

Rekenen-wiskunde aan het eind van het basisonderwijs en het speciaal (basis)onderwijs in schooljaar 2022-2023

Technisch rapport Peil.Rekenen-wiskunde PO 2023

Martina Meelissen, Jolien Valk, Nathalie Maassen (Universiteit Twente)

Remco Feskens, Jesse Koops, Zoë ten Napel (Stichting Cito)

Michel Hop (Cito BV)

Annemarie van Langen, Tessa Jenniskens (KBA Nijmegen)

Colofon

Rekenen-wiskunde aan het eind van het basisonderwijs en het speciaal (basis)onderwijs in schooljaar 2022-2023. Technisch rapport Peil.Rekenen-wiskunde PO 2023

Martina Meelissen, Jolien Valk, Nathalie Maassen (Universiteit Twente)

Remco Feskens, Jesse Koops, Zoë ten Napel (Stichting Cito)

Michel Hop (Cito BV)

Annemarie van Langen, Tessa Jenniskens (KBA Nijmegen)

Volledige referentie:

Meelissen, M., Valk, J., Maassen, N., Feskens, R., Koops, J., Napel, Z. ten, Hop, M., Langen, A. van, Jenniskens, T. (2024). *Rekenen-wiskunde aan het eind van het basisonderwijs en het speciaal (basis)onderwijs in schooljaar 2022-2023. Technisch rapport Peil.Rekenen-wiskunde PO 2023*. Enschede: Universiteit Twente.

Dit onderzoek is uitgezet door het Nationaal Regieorgaan Onderwijsonderzoek (NRO).

© Universiteit Twente 2024

Inhoudsopgave

1. Achtergrond van het onderzoek	6
1.1 Inleiding	6
1.2 Doelen	6
1.3 Context	7
1.4 Onderzoeksvragen	9
1.5 Leeswijzer	10
2. Ontwikkeling toetsinstrumentarium	11
2.1 Theoretische uitgangspunten en inhoud	11
2.1.1 Beoogde itemconstructie	12
2.1.2 Gerealiseerd instrumentarium	13
2.2 Afnamedesign proeftoets	13
2.3 Resultaten proeftoets	14
2.4 Afnamedesign hoofdonderzoek	15
3. Ontwikkeling vragenlijsten	17
3.1 Conceptueel model	17
3.2 Uitgangspunten constructie vragenlijsten	17
3.3 Vragenlijsten proefafname	18
3.4 Vragenlijsten hoofdonderzoek	19
4. Steekproefdesign	24
4.1 Steekproef regulier basisonderwijs	24
4.2 Steekproef sbo en so4	25
4.3 Uitsluitingscriteria	25
5. Werving, dataverzameling en respons hoofdonderzoek	27
5.1 Organisatie scholenwerving	27
5.2 Dataverzameling	28
5.2.1 Voorbereidende dataverzameling	28
5.2.2 Training toetsleiders	29
5.2.3 Afname toets en leerlingvragenlijst	29
5.2.4 Afname leerkracht- en schoolvragenlijst	30
5.3 Respons op schoolniveau	30
5.3.1 Regulier basisonderwijs	30
5.3.2 Speciaal basisonderwijs	31
5.3.3 Speciaal onderwijs	31
5.4 Redenen voor niet-deelname	32
5.5 Respons op schoolleider-, leerkracht- en leerlingniveau	34
5.6 Representativiteitsanalyse	36
6. Schaling Rekenen-wiskunde	37
6.1 Inleiding	37
6.2 De toets en afname	37
6.2.1 Klassieke analyse	39

6.3 Schaling	40
6.3.1 DIF tussen 2019 en 2022	41
6.3.2 Modusonderzoek	43
6.3.3 DIF tussen de proeftoets 2022 en de peiling in 2023	45
6.3.4 DIF naar schooltype	48
6.3.5 Model en itempassing	50
6.3.6 Samenhang tussen onderdelen	54
6.4 Betrouwbaarheid	55
6.5 Referentieniveaus	56
7. Uitkomsten rekenvaardigheid	57
7.1 Vergelijking met 2019	57
7.2 Referentieniveaus	61
7.3 Prestaties op domeinen en onderdelen	62
7.3.1 Rekendomeinen	62
7.3.2 Rekenmachine	64
7.3.3 Kaal/context	67
7.4 Beschrijving van de schaal	67
8. Uitkomsten contextvragenlijsten	69
8.1 Inleiding	69
8.2 Leerlingvragenlijst	69
8.2.1 Deelname en ontbrekende waarden	69
8.2.2 Achtergrondkenmerken leerlingen	71
8.2.3 Domeinspecifieke leerlingkenmerken	72
8.2.4 Domeinspecifieke onderwijskenmerken	74
8.2.5 Overzicht constructen leerlingvragenlijst	76
8.3 Leerkrachtvragenlijst	77
8.3.1 Deelname en ontbrekende waarden	77
8.3.2 Achtergrondkenmerken leerkrachten	78
8.3.3 Schoolkenmerken (perspectief leerkracht)	80
8.3.4 Domeinspecifieke leerkrachtkenmerken	83
8.3.5 Domeinspecifieke onderwijskenmerken	85
8.3.6 Overzicht constructen leerkrachtvragenlijst	94
8.4 Schoolvragenlijst	96
8.4.1 Deelname en ontbrekende waarden	96
8.4.2 Achtergrondkenmerken schoolleiders	97
8.4.3 Schoolkenmerken	98
8.4.4 Domeinspecifieke schoolkenmerken	100
8.4.5 Overzicht constructen schoolvragenlijst	102
9. Samenhang kenmerken leerlingen, leerkrachten, scholen en onderwijsleerproces met leerprestaties	103
9.1 Inleiding	103
9.2 Interpretatie van effecten in multilevel modellen	103
9.3 Spreiding toetsresultaten	104

9.3.1 Hoeveel variabiliteit in leerlingprestaties is toe te schrijven aan scholen/klassen en leerlingen?	104
9.4 Leerling	106
9.4.1 Welke verschillen in leerlingprestaties zijn er naar leerlingkenmerken?	108
9.5 Leerkracht	109
9.5.1 Welke verschillen in leerlingprestaties zijn er naar algemene leerkrachtkenmerken?	110
9.5.2 Welke verschillen in leerlingprestaties zijn er naar domeinspecifieke leerkrachtkenmerken?	112
9.5.3 Welke verschillen in leerlingprestaties zijn er naar kenmerken van het reken-wiskundeonderwijs?	114
9.6 School	118
9.6.1 Welke verschillen in leerlingprestaties zijn er naar algemene schoolkenmerken?	118
9.7 Alle kenmerken in samenhang	119
9.8 Verklaarde variantie in de verschillende modellen	122
10. Samenvatting	123
Referenties	128
Bijlagen	130
Bijlage 1 Representativiteitsanalyse	131
B1.1 Methode	131
B1.2 Regulier basisonderwijs	132
B1.3 Speciaal basisonderwijs	137
B1.4 Speciaal onderwijs	141
B1.5 Representativiteit op leerlingniveau	144
B1.5 Wel of niet wegen?	148
B1.6 Conclusie	149
Bijlage 2 Prestaties op onderdelen	150
B2.1 Inleiding	150
B2.2 Gestandaardiseerde effecten	152
Bijlage 3 Overzicht variabelen multilevel analyses	157

1. Achtergrond van het onderzoek

1.1 Inleiding

Onder de naam 'Peil.Rekenen-wiskunde PO 2023' hebben in het voorjaar van 2023 bijna 5.000 leerlingen in groep 8 van het regulier basisonderwijs en in de schoolverlatersgroep van het speciaal (basis)onderwijs een digitale toets voor rekenen-wiskunde gemaakt. Daarnaast is met vragenlijsten voor leerlingen, leerkrachten en schoolleiders informatie verzameld over leerling-, leerkracht- en schoolkenmerken, attitudes ten aanzien van rekenen-wiskunde, het onderwijsaanbod voor rekenen-wiskunde en hoe dit aanbod is vormgegeven. De peiling is uitgevoerd door een onderzoeksconsortium bestaande uit de Universiteit Twente, Stichting Cito, Cito BV en KBA Nijmegen, in opdracht van het Nationaal Regieorgaan Onderwijsonderzoek (NRO) en onder regie van de Inspectie van het Onderwijs.

Het is de tweede keer dat in het basisonderwijs een peilingsonderzoek Rekenen-wiskunde onder regie van de Inspectie van het Onderwijs is uitgevoerd. De dataverzameling van het eerste peilingsonderzoek vond plaats in het voorjaar van 2019 (Inspectie van het Onderwijs, 2021; Buisman et al., 2021).

In dit technische rapport worden de opzet, uitvoering en uitkomsten van Peil.Rekenen-wiskunde PO 2023 beschreven. Dit hoofdstuk beschrijft de doelen en de context waarbinnen dit onderzoek is uitgevoerd en de onderzoeksvragen die in dit rapport worden beantwoord.

1.2 Doelen

De Inspectie van het Onderwijs peilt periodiek de kennis en vaardigheden van leerlingen aan het einde van het primair onderwijs, (voortgezet) speciaal onderwijs en voortgezet onderwijs in de leergebieden die in de kerndoelen zijn opgenomen. Een van deze peilingen is Peil.Rekenen-wiskunde PO 2023.

Deze peiling meet het prestatieniveau van rekenen-wiskunde in schooljaar 2022-2023 van leerlingen aan het einde van het basisonderwijs en het speciaal (basis)onderwijs, dus kort voor ze doorstromen naar het voortgezet onderwijs of het voortgezet speciaal onderwijs. Daarnaast heeft deze peiling als doel inzicht te krijgen in de ontwikkeling van het prestatieniveau van rekenen-wiskunde in de afgelopen vier jaar, sinds de peiling in 2019. Ook biedt de peiling de mogelijkheid om de leerlingprestaties 2023 te vergelijken tussen de verschillende onderwijssoorten basisonderwijs (bao), speciaal basisonderwijs (sbo) en speciaal onderwijs. In het speciaal onderwijs gaat het dan voornamelijk om voormalig cluster 4 leerlingen met uitstroomprofiel vervolgonderwijs (voortaan aangeduid met so4).¹

1 Op sommige so4-scholen zitten ook leerlingen uit voormalig cluster 3. Vooraf is met de leerkracht van de schoolverlatersgroep (=laatste leerjaar vóór doorstroom naar v(s)o)) overlegd welke leerlingen wel of niet in staat waren om deel te nemen aan het onderzoek. Dit betekent dat ook een klein aantal voormalig cluster 3 leerlingen met uitstroomprofiel voortgezet onderwijs aan het onderzoek heeft deelgenomen.

