Aangezien deze peiling zich richtte op begrijpend lezen, is het technisch leesniveau niet getoetst. Ter indicatie van het technisch leesniveau van leerlingen is het AVI-niveau van de leerlingen opgevraagd.

De leestoets bevatte zowel open als gesloten opgaven, die verschillende aspecten van leesvaardigheid toetsten, bij verschillende tekstsoorten. De 165 opgaven uit de leestoets zijn ondergebracht in 13 toetsversies, die van elkaar verschillen in moeilijkheidsgraad, maar wel deels overlappende opgaven bevatten. Leerlingen kregen een toetsversie toebedeeld op basis van het vooraf door de school aangeleverde uitstroomprofiel voortgezet onderwijs van de betreffende leerling. Omdat leerlingen deels verschillende opgaven maakten, zijn hun resultaten niet rechtstreeks te vergelijken. Om deze vergelijkingen toch mogelijk te maken, zijn de opgaven en leerlingen op één onderliggende meetschaal geplaatst (met behulp van itemresponsetheorie-modellen (IRT)⁴). Daardoor konden voor alle leerlingen zogeheten vaardigheidsscores worden berekend die wél onderling vergelijkbaar zijn.

Koppeling met de referentieniveaus

De vaardigheidsscores konden worden gekoppeld aan de referentieniveaus, omdat in de leestoets zogeheten ankeropgaven zijn opgenomen. Deze opgaven komen uit twee referentiesets die zijn voorzien van prestatiestandaarden voor 1F en 1S. Door deze ankeropgaven af te nemen, konden de prestatiestandaarden overgebracht worden naar het instrumentarium van deze peiling. Daarbij dient opgemerkt te worden dat voor deze peiling de referentieniveaus bepaald zijn op de gehele schaal leesvaardigheid, waarin alle vier de domeinen zijn opgenomen (begrijpen, interpreteren, samenvatten en evalueren). In de referentiesets, die ook gebruikt worden voor de ijking van de verschillende eindtoetsen (CvtE, 2014), zijn oorspronkelijk echter geen opgaven van het domein evalueren opgenomen en maar weinig opgaven van het domein samenvatten. Daarom hebben we ook gekeken naar de beheersing van de referentieniveaus als de opgaven evalueren buiten beschouwing worden gelaten (zie par. 2.1)

Beheersing op basis van deze nieuwe schaal, zonder de opgaven van het domein evalueren, zou de basis kunnen vormen voor een vergelijking met de resultaten van de eindtoets. Door een dergelijke vergelijking te maken, ontstaat er een beeld van het eventuele verschil in resultaten in een *high stakes* (de eindtoets) en *low stakes* (het peilingsonderzoek) conditie. Waar er immers aan de toets in dit peilingsonderzoek voor leerlingen geen belangrijke consequenties verbonden zijn, geldt dit wel voor de eindtoetsen gezien hun adviesfunctie voor het vervolgonderwijs. Uit onderzoek is bekend dat prestaties tussen een *low* en *high stakes* context kunnen verschillen, omdat de leerling minder belang hecht aan de *low stakes* toets (zie bijvoorbeeld Cole & Osterlind, 2008). Door problemen rondom

³ Samenvatten staat in het referentiekader Taal alleen bij zakelijke teksten opgenomen en geldt als een aspect dat alleen vanaf niveau 2F gepeild kan worden.

⁴ Bij de schaling bleek 1 item niet goed te passen. Het gaat om een meerkeuze opgave uit het domein interpreteren bij een literaire tekst. De leerlingresultaten zijn dus gebaseerd op de 164 overgebleven items. Zie voor meer informatie over de schalingsprocedure het technisch rapport van het uitvoerend consortium (nog te verschijnen).

de normering van de referentieniveaus, is het helaas niet mogelijk om de gegevens van de eindtoetsen 2021 te benutten voor een dergelijke vergelijking (zie Kamerbrief minister Wiersma, 2022). Hoewel we de precieze invloed van de *low stakes* context dus niet kennen, is het bij het duiden van de resultaten uit peilingsonderzoek wel belangrijk om dit conditie-effect mee te wegen.