Een ander doel van de peilingsonderzoeken, naast het meten en vergelijken van leerlingprestaties, is het verzamelen van informatie over houdingen van leerlingen en leerkrachten en over de wijze waarop het onderwijs wordt vormgegeven, aan de hand van leerling-, leerkracht- en schoolvragenlijsten. Tot slot bieden peilingsonderzoeken de mogelijkheid tot het uitvoeren van relationele analyses naar de kenmerken van leerlingen, leerkrachten, scholen en het onderwijsleerproces enerzijds en leerlingprestaties anderzijds.

In Peil.Rekenen-wiskunde PO 2023 is bij leerlingen een digitale (online) toets afgenomen gebaseerd op de kerndoelen, de referentieniveaus en de domeinbeschrijving voor rekenen-wiskunde (Van Zanten & Oldengarm, 2020). Aansluitend is bij deze leerlingen een digitale (online) vragenlijst afgenomen over onder meer hun attitudes en het onderwijsleerproces voor rekenen-wiskunde. De meeste informatie over het onderwijsleerproces is echter verkregen via de online leerkrachtvragenlijst, die daarnaast ook vragen bevatte over de attitude van de leerkracht ten aanzien van reken-wiskundeonderwijs. Ook de schoolleider vulde online een vragenlijst in, voornamelijk over (onderwijs)kenmerken van de school.

1.3 Context

Evenals in de eerste peiling Rekenen-wiskunde, in 2019, heeft de dataverzameling voor de tweede peiling gelijktijdig en op dezelfde reguliere basisscholen plaatsgevonden als het Nederlandse aandeel in het internationaal vergelijkend trendonderzoek TIMSS. TIMSS staat voor *Trends in International Mathematics and Science Study* en meet sinds 1995 om de vier jaar de leerlingprestaties in rekenen-wiskunde en de natuurwetenschappelijke vakken in groep 6 (*grade 4*) van het reguliere basisonderwijs.

De Nederlandse bijdrage aan TIMSS-2023 is door hetzelfde onderzoeksconsortium uitgevoerd als Peil.Rekenen-wiskunde PO 2023. De twee onderzoeken zijn gecombineerd om (reguliere) basisscholen minder vaak te hoeven benaderen voor onderzoek. De afnames in groep 6 en 8 vonden plaats op dezelfde dag en de schoolleiders vulden één schoolvragenlijst in (de TIMSS-versie) ten behoeve van beide onderzoeken. Het combineren van beide onderzoeken op dezelfde basisscholen vergroot ook de mogelijkheden om de uitkomsten van groep 6 en groep 8 met elkaar te vergelijken (geen onderdeel van dit rapport). Tabel 1.1 geeft een overzicht van de wijze waarop de twee onderzoeken voor het regulier basisonderwijs gecombineerd zijn.

Tabel 1.1 – Gecombineerde uitvoering TIMSS-2023 en Peil.Rekenen-wiskunde PO 2023 in regulier basisonderwijs

Onderdeel	Groep 6	Groep 8
Wat wordt er gemeten	Leerlingprestaties en -houdingen rekenen-wiskunde en natuuronderwijs, context waarbinnen deze tot stand komen	Leerlingprestaties en -houdingen rekenen-wiskunde, context waarbinnen deze tot stand komen
Steekproeftrekking scholen	Getrokken door het ISC* o.b.v aangeleverde data consortium	Steekproef TIMSS
Steekproeftrekking leerlingen	Alle groepen 6 van de deelnemende school	Alle groepen 8 van de deelnemende school
Instrumenten (toets en vragenlijsten) gebaseerd op:	Internationaal curriculum (TIMSS-Assessment Framework)** in samenspraak met deelnemende landen	Nederlandse curriculum (kerndoelen, referentieniveaus domeinbeschrijving *** en reviewstudies ****
Toetsontwikkeling	ISC* in samenspraak met deelnemende landen	Consortium in overleg met begeleidingscommissie
Ontwikkeling leerling- en leerkrachtvragenlijst	ISC* in samenspraak met deelnemende landen	Consortium in overleg met begeleidingscommissie
Schoolvragenlijst	ISC* in samenspraak met deelnemende landen	Schoolvragenlijst TIMSS
Afname: planning	Dezelfde dag als groep 8	Dezelfde dag als groep 6
Afnameprocedure	Internationaal voorgeschreven TIMSS-procedure, uitgevoerd door toetsleiders van het consortium	Zoveel mogelijk identiek aan TIMSS-procedure, uitgevoerd door dezelfde toetsleiders als TIMSS
Afnamemodus toets en vragenlijsten	Digitaal (online), platform toets: RM Assessment Master	Digitaal (online), platform toets en leerlingvragenlijst: TAO
Scoring toetsen	Deels computer, deels handmatig (open opgaven)	Computer
Schaling toets en vragenlijsten	ISC*	Consortium

* TIMSS & PIRLS International Study Center; ** Mullis, Martin, & Von Davier, 2021; *** Van Zanten & Oldengarm, 2020; **** Hickendorff et al., 2017; Hickendorff, 2020

De Peil.Rekenen-wiskunde PO 2023 is een vervolg op de peiling Rekenen-wiskunde van 2018-2019. Er zijn echter een aantal verschillen tussen de eerste en tweede peiling in zowel de context als de uitvoering:

- Speciaal onderwijs
De peiling van 2018-2019 richtte zich op groep 8 van het basisonderwijs en de schoolverlatersgroep in het speciaal basisonderwijs. In de huidige rekenpeiling zijn ook schoolverlaters in het speciaal onderwijs meegenomen. Het gaat hier voornamelijk om leerlingen in voormalig cluster 4 met uitstroomprofiel voortgezet onderwijs.

- COVID-19 pandemie

In de periode 2020-2022 hebben leerlingen en scholen te maken gehad met de gevolgen van de coronapandemie zoals scholensluitingen, thuisleren, lessenuitval en beperking van sociale contacten. De resultaten van onder andere NCO-analyses en PISA-2022 wijzen erop dat deze gevolgen in ieder geval op korte termijn negatief van invloed zijn geweest op leerprestaties, ook in rekenen-wiskunde (Haelermans, et.al., 2021; Meelissen et al., 2023). Bij het vergelijken van de leerlingprestaties die uit de peiling in 2019 en die in 2023 naar voren zijn gekomen, kunnen de gevolgen van de pandemie nog steeds een rol spelen.

- Afnamemodus

In de peiling van 2018-2019 werden de leerlingtoets – en vragenlijst op papier afgenomen. In deze tweede peiling zijn beide instrumenten in het hoofdonderzoek volledig digitaal (online) afgenomen. Ten behoeve van de trendanalyse in prestaties is in proefafname in het voorjaar van 2022 een equivalentiestudie uitgevoerd om moduseffecten te kunnen evalueren en indien nodig voor deze effecten te kunnen corrigeren. Op een groot deel van de reguliere basisscholen die deelnamen aan de proefafname in 2022, zijn daarvoor ankeropgaven zowel digitaal als op papier afgenomen. De uitvoering en uitkomsten van de equivalentiestudie zijn beschreven in Hoofdstuk 6.

1.4 Onderzoeksvragen

In totaal zijn er in de aanbesteding van dit peilingsonderzoek veertien onderzoeksvragen geformuleerd. De eerste zes vragen hebben specifiek betrekking op de leerlingprestaties op de peilingstoets, een vergelijking tussen de drie onderwijssoorten (bo, sbo en so4) en een vergelijking met 2019. Deze onderzoeksvragen worden beantwoord in hoofdstuk 7.

1. Hoeveel procent van de leerlingen einde basisonderwijs en schoolverlaters speciaal basisonderwijs en speciaal onderwijs beheerst de referentieniveaus (1F en 1S) voor rekenen-wiskunde en wat is de vaardigheidsverdeling op de onderliggende schaal?
2. Hoe kan de prestatie van leerlingen op de referentieniveaus worden getypeerd naar domeinen, subdomeinen, rekenonderwerpen en onderdelen (*described scale*)? Beschreven worden de prestatie van de referentieniveaugroepen <1F, 1F en 1S en per onderwijssoort, de prestatie van de groep laagvaardige (de 10% laagst scorende leerlingen) leerlingen, de gemiddelde leerling en de groep hoogvaardige (de 10% hoogst scorende leerlingen) leerlingen.
3. Hoe kan de prestatie van leerlingen voor rekenen met de rekenmachine worden getypeerd? Per onderwijssoort worden de prestaties van de laagvaardige (de 10% laagst scorende leerlingen), de gemiddelde en de hoogvaardige (de 10% hoogst scorende leerlingen) leerlingen beschreven?
4. Hoe kunnen per onderwijssoort de prestaties en het oplossingsproces van leerlingen voor probleemoplossen worden getypeerd?
5. Wat is de onderlinge samenhang in prestaties op domeinen, subdomeinen, rekenonderwerpen en onderdelen (inclusief rekenen met de rekenmachine en probleemoplossen)?
6. Welke verschillen zijn er tussen de leerlingprestaties einde basisonderwijs en einde speciaal basisonderwijs anno 2023 en de leerlingprestaties in Peil.Rekenen-wiskunde einde bo en sbo

2018-2019, ook gecorrigeerd voor eventuele verschillen in geslacht, leeftijd en thuistaal van de leerlingpopulaties?

In dit peilingsonderzoek is met vragenlijsten informatie verzameld over algemene kenmerken, domeinspecifieke kenmerken en domeinspecifieke onderwijskenmerken. De vragen 7 tot en met 14 hebben betrekking op deze kenmerken op het niveau van leerlingen, leerkrachten en scholen en hun mogelijke relatie met leerlingprestaties. Onderzoeksvraag 7 wordt beantwoord in paragraaf 8.4.

7. Hoe kunnen de deelnemende scholen worden getypeerd als het gaat om kenmerken van het onderwijs op het gebied van rekenen-wiskunde?

Onderzoeksvraag 8 tot en met 14 worden beantwoord aan de hand van multilevel analyses waarbij naar de samenhang wordt gekeken tussen domeinspecifieke (onderwijs-)kenmerken en toetsprestaties, gegeven niet-manipuleerbare algemene kenmerken van leerlingen, leerkrachten en scholen. In de oorspronkelijke aanbesteding was in de vraagstelling geen onderscheid gemaakt tussen domeinspecifieke kenmerken en domeinspecifieke onderwijskenmerken op leerlingniveau. Dit onderscheid is in deze rapportage wel gemaakt. Deze onderzoeksvragen worden beantwoord in hoofdstuk 9.

8. Hoeveel variabiliteit in de leerlingprestaties is toe te schrijven aan scholen/klassen en leerlingen?
9. Welke verschillen in leerlingprestaties zijn er tussen leerlingen naar algemene leerlingkenmerken?
10. Welke verschillen in leerlingprestaties zijn er tussen leerlingen naar domeinspecifieke leerlingkenmerken en domeinspecifieke onderwijskenmerken (bevraagd op niveau van de leerling)?
11. Welke verschillen in leerlingprestaties zijn er naar algemene leerkrachtkenmerken?
12. Welke verschillen in leerlingprestaties zijn er naar domeinspecifieke leerkrachtkenmerken?
13. Welke verschillen in leerlingprestaties zijn er naar algemene schoolkenmerken?
14. Welke verschillen in leerlingprestaties zijn er naar kenmerken van het onderwijs (domeinspecifieke onderwijskenmerken) op het gebied van rekenen-wiskunde (op niveau van de klas)?

1.5 Leeswijzer

De ontwikkeling van de toets en de vragenlijsten voor dit peilingsonderzoek komen aan bod in het tweede, respectievelijk derde hoofdstuk van deze technische rapportage. In hoofdstuk 4 wordt het steekproefdesign beschreven. Hoofdstuk 5 geeft informatie over de wijze waarop de werving en dataverzameling is uitgevoerd en geeft een overzicht van de uiteindelijke respons op school- en respondentniveau. De toelichting op de schaling van de reken-wiskundetoets die in deze peiling is afgenomen, is te vinden in hoofdstuk 6. De uitkomsten van de toets en vragenlijsten staan in hoofdstuk 7 en 8. Hoofdstuk 9 beschrijft de opzet, uitvoering en uitkomsten van de multilevel analyses. De antwoorden op de verschillende onderzoeksvragen worden gegeven in het laatste en samenvattende hoofdstuk (Hoofdstuk 10).

2. Ontwikkeling toetsinstrumentarium

2.1 Theoretische uitgangspunten en inhoud

Het doel van Peil.Rekenen-wiskunde PO 2023 is het verkrijgen van een landelijk beeld van het onderwijsaanbod en de leerresultaten voor rekenen-wiskunde van leerlingen einde baso, sbo en so4 die doorstromen naar het vo of vso. Voor het peilen van het onderwijsaanbod is onder meer de reviewstudie van Hickendorff e.a., (2017) van belang. Voor het peilen van de leerresultaten is gebruik gemaakt van de domeinbeschrijving van SLO ten behoeve van dit peilingsonderzoek. Het wettelijk beoogde curriculum, bestaande uit de kerndoelen en referentieniveaus, vormt het uitgangspunt voor het peilingsonderzoek.

Het Referentiekader beschrijft de doelen die leerlingen aan het einde van een bepaald onderwijsniveau moeten hebben bereikt. Hierbij richt de F-lijn zich op het functioneel rekenen dat noodzakelijk is voor participatie in de samenleving waarbij referentieniveau 2F gezien wordt als het minimale niveau dat mensen moeten beheersen om deel te kunnen nemen aan de maatschappij. De S-lijn richt zich op het streefniveau en kan omschreven worden als de inhoudslijn die leerlingen voorbereidt op de formele en abstracte wiskunde. 1F en 1S zijn de uitstroomniveaus van het basisonderwijs, behalve voor leerlingen die uitstromen naar het praktijkonderwijs en het speciaal onderwijs voor zeer moeilijk lerende leerlingen en meervoudig gehandicapte leerlingen. Hierbij geldt dat 1F het instroomniveau is voor vmbo-basis/kader en dat 1S het gewenste instroomniveau is voor vmbo-gemengd/theoretisch, havo en vwo. De ambitie is dat 85% van alle leerlingen aan het einde van het basisonderwijs fundamenteel niveau 1 (1F) en 65% streefniveau 1 (1S) behaalt. De eerste ambitie is inmiddels gerealiseerd, de tweede tot op heden niet. In de peiling Rekenen-wiskunde 2019 werd door 94% van de leerlingen 1F gehaald en door 47% 1S. In deze peiling is wederom gebruik gemaakt van het Referentiekader rekenen voor 1F en 1S, de andere niveaus zijn buiten beschouwing gelaten.

Rekenen-wiskunde in het primair onderwijs is gericht op 10 kerndoelen (23 tot en met 33). Daarnaast zijn er voor rekenen-wiskunde vier domeinen beschreven: getallen, verhouding, meten en meetkunde en verbanden. Aanvullend is met het oog op de verwachte herziening van het beoogde curriculum ook *probleemoplossen* meegenomen als nieuw domein in deze peiling. Probleemoplossen is onderdeel van de verwachte herziening en bovendien al opgenomen in de huidige kerndoelen. Bij de peiling in 2019 is er al voor gepleit om probleemoplossen als experimenteel onderdeel op te nemen, maar dit is om praktische en methodologische redenen niet gelukt. Op basis van voorgaande peilingsonderzoeken is de verwachting dat de verschillende vaardigheidsniveaus en de verschillende domeinen gezamenlijk één rekenvaardigheidsschaal vormen.

Naast de referentieniveaus en de domeinen is voor het opstellen van het toetsinstrumentarium rekening gehouden met verschillende inhoudsoverstijgende vaardigheden. Daarmee kan er afzonderlijk worden gerapporteerd op de beheersing van kale en contextopgaven, op de beheersing van hele getallen, decimale getallen en breuken, op de beheersing van exact en schattend rekenen en op de beheersing van het gebruik van de rekenmachine.

Om vergelijkingen met eerdere peilingen te kunnen maken, is gebruikgemaakt van items uit de eerste peiling Rekenen-wiskunde in het primair onderwijs (2019). Deze ankeropgaven zijn verdeeld over de verschillende domeinen en referentieniveaus. Tevens zijn in de vorige peiling 36 items uit de niet-openbare referentieset ingezet om de prestaties van de leerlingen op de referentieniveaus te bepalen. In deze peiling zijn de referentiecores zoals bepaald in de peiling van 2019 gebruikt. Dit wordt beschreven in hoofdstuk 6.

2.1.1 Beoogde itemconstructie

Op basis van de hiervoor omschreven theoretische uitgangspunten was beoogd om 119 nieuwe items te maken, verdeeld over de verschillende domeinen en kenmerken. In Tabel 2.1 is de beoogde verdeling gegeven. Van deze opgaven zijn er 94 verdeeld over de gebruikelijke vier domeinen (getallen, verhoudingen, meten en meetkunde en verbanden). Daarbinnen zijn er 16 opgaven gemaakt over de beheersing van het gebruik van de rekenmachine. Deze opgaven worden enkel in een paar toetsversies opgenomen en niet door alle leerlingen gemaakt. Daarnaast zijn er 25 nieuwe items gemaakt om de beheersing van probleemoplossen te peilen. Alle nieuwe opgaven zijn samen met een kleine set opgaven uit de vorige peiling door middel van de proeftoets getest.

Tabel 2.1 – De beoogde toetsmatrijs voor constructie, Peil.Rekenen-wiskunde PO 2023

Domein	Subdomein	Rekenonderwerp	Kaal	Context	Exact	Schattend	Hele getallen	Decimale getallen	Breuken	Rekenmachine	Totaal
Getallen	Getallen en getalbegrip		6	6	6	6	4	4	4	2	14
	Bewerkingen	Optellen en aftrekken	6	6	6	6	4	4	4	2	14
		Vermenigvuldigen en delen	6	6	6	6	4	4	4		12
		Combinaties	6	6	6	6	4	4	4		12
Verhoudingen	Verhoudingen		4	4	4	4				2	10
	Procenten		4	4	4	4				2	10
Meten en meetkunde	Meten	Lengte en omtrek								2	2
		Inhoud								2	2
		Tijd en snelheid								2	2
	Meetkunde				4	4					8
Verbanden					4	4				2	8
Probleem-oplossen											25

2.1.2 Gerealiseerd instrumentarium

De verdeling over de domeinen van het gerealiseerde instrumentarium is weergegeven in Tabel 2.2, inclusief de ankeritems. Ten aanzien van de leerlingen in het sbo en so4 is er rekening gehouden met het spanningsveld dat kan optreden tussen de wettelijk beoogde doelen en de daadwerkelijke capaciteiten van de leerlingen. Er is voor gekozen dat deze leerlingen een verkorte toets maken. Tabel 2.3 geeft de herkomst weer van de uiteindelijke itemverzameling, uitgesplitst naar domeinen.

Tabel 2.2 – De gerealiseerde toetsmatrijs, Peil.Rekenen-wiskunde PO 2023

Domein	Kaal	Context	Exact	Schattend	Hele getallen	Decimale getallen	Breken	Rekenmachine	Totaal	Percentage
Getallen	64	62	72	39	39	40	32	7	131	42
Verhoudingen	13	31	30	11				6	48	15
Metten en meetkunde			75	6				5	88	28
Verbanden			15	3				1	20	6
Probleem-oplossen									26	8

Tabel 2.3 – Herkomst van de items uitgesplitst naar domeinen, Peil.Rekenen-wiskunde PO 2023

	Getallen	Verhoudingen	Metten en meetkunde	Verbanden	Probleem-oplossen	Totaal
Peil 2019	50	26	71	12		159
TIMSS	12		1			13
Nieuw	69	22	16	8	26	141
Totaal	131	48	88	20	26	313

2.2 Afnamedesign proeftoets

Het design dat gebruikt is voor de proefafname wordt geïllustreerd in Tabel 2.4. In totaal waren er twaalf verschillende toetsversies, ook wel boekjes genoemd. De eerste zes boekjes zijn afgenomen op 'e-scholen' en bestaan uit drie verschillende itemsets van 20 items: twee daarvan (één digitaal en één op papier) bevatten ankeritems, de derde set (digitaal) bevat nieuw ontwikkelde items. Een leerling heeft nooit dezelfde opgave zowel op papier als digitaal hoeven maken. De papieren itemset is afwisselend als eerste en als tweede afgenomen. De derde itemset is altijd als laatste afgenomen en bevat een aantal nieuwe items waarvan de kwaliteit getest moest worden. De laatste zes boekjes zijn afgenomen op 'p-scholen' en bevatten twee itemsets van elk 20 nieuwe items die geheel digitaal zijn afgenomen.