Berekening van de trend

Naast opgaven uit de referentiesets, bevatte de leestoets ook ankeropgaven uit vorige peilingsonderzoek in het bo in 2011 en het sbo in 2012 (Kuhlemeier et al., 2014). De opname van deze opgaven maakte de trendvergelijking over tijd mogelijk. We kunnen de trendanalyse echter alleen uitvoeren voor de domeinen interpreteren en begrijpen, omdat niet alle leesdomeinen in de leestoets van zowel het huidige als de vorige peilingsonderzoeken terugkomen⁵. Om de prestaties direct te kunnen vergelijken, zijn deze eerst op één meetschaal geplaatst met behulp van de ankeropgaven. Het is daarbij belangrijk dat de ankeropgaven vergelijkbaar functioneren in de verschillende peilingen (m.a.w. geen differential item functioning (DIF) laten zien). Omdat een aantal opgaven anders functioneerden in de voorgaande peilingen, zijn deze buiten beschouwing gelaten bij het construeren van de meetschaal. Uiteindelijk zijn er 23 opgaven die de link tussen de bo-peilingen verzorgen en 11 opgaven die de link tussen de sbo-peilingen verzorgen. Van deze opgaven is er 1 opgave die een link vormt tussen alle peilingen.

1.4.2

De vragenlijsten

Voor het peilingsonderzoek zijn vragenlijsten ontwikkeld om algemene en domeinspecifieke kenmerken van leerlingen, leerkrachten, de school en het onderwijsleerproces in kaart te brengen. In de factsheet worden de uitkomsten van de leerlingvragenlijst in verband gebracht met de leesprestaties. De leerlingvragenlijst bestond uit zeventien vragen. De meerderheid van de vragen was gebaseerd op de leerlingvragenlijst van PIRLS-2021, zodat een vergelijking tussen algemene en domeinspecifieke leerlingkenmerken van groep 8 (peilingsonderzoek Leesvaardigheid) en groep 6 (PIRLS-2021) mogelijk is (vergelijking verwacht in 2023).

De volgende variabelen werden bevraagd: leeftijd, thuistaal en herkomst, taalrijkheid van de thuisomgeving, lezen voor eigen plezier, leesmotivatie- intrinsiek en extrinsiek, attitude ten opzichte van lezen, leesvertrouwen, gebruik van digitale middelen, digitale vaardigheden, tijd besteed aan digitale middelen buiten school, waardering van de lessen begrijpend lezen op school en leesactiviteiten op school. Daarnaast werden het meest recente AVI-niveau, het voorlopig schooladvies (bo) en het verwachte uitstroomprofiel (sbo) opgevraagd. 3318 bo leerlingen en 1000 sbo leerlingen vulden de vragenlijst in. In bijlage 1 is te vinden welke vragenlijstitems de basis bieden voor de schalen taalrijkheid van de thuisomgeving, lezen voor eigen plezier, leesmotivatie, leesattitude, leesvertrouwen en de waardering van de lessen begrijpend lezen. Voor alle schalen geldt dat deze als voldoende betrouwbaar zijn beoordeeld ($\text{Lambda-2} > 0,7$ conform COTAN-richtlijn: Evers et al., 2010).

⁵ Het domein evalueren is, net als begrijpen en interpreteren, ook in de huidige en voorgaande peilingen meegenomen. Echter, door structurele ontbrekende waarden in de oude peilingen bij opgaven van dit domein, kan er voor evalueren geen trendanalyse uitgevoerd worden.

2 De leerlingresultaten uit de factsheet

Zoals gezegd focust de factsheet op de leerlingresultaten op het gebied van begrijpend lezen, de leesprestaties over tijd en samenhangen tussen leerlingkenmerken en – prestaties. In deze paragraaf worden de bevindingen die in de factsheet worden gepresenteerd, onderbouwd.

2.1 Beheersing referentieniveaus leesvaardigheid

Na het overbrengen van de prestatiestandaarden (zie paragraaf 1.4.1) is in kaart gebracht op welk referentieniveau bo- en sbo-leerlingen functioneren op het gebied van begrijpend lezen. In Tabel 2.1a is per schooltype voor de verschillende referentieniveaus het aantal en percentage leerlingen weergegeven dat een bepaald referentieniveau heeft behaald. Omdat binnen het taaldomein het streefniveau 1S gelijk is aan het 2F-niveau, spreken we hier en in de factsheet van "2F" (in plaats van 1S).

Tabel 2.1a Behaalde referentieniveaus bij leesvaardigheid per schooltype

	bo		sbo		Totaal	
Referentieniveau	N	%	N	%	N	%
Lager dan 1F	402	12,1	606	60,6	1008	23,3
1F	1254	37,8	325	32,5	1579	36,6
2F	1662	50,1	69	6,9	1731	40,1
<i>Totaal</i>	<i>3318</i>	<i>100,0</i>	<i>1000</i>	<i>100,0</i>	<i>4318</i>	<i>100,0</i>

Zoals aangekondigd in paragraaf 1.4 kijken we ook naar de beheersing van de referentieniveaus als de opgaven evalueren buiten beschouwing worden gelaten. Dat levert een lichte verschuiving in de beheersing op. Tabel 2.1a geeft het resultaat weer.

Tabel 2.1a Verschuiving in beheersing referentieniveaus na uitsluiten opgaven uit het domein evalueren.

	< 1F	1F	2F
Originele schaal met Evalueren-items	1008	1579	1731
Nieuwe schaal zonder Evalueren items	1264	1285	1769

Te zien is dat als de evalueren-opgaven buiten beschouwing gelaten worden bij het bepalen van de referentieniveaus, meer leerlingen het niveau 1F niet behalen. Daarnaast is er een klein deel van de leerlingen dat juist hoger scoort; er zijn iets meer leerlingen met niveau 2F. In de factsheet gaan we voor de beheersing van de referentieniveaus uit van de schaal uit van de totaalschaal inclusief opgaven evalueren, omdat dit domein ook uitdrukkelijk onderdeel vormt van het referentiekader.

2.2 Relatie met achtergrondkenmerken

Voor het onderzoeken van samenhangen tussen de leesprestaties enerzijds en kenmerken van scholen, leerkrachten, onderwijsleerproces en leerlingen anderzijds, is een reeks meerniveau-analyses uitgevoerd. In deze analyses is uitgegaan van twee niveaus (leerlingen genest binnen klassen). Ontbrekende waarden op

predictorniveau zijn – na exclusie van leerlingen met teveel missings - voorafgaand aan deze multilevel analyses geïmputeerd met behulp van het pakket Mice (van Buuren & Groothuis-Oudshoorn, 2011). Daarbij is rekening gehouden met de multilevel structuur zoals aanwezig in de dataset. Voorafgaand aan de analyses zijn zowel de uitkomstmaten (WML-schatters voor leesvaardigheid algemeen en de domeinen begrijpen, interpreteren, evalueren en samenvatten afzonderlijk) als de predictorvariabelen getransformeerd tot z-scores om zo gestandaardiseerde regressiecoëfficiënten te verkrijgen. Gestandaardiseerde regressiecoëfficiënten hebben het voordeel dat ze vergeleken kunnen worden en dat de grootte van het effect bepaald kan worden (Lorah, 2018). Deze gestandaardiseerde coëfficiënten (ook wel beta weights genoemd) kunnen op vergelijkbare wijze als (partiële) correlaties worden geïnterpreteerd met de daarbij geldende richtlijnen: $< .02$ = klein effect, $0.2 - 0.5$ = matig effect, > 0.5 = sterk effect⁶ (Acock, 2014).