In Tabel 2.4 is ook een kolom 'vragenlijst' opgenomen. Deze verwijst naar de leerlingvragenlijst die tijdens de proefafname alleen op p-scholen is afgenomen, vanwege de organisatorische complexiteit van de proefafname op de e-scholen.

Tabel 2.4 – Design van de proeftoets, Peil.Rekenen-wiskunde PO 2023*

Boekje	n	Modus-/Equivalentiestudie						Nieuwe opgaven						vragenlijst
		1	2	3	1	2	3	4	5	6	7	8	9	
Boekje 1	130	20	20					20						
Boekje 2	116		20	20					20					
Boekje 3	136			20	20					20				
Boekje 4	106				20	20					20			
Boekje 5	114					20	20					20		
Boekje 6	129	20					20						20	
Boekje 7	41							20	20					Ja
Boekje 8	49								20	20				Ja
Boekje 9	72									20	20			Ja
Boekje 10	64										20	20		Ja
Boekje 11	51											20	20	Ja
Boekje 12	40							20					20	Ja

* De kolommen genaamd 1-9 geven de verschillende itemsets. De blauwe cellen geven aan dat de items op papier zijn afgenomen. De getallen in de cellen geven het aantal items in een blokje. Dit design is exclusief 4 experimentele items, deze zijn verdeeld over de boekjes met nieuwe opgaven.

2.3 Resultaten proeftoets

Het resultaat van de equivalentie- of modusstudie wordt besproken in hoofdstuk 6. De nieuwe opgaven in de proeftoets zijn geëvalueerd aan de hand van klassieke itemstatistieken. De items in het domein probleemoplossen bleken duidelijk moeilijker dan de andere domeinen (Tabel 2.5). Daarnaast lijken de nieuwe ontwikkelde opgaven over het algemeen iets moeilijker dan de ankeropgaven (idem). Op basis van de proeftoets is besloten om 8 items niet op te nemen in het design voor het peilingsonderzoek.

Tabel 2.5 – Gemiddelde p-waarde en item-rest correlatie proeftoets, Peil.Rekenen-wiskunde PO 2023*

Itemset	Schooltype	Domein	items	rir	p-waarde
anker	bao	getallen	38	0.37	0.64
		meten en meetkunde	40	0.31	0.69
		verbanden	16	0.36	0.60
		verhoudingen	26	0.47	0.69
nieuw	bao	getallen	62	0.30	0.49
		meten en meetkunde	11	0.30	0.47
		probleemoplossen	26	0.36	0.33
		verbanden	7	0.29	0.52
		verhoudingen	18	0.42	0.54
nieuw	sbo	getallen	62	0.31	0.30
		meten en meetkunde	11	0.30	0.30
		probleemoplossen	26	0.33	0.14
		verbanden	7	0.27	0.34
		verhoudingen	18	0.42	0.26

* Op papier en digitaal afgenomen items zijn als verschillende items geteld.

2.4 Afnamedesign hoofdonderzoek

Voor het hoofdonderzoek zijn 12 boekjes voor het bao en 8 boekjes voor het sbo en so4 gebruikt (Tabel 2.6). Naast trenditems en items uit de proeftoets zijn in het hoofdonderzoek 20 items opgenomen die afkomstig zijn uit de reken-wiskundetoets van TIMSS-2023 voor groep 6. Hiervoor is toestemming verkregen van de IEA (*International Association for the Evaluation of Educational Achievement*)².

2 Trends in International Mathematics and Science Study – TIMSS-2023 Copyright © 2024 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College.

Tabel 2.6 – Samenstelling van de boekjes in het hoofdonderzoek, Peil.Rekenen-wiskunde PO 2023 *

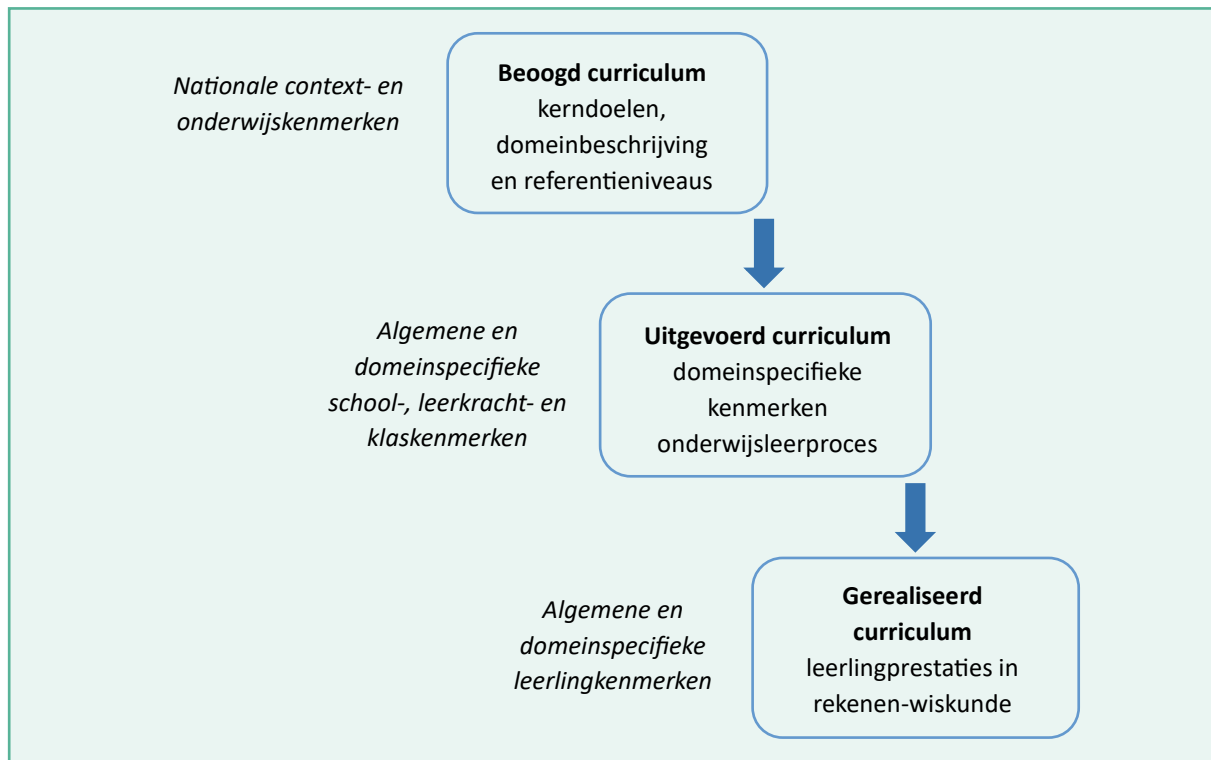
Schooltype	domein	context	kaal	breuken	decimale getallen	hele getallen	Exact	Schattend	rekenmachine	totaal
bao 10 boekjes	getallen	5-13	4-13	4-8	3-9	3-8	9-13	2-8		17-20
	meten en meetkunde	9-12		0-1	1-5	3-7	9-12	0-2		11-14
	probleemoplossen	3-4								3-4
	verbanden	1-3				1-3	1-3	0-1		2-4
	verhoudingen	3-7	1-3	0-2		1-4	3-6	0-3		5-8
bao rekenmachine 2 boekjes	getallen	4-6	9-10	2-3	5-6	3-4	4-7	4-8	5-7	19-20
	meten en meetkunde	7			1-3	3-4	7-8		5	12-13
	probleemoplossen	3								3
	verbanden	1-2				1-2	2		1	3
	verhoudingen	2-4	1	1-2		1	2-3	0-1	4-6	7-9
sbo/so4 6 boekjes	getallen	7-10	4-7	2-4	3-7	4-8	8-9	3-5		13-14
	meten en meetkunde	8-13		0-1	1-3	5-7	9-11	0-2		10-14
	probleemoplossen	2-3								2-3
	verbanden	1-3				1-2	1-2	0-2		2-3
	verhoudingen	2-5	1-4	0-2		3-4	3-5	0-3		5-6
sbo/so4 rekenmachine 2 boekjes	getallen	6-7	3	0-2	1-3	4-6	6	1-3	5-7	14-15
	meten en meetkunde	7			1-2	3-4	6-7	0-1	5	12
	probleemoplossen	1-2								1-2
	verbanden	0-1				0-1	1		1	2
	verhoudingen	4		0-2		1-2	1-2	1	4-6	8

* De cellen geven de minimum en maximum aantallen items in de boekjes voor de betreffende categorieën.

3. Ontwikkeling vragenlijsten

3.1 Conceptueel model

Aan de instrumentontwikkeling voor deze peiling, en dan met name de vragenlijsten, ligt het volgende globale conceptuele model ten grondslag, dat ook het uitgangspunt vormt van het internationale peilingsonderzoek TIMSS (Mullis & Martin, 2017).



Figuur 3.1 – Conceptueel model Peil.Rekenen-wiskunde PO 2023 (gebaseerd op Mullis & Martin, 2017).

Het beoogde curriculum is in dit peilingsonderzoek een gegeven, uitgewerkt in de domeinbeschrijving (Van Zanten & Oldengarm, 2020) en in de referentieniveaus die voor rekenen-wiskunde zijn vastgesteld. Kenmerken van de rekenlessen (domeinspecifieke onderwijskenmerken) kunnen worden beschouwd als indicatoren van het uitgevoerde curriculum. De leerlingprestaties op de peilingstoets representeren het gerealiseerde curriculum. Tot slot laat dit model ook de mogelijke relaties zien tussen de algemene en domeinspecifieke kenmerken op school, klas- en leerlingniveau, het uitgevoerde curriculum en het gerealiseerde curriculum. In de multilevel analyses (Hoofdstuk 9) worden deze mogelijke relaties nader geëxploreerd.

3.2 Uitgangspunten constructie vragenlijsten

De samenstelling van de vragenlijsten is gebaseerd op het bovengenoemde conceptuele model, de onderzoeksvragen van deze peiling en de adviezen voortkomend uit de reviewstudies van Hickendorff

et al. (2017; 2020) over effectieve rekeninstructie. De conceptvragenlijsten voor de leerling en de leerkracht zijn vervolgens voorgelegd aan een externe expert en de begeleidingscommissie van deze peiling. Kenmerken op schoolniveau zijn voor ba0 gemeten met de schoolvragenlijst van TIMSS. In het sbo en so4 is een verkorte en aangepaste versie van de TIMSS-schoolvragenlijst afgenomen.

Voor de ontwikkeling van vragen is zoveel mogelijk gebruikgemaakt van de vragenlijsten van de vorige rekenpeiling en TIMSS³. Dit heeft als voordeel dat er informatie bekend was over wat deze vragen aan uitkomsten kunnen opleveren en over de consistentie en betrouwbaarheid van de constructen die zijn samengesteld uit meerdere items. Het biedt tevens de mogelijkheid om veranderingen in het onderwijsleerproces tussen 2019 en 2023 of tussen groep 6 en 8 in kaart te brengen; bijvoorbeeld wat betreft de tijd die aan rekenonderwijs wordt besteed, het gebruik van digitale apparaten en de hoeveelheid rekenhuiswerk die leerlingen krijgen. Het vergelijken van de bevindingen over het onderwijsleerproces uit 2023 met de uitkomsten van de peiling van 2019 of van TIMSS 2023 (groep 6) maakt echter geen deel uit van deze peiling.

Op basis van de adviezen van een externe expert en de begeleidingscommissie is er ook een beperkt aantal nieuwe vragen of items opgenomen. Om er voor te zorgen dat de respondent gemotiveerd is (en blijft) om de vragenlijst zo volledig mogelijk in te vullen, is geprobeerd de gevraagde inspanning van de respondent zoveel mogelijk te beperken. Het gaat hierbij zowel om de lengte van de vragenlijsten als de wijze waarop de vragen worden voorgelegd. Dit laatste betekent zo weinig mogelijk open vragen, afwisseling van vraagvormen en het – op basis van eerder gegeven antwoorden – automatisch overslaan van vragen die niet van toepassing zijn op de respondent.

3.3 Vragenlijsten proefafname

In het voorjaar van 2022 (maart tot half april) zijn de toets en alle (online) vragenlijsten uitgetest in de proefafname. De proefafname in het regulier basisonderwijs is gelijktijdig en op dezelfde scholen uitgevoerd als de proefafname van TIMSS-2023. De steekproef voor de proefafname (2022) en hoofdonderzoek (2023) is door TIMSS in één keer getrokken om overlap te voorkomen en is half november 2021 getrokken op basis van de op dat moment meest actuele beschikbare DUO-data (zie verder hoofdstuk 4).

Voor de proefafname van TIMSS en dit peilingsonderzoek zijn in eerste instantie 35 reguliere basisscholen voor beide onderzoeken benaderd. In geval van weigering was per basisschool één vervanger beschikbaar die in diens plaats werd benaderd. De werving voor de proefafname begon eind 2021, toen de COVID-19 pandemie nog gaande was. Vanwege de gevolgen daarvan (zoals lesuitval) was de deelnamebereidheid voor de proefafname zeer laag. Om toch voldoende data te verkrijgen voor de uitvoering van de equivalentiestudie is aan TIMSS gevraagd voor de proefafname nogmaals 2 x 35 scholen te trekken. In totaal hebben 39 groepen 8 (731 leerlingen) afkomstig van 29 basisscholen aan de proefafname deelgenomen.

3 Voor het gebruik van vragen uit de TIMSS-vragenlijsten is toestemming verkregen van IEA. Referentie: *Trends in International Mathematics and Science Study* – TIMSS-2007/2015/2019/2023 Copyright © 2007/2015/2019/2023 International Association for the Evaluation of Educational Achievement (IEA). Publisher: TIMSS & PIRLS International Study Center, Lynch School of Education and Human Development, Boston College.

Voor de proefafname in het sbo en so4 is geen steekproef getrokken, maar zijn 8 so4- en 8 sbo-scholen uit het netwerk van het consortium bereid gevonden om aan de proefafname deel te nemen. De equivalentiestudie is uitgevoerd in het bao, in sbo en so4 was de afname volledig digitaal. Omdat in de equivalentiestudie de toets al uit drie sessies bestond, is de leerlingvragenlijst alleen in het sbo en so4 afgenomen. In de proefafname hebben 25 schoolleiders (bao en s(b)o), 52 leerkrachten (bao en s(b)o) en 312 leerlingen (s(b)o) de vragenlijsten ingevuld.

De vragenlijstdata van de proefafname zijn geanalyseerd door te kijken naar frequentieverdelingen, antwoorden op open vragen en ontbrekende waarden. Voor de constructen in de leerlingvragenlijst (alleen gevalideerde schalen), is de betrouwbaarheid van de schalen (λ^2) berekend om na te gaan of deze schalen ook bruikbaar waren voor het s(b)o.

De review op basis van de proefafname-data resulteerde erin dat de leerkrachtvragenlijst voor het hoofdonderzoek is ingekort vanwege het grote aantal missende waarden in met name het laatste deel van de vragenlijst.

3.4 Vragenlijsten hoofdonderzoek

De online leerlingvragenlijsten voor bao, sbo en so4 die in het hoofdonderzoek van 2023 zijn afgenomen, zijn grotendeels identiek. De versies bestemd voor de sbo- en so4-leerlingen hadden alleen een uitgebreidere invulinstructie dan de versie voor groep 8. De online leerkrachtvragenlijsten voor de drie onderwijssoorten zijn eveneens gelijk, met uitzondering van de vraag naar de onderwijservaring van de leerkracht, deze is het s(b)o iets uitgebreider (zie Tabel 3.2). Omdat in het reguliere basisonderwijs de schoolvragenlijst van TIMSS is afgenomen (naast rekenen-wiskunde ook gericht op natuurwetenschappen) zijn de schoolvragenlijsten voor sbo en so4 korter dan de bao-versie.

De invultijd van vragenlijsten voor het hoofdonderzoek was voor de leerlingvragenlijst ongeveer 25 minuten en voor de leerkrachtvragenlijst 30 minuten. De online schoolvragenlijst voor bao (TIMSS-versie) kostte ongeveer 30 minuten om in te vullen, voor s(b)o was de invultijd ongeveer 25 minuten. Daarnaast hebben de leerkrachten van de getoetste klassen (of een collega, zoals een administratieve kracht of IB'er) vooraf de zogenoemde leerlingadministratielijst ingevuld waarmee aanvullende informatie over de leerlingen is verzameld: geslacht, geboortedatum (maand en jaar), LVS-score voor begrijpend lezen, voorlopig uitstroomadvies en reden voor eventuele uitsluiting van de toets.

Hieronder wordt een overzicht gegeven van de contextinformatie die in het hoofdonderzoek met (voornamelijk) de vragenlijsten is verzameld.

Tabel 3.1 – Variabelen leerlingvragenlijst, Peil.Rekenen-wiskunde PO 2023

Categorie	Kenmerk	Gemeten als:*
Algemene leerlingkenmerken	Geslacht	Geslacht (meisje, jongen anders)**
	Leeftijd/vertraagd	Geboortjaar en maand**
	Thuis taal	Frequentie thuis Nederlands gesproken
	Sociaal-cultureel kapitaal	Aantal boeken thuis
	Vaardigheid begrijpend lezen	LVS-score begrijpend lezen groep 7 inclusief versie en generatie***
	Voorlopig schooladvies	VO-advies of uitstroomprofiel***
Domeinspecifieke leerlingkenmerken	Attitudes	<i>Plezier in rekenen</i> <i>Zelfvertrouwen in rekenen</i>
	Attributies positieve prestaties	<i>Attributies positieve rekenprestaties aan eigen inzet</i>
	Mindset	<i>Fixed/growth mindset t.o.v.rekenen</i>
Domeinspecifieke kenmerken onderwijsleerproces (uitgevoerd curriculum)	Ordelijkheid les	<i>Ordelijkheid rekenles</i>
	Helderheid instructie	<i>Helderheid instructie rekenles</i>
	Zelfstandig werken	Frequentie zelfstandig werken

* Kenmerken die gemeten zijn aan de hand van een set van items (constructen) zijn cursief weergegeven; ** Zowel gemeten met de leerlingvragenlijst als leerlingadministratielijst; Leerlingen met een geboortedatum voor 1 oktober 2010 zijn aangeduid als vertraagd. ***Gemeten met leerlingadministratielijst

Tabel 3.2 – Variabelen leerkrachtvragenlijst, Peil.Rekenen-wiskunde PO 2023

Categorie	Kenmerk	Gemeten als:*
Algemene leerkrachtkenmerken	Geslacht	Geslacht (vrouw, man, anders)
	Leeftijd	Leeftijd in categorieën
	Opleiding	Hoogste opleiding met diploma afgerond Lerarenopleiding(en) met diploma afgerond Specialisatie rekenen-wiskunde tijdens lerarenopleiding Post HBO rekencoördinator gevolgd
	Leservaring	Gemiddeld aantal jaar lesgegeven in betreffende onderwijssoort S(b)o: ook lesgegeven in andere onderwijssoorten
Algemene schoolkenmerken (perspectief leerkracht)	Onderwijsconcept	Onderwijsconcept
	Prestatiegerichtheid school	<i>Prestatiegerichtheid school (alle items)</i> <i>Prestatiegerichtheid school: leerkrachten</i> <i>Prestatiegerichtheid school: ouders</i> <i>Prestatiegerichtheid school: leerlingen</i>
	Werkdruk	<i>Door leerkracht ervaren werkdruk op school</i>
Domeinspecifieke leerkrachtkenmerken	Ondersteuning rekencoördinator	Ervaren ondersteuning rekencoördinator
	(Bij)scholing rekenen-wiskunde	Inhoud (bij)scholing rekenen-wiskunde afgelopen twee jaar Toekomstige behoefte (bij)scholing Tijdsbesteding (bij)scholing afgelopen twee jaar
	Zelfvertrouwen	<i>Zelfvertrouwen vakdidactische vaardigheden rekenen</i>
	Mindset ten opzichte van leren rekenen	<i>Fixed/growth mindset</i>
	Geschiktheid rekenmethode	<i>Oordeel leerkracht geschiktheid reken-wiskundemethode</i>
	Beoordelingsstrategieën	Gepercipieerd belang leerkracht beoordelingsstrategieën voor rekenen