De meerniveau-analyses zijn in verschillende stappen uitgevoerd. Eerst is er een 0-model geschat, met als enige predictor het schooltype: bo of sbo. Vervolgens zijn in verschillende deelmodellen predictoren aan het 0-model toegevoegd om inzicht te krijgen in samenhangen tussen de leesprestaties en respectievelijk: algemene leerlingkenmerken, domeinspecifieke leerlingkenmerken, algemene leerkrachtkenmerken, domeinspecifieke leerkrachtkenmerken, algemene schoolkenmerken, algemene kenmerken van het onderwijsleerproces, kenmerken rondom de vormgeving van het onderwijsleerproces en kenmerken rondom de inhoud van het onderwijsleerproces. Voor alle domeinspecifieke kenmerken geldt dat er een interactieterm met schooltype in de deelmodellen is opgenomen. Significante interacties en predictoren uit deze deelmodellen zijn vervolgens gecombineerd tot één totaalmodel volgens onderstaand werkwijze:

- Start: Leeg model met enkel type onderwijs als predictor;
- Blok 1: Alle significante kenmerken van het onderwijsleerproces (algemeen, vormgeving en inhoud) worden toegevoegd, inclusief de interacties met schooltype. Kenmerken worden meegenomen als ze in minimaal met een van de uitkomstmaten (leesvaardigheid algemeen, begrijpen, interpreteren, evalueren en samenvatten) een significante samenhang vertoonden.
- Blok 2: Alle significante domeinspecifieke leerkracht- en leerlingkenmerken worden toegevoegd, inclusief interacties met type onderwijs. Kenmerken worden meegenomen als ze in minimaal met een van de uitkomstmaten een significante samenhang vertoonden.
- Blok 3: Alle significante algemene leerling-, leerkracht- en schoolkenmerken wordt toegevoegd. Kenmerken worden meegenomen als ze in minimaal met een van de uitkomstmaten een significante samenhang vertoonden.
- Compact totaalmodel: Verwijderen van alle niet-significante interactie- en hoofdeffecten. Er wordt een significantieniveau van $p < .05$ aangehouden.
- Significante interactie-effecten worden geduid door apart voor het bo en sbo een model te draaien.

In de factsheet worden robuuste samenhangen met leerlingkenmerken uitgelicht, omdat we op dit niveau de sterkste samenhangen vinden. Onder robuuste samenhangen verstaan we hier kenmerken die zowel in het bo als het sbo een significante samenhang vertonen met de totaalschaal voor leesvaardigheid in het compacte totaalmodel. Daarnaast geldt dat deze samenhangen een eenduidig beeld laten zien over de leesdomeinen.

⁶ In sommige rapporten wordt er ook gebruik gemaakt van benchmarks geldend bij Cohen's d, waarbij $\beta = 0,2$ een klein effect $\beta = 0,5$ een matig effect en $\beta = 0,8$ een groot effect (Cohen, 1988).

Tabel 2.2a geeft de robuuste samenhangen tussen leerlingkenmerken en de totale vaardigheidsschaal met daarbij de gestandaardiseerde regressie coëfficiënt en een duiding van de grootte van het effect. Deze tabel geeft dus een selectie van de uitkomsten van het compacte totaalmodel dat de samenhangen met leesvaardigheid in beeld brengt. De volledige modelresultaten (waarin ook kenmerken op het niveau van het onderwijsleerproces, de leerkracht en de school zijn opgenomen) volgen in het publieksrapport bij dit peilingsonderzoek in december 2022.

Tabel 2.2a Robuuste samenhangen tussen leerlingkenmerken en leesprestaties

	Totaal	
	β	p
Fixed effecten		
(Intercept)	-0.006	0.882
Type onderwijs: sbo	-0.928	0.000
Lezen voor eigen plezier ^a	-0.098	0.000
Leesmotivatie: extrinsiek ^a	0.047	0.002
Leesvertrouwen ^a	-0.182	0.000
Geslacht leerling: meisje	0.218	0.000
Thuis taal: soms of nooit NL	-0.131	0.001
Aantal boeken thuis ⁷		
Aantal Boeken hoog (meer dan 100)	0.255	0.000
Aantal Boeken midden (26 – 100)	0.215	0.000
Random effecten		
Klas (niveau 2)	0.059	
Leerling (niveau 2)	0.595	

^a Hoe hoger de factorscore op deze schaal, des te lager is de score op het construct (en vice versa).