* Kenmerken die gemeten zijn aan de hand van een set van items (constructen) zijn cursief weergegeven

Tabel 3.2 (vervolg) – Variabelen leerkrachtvragenlijst, Peil.Rekenen-wiskunde PO

Categorie	Kenmerk	Gemeten als:*
Domeinspecifieke onderwijskenmerken (uitgevoerd curriculum)	Rekencoördinator	Aanwezigheid rekencoördinator
	Onderwijstijd rekenen-wiskunde	Minuten rekenonderwijs gemiddeld per week
	Methodegebruik	Hoofdzakelijk gebruikte methode voor instructie en verwerking rekenonderwijs Wel/niet traditionele methode
	Werkvormen tijdens rekenlessen	<i>Leerkracht-gestuurde klassikale instructie</i> Werken in heterogene en/of homogene vaardigheidsgroepen
	Instructiemodel	Frequentie directe instructie Frequentie ontdekkend of onderzoekend leren
	Oplossingsstrategieën rekenen	Aanbod oplossingsstrategieën rekenen Behandelde strategieën lang (kolom) optellen, aftrekken, vermenigvuldigen en delen Behandelde strategieën kort (cijferend) optellen, aftrekken, vermenigvuldigen, delen <i>Behandelde strategieën lang (zonder delen)</i> <i>Behandelde strategieën kort (zonder delen)</i>
	Digitale apparaten in de rekenles	Gebruik digitale apparaten Wijze van toegang tot digitale apparaten Frequentie gebruik softwareprogramma's-/rekenwebsites Gebruikte softwareprogramma's/rekenwebsites <i>Leeractiviteiten met digitale apparaten</i>
	Rekenmachine in de rekenles	Mate waarin gebruik rekenmachine is toegestaan <i>Leeractiviteiten met rekenmachine</i>
	Huiswerk	Frequentie huiswerk rekenen <i>Feedback huiswerk rekenen</i>
	Evalueren rekenprestaties	<i>Evaluatie</i>
	Differentiatie rekenonderwijs	<i>Differentiatie</i>

* Kenmerken die gemeten zijn aan de hand van een set van items (constructen) zijn cursief weergegeven

Tabel 3.3 – Variabelen schoolvragenlijst, Peil.Rekenen-wiskunde PO 2023

Categorie	Kenmerk	Gemeten als:*
Algemene schoolkenmerken	Schoolgewicht	<i>Schoolgewicht**</i>
	Schoolgrootte	<i>Schoolgrootte**</i>
	Belemmeringen onderwijs	Belemmering door gebrek aan of gebrekkige digitale apparaten
	Prestatiegerichtheid school	<i>Prestatiegerichtheid school (alle items)</i> <i>Prestatiegerichtheid school: leerkrachten</i> <i>Prestatiegerichtheid school: ouders</i> <i>Prestatiegerichtheid school: leerlingen</i>
	COVID-19 pandemie	Duur scholensluiting
Domeinspecifieke schoolkenmerken	Belemmeringen rekenonderwijs	Belemmering door gebrek/onvolkomenheden leermaterialen rekenonderwijs
Domeinspecifieke onderwijskenmerken (uitgevoerd curriculum)	Rekencoördinator	Beschikbaarheid Taken: kwaliteitsbewaking Taken: zorgniveau Taken: leerkrachtniveau

* Kenmerken die gemeten zijn aan de hand van een set van items (constructen) zijn cursief weergegeven; **Op basis van Duo-bestanden 2021

4. Steekproefdesign

4.1 Steekproef regulier basisonderwijs

Voor de uitvoering van het hoofdonderzoek in het regulier basisonderwijs (groep 8) zijn dezelfde scholen benaderd als die voor TIMSS-2023 (groep 6). Dit betekent dat de steekproef is getrokken door *RTI International* in opdracht van *TIMSS & PIRLS International Study Center*. Hiervoor heeft het consortium gegevens aangeleverd over het Nederlandse onderwijssysteem en in het najaar van 2021 een scholenbestand aangeleverd op basis van DUO-bestanden (aantal 8-jarigen per school, schooljaar 2020-2021). De steekproef voor de proefafname en hoofdonderzoek zijn in één keer getrokken om overlap te voorkomen. In het reguliere basisonderwijs is een gestratificeerde steekproef getrokken op basis van het schoolgewicht van de school met een trekkingskans proportioneel naar grootte van de school (gebaseerd op het aantal 8-jarigen per school). Scholen met minder dan zes (8-jarige) leerlingen zijn uitgesloten van de steekproef.

Voor het hoofdonderzoek zijn 152 scholen (eerste tranche) getrokken. Om te worden opgenomen in de internationale vergelijkingstabellen, vereist TIMSS dat minimaal 85% van de scholen uit de eerste tranche daadwerkelijk deelneemt. In Nederland (en veel andere TIMSS-landen) is dit percentage niet haalbaar. Om deze reden is gebruikgemaakt van reservescholen. Voor elke originele school in het hoofdonderzoek heeft TIMSS gelijktijdig twee reservescholen getrokken (tweede en derde tranche). Dit zijn geen willekeurige scholen, maar scholen die qua schoolgewicht en schoolgrootte sterk overeenkomen met de originele school. Een reserveschool mag pas worden benaderd als de originele school deelname weigert.

Tabel 4.1 toont hoe de scholen in de onderzoekspopulatie en in de steekproef verdeeld zijn over de vier strata voor schoolgewicht.

Tabel 4.1– Verdeling populatie en steekproef bao naar schoolgewicht, Peil.Rekenen-wiskunde PO 2023

Schoolgewicht bao 2021	% scholen in populatie (n=6080)	% scholen in steekproef (eerste tranche, n=152)
1= Hoog (> 36)	5,3	5,3
2= Gemiddeld/hoog (30-36)	39,7	34,2
3= Gemiddeld/laag (25-30)	45,1	46,1
4= Laag (< 25)	9,9	14,5
Totaal	100	100

In TIMSS kunnen landen kiezen of zij binnen een getrokken school willekeurig één of twee klassen trekken of dat alle klassen van de doelpopulatie (*grade 4*, in Nederland groep 6) binnen een school worden benaderd voor het onderzoek. In Nederland wordt sinds TIMSS-2007 aan de scholen gevraagd om met alle aanwezige klassen deel te nemen. Hiermee wordt voorkomen dat er (nog) meer dan de huidige 152 basisscholen moeten worden getrokken om aan de TIMSS-eis te voldoen van minimaal 4.000 deelnemende leerlingen per land. Bovendien is gebleken dat, bij het random

trekken van één of twee klassen, dit tot discussies met de basisscholen kan leiden over welke klas meedoet. Scholen met meerdere groepen 6 of 8 vinden bovendien het kunnen vergelijken van de uitkomsten tussen hun klassen vaak van toegevoegde waarde. De geselecteerde basisscholen zijn dus benaderd met het verzoek om met alle aanwezige groepen 6 en 8 deel te nemen.

4.2 Steekproef sbo en so4

De steekproef voor speciaal basisonderwijs is getrokken door de Inspectie van het Onderwijs op basis van de stratificatievariabelen regio en uitstroomprofiel. De doelpopulatie omvatte 267 vestigingen. De steekproef voor het hoofdonderzoek bestond uit 65 scholen plus twee reservetranches van 65 scholen. Een reserveschool is pas benaderd als de originele school deelname weigerde. Indien er op deze scholen sprake was van meerdere groepen schoolverlaters, werden ze benaderd met het verzoek met al deze groepen aan het peilingsonderzoek deel te nemen.

Tabel 4.2 toont hoe de scholen in de steekproef verdeeld zijn over de drie strata voor uitstroomprofiel. De gemiddelde uitstroom naar vso/pro van de doelpopulatie was 49,9% (sd=15,3), die van de scholen in de steekproef 50,3% (sd=15,4).

Tabel 4.2 – Verdeling steekproef sbo naar uitstroomprofiel, Peil.Rekenen-wiskunde PO 2023

Percentage uitstroom vso/pro	sbo-scholen in de steekproef
	%
1= Scholen met een gemiddelde uitstroom naar vso/pro van 40% of lager	27,6
2= Scholen met een gemiddelde uitstroom naar vso/pro tussen 40 en 60%	41,5
3= Scholen met een gemiddelde uitstroom naar vso/pro van 60%	30,8
Totaal	100 (n=65)

De omvang van de totale populatie so4-scholen is vooraf op ongeveer 90 scholen geschat. Er is daarom geen steekproef getrokken, maar alle so4-scholen (minus de scholen die aan de proefafname hadden deelgenomen) zijn benaderd voor deelname aan het hoofdonderzoek. Indien op deze scholen sprake was van meerdere groepen met schoolverlaters (en uitstroomprofiel vervolgonderwijs), is aan de scholen gevraagd om met hen allemaal aan het peilingsonderzoek deel te nemen. Enkele scholen in de steekproef bleken schoolverlatersgroepen te hebben met leerlingen uit zowel voormalig cluster 3 als cluster 4. Hiervan heeft één school aan het peilingsonderzoek deelgenomen.

4.3 Uitsluitingscriteria

De leerlingen in het sbo en so4 hebben de relatief ‘makkelijkere’ toetsversies (met meer opgaven op eenvoudig niveau) voorgelegd gekregen. In het reguliere basisonderwijs zijn zowel ‘makkelijkere’ als ‘moeilijkere’ toetsversies afgenomen. Ondanks dat er met deze toewijzing van toetsversies rekening is gehouden met onderwijssoort, zijn niet alle leerlingen in staat om de peilingstoets te maken. Dit geldt zowel voor leerlingen in het reguliere basisonderwijs als in het sbo en so4.

Vooraf konden leerkrachten op de leerlingadministratielijsten aangeven of er leerlingen moesten worden uitgesloten van toetsdeelname. Een leerling mocht alleen mag worden uitgesloten van toetsdeelname als deze ook niet in staat is om de reguliere rekentoetsen van de klas te maken, vanwege a) ernstige fysieke beperkingen; b) ernstige intellectuele beperkingen en/of c) minder dan één jaar ervaring met de Nederlandse taal. In het s(b)o kunnen er nog andere beperkingen zijn waardoor een leerling niet in staat is om de peilingstoets te maken. Vooraf is hierover met de betreffende leerkracht overlegd. Leerlingen die om bovengenoemde redenen zijn uitgesloten maken geen deel uit van de steekproef, evenals leerlingen die tussentijds (na aanmelding en vóór afname) de school hebben verlaten.

5. Werving, dataverzameling en respons hoofdonderzoek

5.1 Organisatie scholenwerving

Er is een intensieve wervingsprocedure ingezet om ervoor te zorgen dat er voldoende scholen en leerlingen aan het onderzoek zouden deelnemen. De werving startte met een aankondiging van het onderzoek in digitale nieuwsbrieven en op sociale media. De scholen uit de steekproef ontvingen vervolgens een informatiebrief voor de schoolleider waarin het belang en de praktische organisatie van het onderzoek werden beschreven en waarin zij uitgenodigd werden om zich aan te melden voor het onderzoek. Bij deze uitnodiging was een aanbevelingsbrief gevoegd, ondertekend door de minister voor Primair en Voortgezet onderwijs, waarin nogmaals het belang en de anonimiteit van het onderzoek werd benadrukt. Deze brieven zijn in het najaar van 2022 zowel op papier als digitaal verstuurd naar de schoolleiders van alle scholen uit de hoofdstekproef voor het basisonderwijs en voortgezet onderwijs en aan alle so4-scholen met uitzondering van de deelnemers aan de proefmeting. In de brieven werd verwezen naar een speciaal ontwikkelde website waarop informatie over het onderzoek op een overzichtelijke manier werd gepresenteerd voor scholen (<https://timss-en-peil.nl/>). In de brieven en op de website werd benadrukt dat de onderzoekbelasting voor de deelnemende scholen zo laag mogelijk zou worden gehouden, bijvoorbeeld door het verstrekken van voorbeeldbrieven waarmee scholen betrokkenen (leerlingen, leerkrachten, ouders) kunnen informeren en het inschakelen van onderzoeksassistenten om de afname bij de leerlingen te organiseren. Ook werden de tegenprestaties voor deelname aan het onderzoek benoemd, bestaande uit kleine presentjes voor de leerlingen en voor de school een terugkoppeling van de resultaten van de eigen leerlingen in vergelijking tot landelijke gemiddelden in de vorm van een schoolrapport. Bij volledige deelname aan alle onderdelen van het onderzoek dongen scholen ook mee naar de verloting van een aantal Chromebooks.

Enkele dagen na het versturen van de brieven werd de schoolleider van elke benaderde school telefonisch benaderd door een onderzoeksassistent van KBA Nijmegen. In dit telefoongesprek werd desgewenst meer informatie gegeven over het onderzoek en werd de school gevraagd om hieraan deel te nemen. De assistenten voerden dit gesprek aan de hand van een belscript, waarin ook de antwoorden op veel gestelde vragen waren opgenomen. Van tevoren hadden zij een uitgebreide training ontvangen over het onderzoek, de wijze van benaderen van de schoolleiders en het werken met het belscript. Ook waren met de assistenten afspraken gemaakt over het gebruik van het digitale logboek om contacten met de scholen te registreren.

De onderzoeksassistenten ondernamen soms een groot aantal belpogingen om een school bereid te vinden zich aan te melden voor het onderzoek. Onder een belpoging verstaan we de actie dat een school is gebeld, ongeacht of de schoolcontactpersoon daadwerkelijk is gesproken. Uit de logboekdata blijkt dat een ruime meerderheid van de uiteindelijk deelnemende scholen maximaal tien keer gebeld is voordat de toestemming tot deelname volledig rond was. Bij één school waren echter 48 belpogingen nodig om de aanmelding voor het onderzoek volledig af te ronden.

Voor zowel het basisonderwijs als het voortgezet onderwijs geldt dat wanneer een school zich afmeldde voor het onderzoek, de eerste vervangende school werd benaderd. Voor so4-scholen was er geen vervangende steekproef aangezien alle (geschikte) so4-scholen direct benaderd werden. De wervingsprocedure van

Peil.Rekenen-wiskunde PO 2023 was (voor het baso-deel) gekoppeld aan TIMSS. Tegen het einde van de wervingsperiode bleek dat de internationale responseis van TIMSS niet gehaald zou worden. Daarop is besloten om alle basisscholen uit de eerste tranche opnieuw te benaderen met de vraag of ze alsnog konden deelnemen aan het TIMSS-onderzoek. Hierdoor komt het voor dat een school uit de vervangende steekproeven al deelname had toegezegd als de (gekoppelde) school uit de hoofdstekproef alsnog toestemde in deelname met groep 6 aan het TIMSS-onderzoek. Dit had echter geen gevolgen voor Peil.Rekenen-wiskunde PO 2023, omdat dergelijke opnieuw benaderde scholen dan niet met groep 8 deelnamen.

Scholen die besloten deel te nemen aan het onderzoek kregen het verzoek om hun aanmelding te bevestigen via een digitale aanmeldlink. Tegelijkertijd konden ze daarbij de gegevens van de contactpersoon voor het onderzoek op de school verstrekken en enkele administratieve gegevens zoals het aantal groepen 8 en het aantal leerlingen per groep opgeven. Aan de scholen die zich hadden aangemeld is gevraagd om de ouders van de leerlingen in de te toetsen klassen te informeren over het onderzoek. Hiervoor is door de onderzoekers een brief aangeleverd die de scholen aan de ouders konden sturen. Indien ouders geen toestemming gaven voor deelname van hun kind aan het peilingsonderzoek, konden zij dat aan de onderzoekers of de school doorgeven. Deze leerlingen werden uitgesloten van deelname aan het onderzoek.

5.2 Dataverzameling

5.2.1 Voorbereidende dataverzameling

Bij de deelnemende scholen is vooraf contactinformatie en informatie over het aantal klassen, het aantal leerlingen en de ICT-infrastructuur (aanwezige Chromebooks, internetverbinding) verzameld met behulp van het digitale aanmeldformulier (zie ook hierboven). Scholen konden daarnaast drie voorkeursdata opgeven voor de afname van de toets en leerlingvragenlijst op hun school.

Voor het verkrijgen van enkele achtergrondgegevens van leerlingen is aan de leerkracht van elke te toetsen klas (of een collega, zoals de IB'er) gevraagd om twee leerlingadministratielijsten in te vullen. Op beide lijsten stonden de vooraf toegekende leerlingnummers (op basis van het opgegeven aantal leerlingen in de klas), de toegewezen toetsversie en de individuele inlogcode al vermeld. Op de eerste lijst vulde de leerkracht vervolgens de voornamen in van de leerlingen in de klas. Deze lijst is tijdens de toetsafname door de toetsleider gebruikt voor het uitdelen van de inlogcodes voor de leerlingen en bleef na de afname achter bij de leerkracht. Hiermee was deze de enige die de koppeling leerlingnummer en leerling kende. Op de tweede administratielijst vulde de leerkracht de volgende gegevens in over elke leerling uit de klas: geslacht, geboortedatum (maand en jaar), LVS-score voor begrijpend lezen (inclusief toetsversie en -generatie), voorlopig uitstroomadvies en reden voor eventuele uitsluiting van de toetsafname. Tijdens de afname gaf de toetsleider op deze lijst aan of de leerling de toets en vragenlijst had gemaakt, de eventuele reden voor afwezigheid (indien van toepassing) en andere bijzonderheden. De informatie op deze anonieme lijst is na afloop ingevoerd in een databestand. Daarnaast vulde de toetsleider tijdens en kort na de afname het Toetsafnameformulier (TAF) in. Hierop worden de begin- en eindtijden van de afname genoteerd, veel voorkomende vragen van leerlingen en of er zich bijzonderheden voordeden tijdens de afname. Ook deze informatie is verwerkt in een databestand.

5.2.2 Training toetsleiders

Alle afnames van de toets en leerlingvragenlijst zijn uitgevoerd door toetsleiders van het consortium. Om er voor te zorgen dat de afnameomstandigheden op de scholen zoveel mogelijk hetzelfde waren zijn de toetsleiders getraind aan de hand van een twee uur-durende training die bestond uit drie onderdelen. Om toetsleiders vertrouwd te maken met het onderzoek, ontvingen ze eerst algemene informatie over het peilingsonderzoek en TIMSS. Hierbij werd stilgestaan bij de historie van de onderzoeken, de inhoud van de onderzoeken en waarom deze onderzoeken belangrijk zijn. Ook kregen ze uitleg over de verschillende schooltypen in het onderzoek (bao, sbo en so4) en de leerlingen op deze scholen. In het tweede deel werd ingegaan op de taken van de toetsleider en op de stappen die ze per school (ophalen materialen, afnemen onderzoek, sorteren materialen, terugbrengen materialen) op welk moment (week voor de afname, dag voor de afname, dag van de afname, na de afname) moesten doorlopen. Tenslotte werd de afnamesessie en alle aandachtspunten die daarbij hoorden met de toetsleiders besproken. Hierbij werd extra aandacht besteed aan de afname op de sbo- en so4-scholen en de zaken waarmee toetsleiders bij de leerlingen van deze scholen rekening moesten houden, zoals aanpassingen aan de afnameprocedure (bijvoorbeeld het inlassen van extra of langere pauzes) of het voorkomen van toetsangst. De toetsleiders werden erin getraind hierover in gesprek te gaan met de scholen en de leerkrachten. Naast de training is er ook een toetshandleiding ontwikkeld voor de toetsleiders. Omdat de toetsleiders in het basisonderwijs op dezelfde dag zowel de toetsen en leerlingvragenlijst in groep 6 als in groep 8 afnamen, was de afnameprocedure in beide groepen zoveel mogelijk hetzelfde. De toetshandleiding van TIMSS, met daarin de voorgeschreven procedures, is daarom ook als basis gebruikt voor de toetshandleidingen voor groep 8, sbo en so. Tijdens de proefafname heeft één van de onderzoekers (verantwoordelijk voor de training en handleidingen) een aantal van de afnames zelf uitgevoerd.