<.02 = klein effect, 0.2-0.5 = matig effect, > 0.5 = sterk effect

2.3

Trend over tijd

Zoals in paragraaf 1.4.1 aangegeven, kon er een trendvergelijking worden gemaakt door de prestaties van leerlingen uit het huidige en uit de voorgaande peilingsonderzoeken op één meetschaal te plaatsen. Op die manier zijn de prestaties van bo- en sbo-leerlingen uit het huidige peilingsonderzoek rechtstreeks te vergelijken met de prestaties uit 2011 en 2012, toen de leesvaardigheid voor het laatst is gepeild (Kuhlemeier et al., 2014). Tabel 2.3 geeft de effectgroottes van deze vergelijking over tijd in het bo en in het sbo. In lijn met de vuistregels voor interpretatie van Cohen's d , worden beide effecten als "klein" beschouwd (want tussen .20 en .49; Cohen, 1988).

Tabel 2.3 effectgrootte trendvergelijking

Populatie	effect	significantie
bo	-0,30	$p < .001$
sbo	-0,48	$p < .001$

Helaas is het niet mogelijk om een concreet getal toe te kennen aan de daling van de leesvaardigheid, omdat de vaardigheidsscores op de meetschaal geen

⁷ Voor de domeinen begrijpen en samenvatten geldt dat de samenhang met het aantal boeken thuis alleen significant is in het bo. In het model met de totale vaardigheidsschaal als uitkomstmaat, bestaat dit verschil alleen voor het contrast laag vs. midden. In tabel 2.2a wordt het hoofdeffect getoond.

eigenstandige betekenisvolle interpretatie hebben. Aangezien het in het vorige peilingsonderzoek nog niet mogelijk was om in termen van behaalde referentieniveaus te rapporteren, kunnen we ook geen trendvergelijking maken in termen van het percentage beheersing van de referentieniveaus.

2.4 Voorbeeldopgaven

In de factsheet staan twee voorbeeldopgaven opgenomen om de vaardigheid van leerlingen op niveau 1F en 2F te illustreren. Beide opgaven komen uit het domein evalueren en horen bij een zakelijke, instructieve tekst. Omwille van de beschikbare ruimte op de factsheet is er voor gekozen om een korte tekst te tonen. In werkelijkheid bestond de leestoets uit een brede variëteit aan tekstsoorten en opgaven. Tabel 2.4a geeft een overzicht van de tekstsoorten in de leestoets en 2.4b van de verdeling van open en gesloten opgaven over domeinen. Waar voorkennis soms een probleem kan zijn bij het meten van leesvaardigheid, zou dit door de grote diversiteit aan teksten geen rol moeten spelen.

Tabel 2.4a Verdeling tekstsoorten over domeinen in leestoets

	Begrijpen	Evalueren	Interpreteren	Samenvatten
Fictie	28	10	23	
fabel	1	1	2	
literair	3	3	5	
mop	2		2	
narratief	10	5	13	
parabelachtig	3		1	
poëzie	9	1		
zakelijk	20	16	45	23
Advertentie		1	4	
betogend			7	
informatief	12	13	20	18
instructief	8	2	14	5
	48	26	68	23

Tabel 2.4b Verdeling open en gesloten opgaven over domeinen

	Begrijpen	Evalueren	Interpreteren	Samenvatten
Meerkeuze	38	1	61	20
Open opgave	10	25	7	3
Totaal	48	26	68	23

3 Referenties

- Acock, A. C. (2014). *A Gentle Introduction to Stata (4th ed.)*. Texas: Stata Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences (2nd ed.)*. Hillsdale, NJ: L. Erlbaum Associates.
- Cole, J. S., & Osterlind, S. J. (2008). Investigating differences between low-and high-stakes test performance on a general education exam. *The Journal of General Education, 57*(2), 119-130.
- CvTE (2014) *Rapportage | referentiesets taal (lezen) en rekenen Verantwoording project*, geraadpleegd van:
http://www.toetsspecials.nl/html/referentiesets/rapportage_referentiesets.shtm
- Evers, A., Lucassen, W., Meijer, R., & Sijtsma, K. (2010). *COTAN beoordelingssysteem voor de kwaliteit van tests* (COTAN ratingsystem for the quality of tests).
- Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2008b). *Over de drempels met taal en rekenen. Hoofdrapport van de Expertgroep Doorlopende Leerlijnen Taal en Rekenen*. Enschede: SLO.
- Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2009). *Het referentiekader taal en rekenen. De referentieniveaus*. Enschede: SLO in opdracht van het ministerie van Onderwijs, Cultuur en Wetenschap (OCW).
- Greven, J., & Letschert, J. (2006). *Kerndoelen primair onderwijs*. Ministerie van Onderwijs, Cultuur en Wetenschap, geraadpleegd van:
<https://www.rijksoverheid.nl/documenten/rapporten/2006/04/28/kerndoelenboekje>
- Haelermans, C., Abbink, H., Baumann, S., Bijlsma, I., Havermans, W., Jacobs, M., Van Vugt, L., & Van Wetten, S. (2022). *Leergroei begrijpend lezen in tweede COVID-jaar hersteld, spelling en rekenen-wiskunde nog niet*. NCO Factsheet No. 2022-2.
- Houtveen, A. A. M., Steensel, R. V., & de la Rie, S. (2019). *De vele kanten van leesbegrip: Literatuurstudie naar onderwijs in begrijpend lezen in opdracht van het Nationaal Regieorgaan Onderwijsonderzoek en de Inspectie van het Onderwijs* (p. 120). Erasmus University Rotterdam.
- Kuhlemeier, H., Jolink, A., Krämer, I., Hemker, B., Jongen, I., Berkel, S. van, & Bechger, T. (2014). *Balans van de leesvaardigheid in het basis- en speciaal basisonderwijs 2. Uitkomsten van de peilingen in 2011 en 2012 in groep 8, groep 5 en de eindgroep van het SBO*. PPONreeks nummer 54. Arnhem: Cito.
- Lorah, J. (2018). Effect size measures for multilevel models: Definition, interpretation, and TIMSS example. *Large-Scale Assessments in Education, 6*(1), 1-11.
- Pulles, M., Prenger, J. (2019). *Leesvaardigheid in het basisonderwijs: Domeinbeschrijving ten behoeve van peilingsonderzoek*. Enschede: SLO.

Staatsblad van het Koninkrijk der Nederlanden (2010). *Besluit referentieniveaus Nederlandse taal en rekenen*. <https://zoek.officielebekendmakingen.nl/stb-2010-265.html>

Swart, N., Gubbels, J., in 't Zandt, M, Kuijpers, R., Hemker, B., Hommelberg, R., van Langen, A., Wolbers, M., Segers, E., Droop, M. (nog te verschijnen)
Leesvaardigheden in (s)bo. Technisch rapport Peil.onderwijs Leesvaardigheid einde (speciaal) basisonderwijs 2021

Van Buuren, S., & Groothuis-Oudshoorn, C. G. M. (2011). mice: Multivariate Imputation by Chained Equations in R. *Journal of statistical software*, 45(3).

Wiersma, A.D. (2022, 28 januari). Normering eindtoetsen en referentieniveaus primair onderwijs [Kamerbrief]. Geraadpleegd van:
<https://www.rijksoverheid.nl/documenten/kamerstukken/2022/01/28/aan-de-tweede-kamer-normering-eindtoetsen-en-referentieniveaus-primair-onderwijs>

4 Bijlage 1.

Overzicht geconstrueerde schalen leerlingvragenlijst en bijbehorende items.

Variabele	Schaal/ schalen per variabele	Item
Taalrijkheid thuisomgeving	Ouders: ondersteuning bij lezen	Hoe vaak: Word je door je ouder(s) aangemoedigd om te gaan lezen? Word je door je ouder(s) geholpen een boek te kiezen om te lezen? Lees je zelf een boek terwijl (een van) je ouders ook aan het lezen zijn? Praat je met (een van) je ouders over wat je in je vrije tijd hebt gelezen?
Lezen voor eigen plezier	Leesgedrag	Hoeveel tijd besteed je op een schooldag aan lezen buiten school? Hoe vaak leen je boeken (of e-books) van school of bij de bibliotheek? Hoe vaak doe je deze dingen buiten school? Ik lees voor mijn plezier Ik lees om dingen te weten te komen
Leesmotivatie	Intrinsiek	Ik lees omdat ik zo meer kan leren over dingen die ik interessant vind Ik lees omdat ik me graag met bepaalde onderwerpen bezig houd Ik lees omdat ik zo avonturen kan beleven in mijn hoofd Ik lees omdat ik me graag verplaats in de hoofdpersoon van een verhaal
	Extrinsiek	Ik lees omdat ik anderen niet wil teleurstellen Ik lees omdat anderen dit van mij verwachten Ik lees omdat anderen vinden dat ik dat moet doen Ik lees omdat ik dan betere cijfers kan halen op school
Leesattitude	Leesattitude	Ik vind het leuk om met andere mensen te praten over wat ik lees Ik vind het leuk als ik een boek als cadeau krijg Ik vind lezen saai* Ik wou dat ik meer tijd had om te lezen Ik vind lezen leuk Ik leer veel door te lezen Ik vind het leuk om dingen te lezen die me aan het denken zetten Ik vind het leuk om boeken te lezen die me helpen fantaseren over andere werelden
Leesvertrouwen	Leesvertrouwen	Ik haal vaak goede cijfers bij lezen Ik vind lezen makkelijk

Variabele	Schaal/ schalen per variabele	Item
Waardering van de lessen begrijpend lezen	Inhoudelijke interesse	Ik vind het moeilijk verhalen te lezen met moeilijke woorden erin*
		Lezen is voor mij moeilijker dan voor veel andere kinderen in mijn klas*
		Ik vind lezen moeilijker dan alle andere vakken*
		Ik ben gewoon niet goed in lezen*
		Ik vind de dingen die ik op school lees leuk
	Leerkrachthandelen	De dingen die ik van mijn leerkracht moet lezen zijn interessant
		Ik ben geïnteresseerd in wat mijn leerkracht zegt bij de lessen over (begrijpend) lezen
		Ik weet wat mijn leerkracht van me verwacht bij het (begrijpend) lezen
		Mijn leerkracht is makkelijk te begrijpen bij de lessen over (begrijpend) lezen
		Mijn leerkracht moedigt me aan om te zeggen wat ik denk over wat ik gelezen heb
	Klassenklimaat	Mijn leerkracht geeft me de kans te laten zien wat ik geleerd heb bij het (begrijpend) lezen
		Mijn leerkracht doet verschillende dingen om ons te helpen leren bij het (begrijpend) lezen
		Mijn leerkracht vertelt me hoe ik het beter kan doen als ik een fout maak bij het (begrijpend) lezen
		Leerlingen luisteren niet naar wat de leerkracht zegt
		Er is teveel lawaai voor leerlingen om goed te kunnen werken
Mijn leerkracht moet lang wachten voordat de leerlingen stil zijn		
De leerlingen onderbreken de leerkracht		
Mijn leerkracht moet ons steeds vertellen dat we ons aan de regels in de klas moeten houden		

* items is voorafgaand aan schaling omgepoold zodat alle vragen in een schaal dezelfde richting op wijzen.

Colofon

Inspectie van het Onderwijs
Postbus 2730 | 3500 GS Utrecht
www.onderwijsinspectie.nl

Een exemplaar van deze publicatie is te downloaden vanaf de website van de
Inspectie van het Onderwijs: www.onderwijsinspectie.nl.

© Inspectie van het Onderwijs | juni 2022