5.2.3 Afname toets en leerlingvragenlijst

Alle afnames bij de leerlingen hebben in 2023 plaatsgevonden tussen week 10 en week 25. Aanvankelijk was beoogd dat de afnameperiode zou lopen tot 1 juni (week 22), maar deze moest iets worden verlengd als gevolg van de langer lopende wervingsperiode. Alle scholen werden in genoemde periode bezocht door getrainde toetsleiders die de online toets en vragenlijsten bij de leerlingen afnamen. In de meeste gevallen gebeurde dat op devices van de school zelf. Voor de (weinig) scholen die zelf niet over voldoende apparaten beschikten, hadden de toetsleiders extra Chromebooks bij zich. Ook namen ze extra routers mee om eventuele problemen met het internet te voorkomen.

Op de meeste reguliere basisscholen was het mogelijk om op dezelfde dag de afnames in groep 6 en groep 8 te plannen. In sbo en so4 legden de toetsleiders, als de school dat wenste, vóór de afnamedag al een voorbereidend schoolbezoek af om de afname met de leerkracht(en) door te spreken en/of alvast kennis te maken met de leerlingen. Op andere sbo- en so4-scholen werd deze extra voorbereiding telefonisch afgehandeld. In overleg met de school werd bijvoorbeeld bepaald of er vanwege de speciale doelgroepen in sbo en so4 aanpassingen in de afnameprocedure nodig waren, zoals het inlassen van extra pauzes gedurende de toetsafname, het voorlezen van de leerlingvragenlijst (niet de toetsopgaven) en het opsplitsen van de groep in kleinere groepjes (alleen toegestaan in so4).

Op de afnamedag zelf verliep de afname op de meeste basisscholen en s(b)o-scholen als volgt:

- 1) Toetsleider bespreekt met de leerkracht de klas-opstelling en de lengte en activiteiten tijdens de pauzes. Indien nodig zet de toetsleider de meegenomen Chromebooks klaar.
- 2) Toetsleider introduceert zichzelf en het onderzoek aan de leerlingen.
- 3) Toetsleider deelt kladpapier uit met individuele inlogcodes aan de hand van de eerste leerlingadministratielijst met voornamen. Leerlingen loggen in.
- 4) Toetsleider geeft de leerlingen instructie voor het maken van de toets.
- 5) Afname toets, deel 1 (maximaal 40 minuten) aan de hand van een afnamescript.
- 6) Pauze (duur en activiteit in overleg met de leerkracht).
- 7) Afname toets, deel 2 (maximaal 40 minuten) aan de hand van een afnamescript.
- 8) Pauze (duur en activiteit in overleg met de leerkracht).
- 9) Afname leerlingvragenlijst ($\pm$ 20 minuten) aan de hand van een afnamescript.
- 10) Toetsleider vult de tweede leerlingadministratielijst en het Toetsafnameformulier (TAF) verder in.
- 11) Leerlingen leveren hun kladpapier in bij de toetsleider en krijgen een presentje.

Zowel in het basisonderwijs als s(b)o zijn de afnames goed verlopen. Er waren nauwelijks technische problemen, en de technische problemen die zich voordeden konden de toetsleiders ter plekke oplossen. Voor veel leerlingen bleek de gereserveerde toetstijd van 40 minuten per deel overigens te ruim bemeten; vaak waren alle leerlingen van een klas na circa 30 minuten klaar met een toetsdeel.

5.2.4 Afname leerkracht- en schoolvragenlijst

Bij de start van de afnameperiode ontvingen alle schoolleiders en leerkrachten een e-mail met daarin de link en inlogcode naar de online school- resp. leerkrachtvragenlijst en het verzoek de vragenlijst in te vullen. Gedurende de afnameperiode zijn regelmatig herinneringsmails gestuurd naar schoolleiders en leerkrachten die de vragenlijst nog niet (volledig) hadden ingevuld.

5.3 Respons op schoolniveau

5.3.1 Regulier basisonderwijs

Onderstaand overzicht laat de respons op schoolniveau zien voor het regulier basisonderwijs (groep 8). Er zijn in totaal 106 basisscholen met 162 klassen bereid gevonden om deel te nemen aan het onderzoek. Tabel 5.1 geeft per tranche het aantal scholen in de originele steekproef en de gerealiseerde steekproef. Het totale responspercentage is 69,7% met deelname van 106 scholen uit een steekproef van 152 scholen. Van deze 106 scholen zijn er 52 afkomstig uit de originele steekproef. De respons zonder vervanging bedraagt daarmee 34,2%.

Tabel 5.1 - Verdeling steekproef hoofdonderzoek en responsgroep regulier basisonderwijs naar schoolgewicht, Peil.Rekenen-wiskunde PO 2023

Schoolgewicht	Originele steekproef		Gerealiseerde originele steekproef (1e tranche)		Gerealiseerde steekproef met reservescholen (2e en 3 ^e tranche)	
	N	%	N	%	N	%
1 = hoog schoolgewicht (>36)	8	5,3	2	3,9	6	5,7
2 = gemiddeld/hoog schoolgewicht (30-36)	52	34,2	21	40,4	40	37,7
3 = gemiddeld/laag schoolgewicht (25-30)	70	46,1	23	44,2	50	47,2
4 = laag schoolgewicht (<25)	22	14,5	6	11,5	10	9,4
Totaal	152	100	52	100	106	100

5.3.2 Speciaal basisonderwijs

In Tabel 5.2 wordt de respons op schoolniveau voor het speciaal basisonderwijs weergegeven. In totaal hebben 49 sbo-scholen deelgenomen aan het onderzoek. Er hebben 30 scholen uit de originele steekproef deelgenomen, het responspercentage uit de originele steekproef bedraagt 46,2%. De uiteindelijke respons, inclusief vervanging door scholen uit de reservesteekproeven (tweede en derde tranche) is 75,4%.

Tabel 5.2 – Verdeling populatie en steekproef hoofdonderzoek sbo naar uitstroomprofiel, Peil.Rekenen-wiskunde PO 2023

Percentage uitstroom vso/pro	Originele steekproef		Gerealiseerde originele steekproef (1e tranche)		Gerealiseerde steekproef met reservescholen (2e en 3e tranche)	
	N	%	N	%	N	%
1 = scholen met een gemiddelde uitstroom naar vso/pro van 40% of lager	18	27,6	7	23,3	12	24,5
2 = scholen met een gemiddelde uitstroom naar vso/pro tussen 40% en 60%	27	41,5	16	53,3	23	46,9
3 = scholen met een gemiddelde uitstroom naar vso/pro van 60% of hoger	20	30,8	7	23,3	14	28,6
Totaal	65	100	30	100	49	100

5.3.3 Speciaal onderwijs

Binnen so4 zijn alle scholen (minus de scholen die aan de proefafname hadden deelgenomen) benaderd met schoolverlaters met het uitstroomprofiel vo. Er zijn in totaal 92 scholen benaderd, waarvan er 28 hebben deelgenomen. Het responspercentage bedraagt daarmee 30%.

5.4 Redenen voor niet-deelname

Wanneer een school afzag van deelname aan het onderzoek, werd de school uit de eerste vervangende steekproef (tweede tranche) benaderd via dezelfde procedure. Indien ook deze weigerde werd de school uit de tweede vervangende steekproef (derde tranche) benaderd.

In totaal hebben 299 bao-scholen een uitnodigingsbrief ontvangen. Uiteindelijk hebben 106 bao-scholen daadwerkelijk deelgenomen aan het onderzoek. Daarnaast zijn er 15 scholen die aanvankelijk hadden toegezegd, maar zich later wegens tijdgebrek of onverwachte belemmeringen (zoals ziekte van de leerkracht) weer terugtrokken. In totaal 178 benaderde basisscholen hebben afgezien van deelname, waaronder 156 scholen die met opgave van reden(en) hebben geweigerd. De voornaamste redenen voor hun weigering worden in Tabel 5.3 gepresenteerd. Van 22 benaderde scholen is niet bekend waarom zij niet hebben deelgenomen en/of is nooit een duidelijke toe- of afzegging verkregen.

Tabel 5.3 – Voornaamste reden weigering bao-scholen, Peil.Rekenen-wiskunde PO 2023

Voornaamste reden weigering deelname	Aantal scholen (totaal 156)	%
Te druk/kost te veel tijd/moeite	34	22
Zien het belang niet in	18	12
Willen focus op regulier ‘normaal’ onderwijs	16	10
Deelname andere onderzoeken	15	10
Personeelstekort/ziekteverzuim	13	8
Onrust in school/al genoeg problemen	12	8
Te belastend voor leerlingen/leerlingen kunnen het niet aan	9	6
Doen nooit mee aan onderzoek	9	6
Grote veranderingen (zoals verhuizing)	7	4
Prioriteiten elders	4	3
Geen duidelijke reden	19	12

In totaal hebben 124 sbo-scholen een uitnodigingsbrief ontvangen. Aanvankelijk hebben 54 van deze scholen toegezegd deel te nemen aan het onderzoek. Uiteindelijk hebben er 49 sbo-scholen daadwerkelijk deelgenomen, 5 scholen hebben zich op het laatste moment teruggetrokken. De overige 70 benaderde sbo-scholen hebben aangegeven niet mee te willen werken aan het onderzoek. In Tabel 5.4 zien we wat de voornaamste reden voor deze weigering was.

Tabel 5.4 – Voornaamste reden weigering sbo-scholen, Peil.Rekenen-wiskunde PO 2023

Voornaamste reden weigering	Aantal scholen (totaal 70)	%
Te belastend voor leerlingen/leerlingen kunnen het niet aan	32	46
Personeelstekort/ziekteverzuim	8	11
Te druk/kost te veel tijd/moeite	6	9
Onrust in school/al genoeg problemen	5	7
Grote veranderingen (zoals verhuizing)	3	4
Prioriteiten elders	3	4
Deelname andere onderzoeken	3	4
Willen focus op regulier 'normaal' onderwijs	1	1
Zien het belang niet in	1	1
Geen duidelijke reden	8	11

De uitnodigingsbrief voor so4 werd verstuurd naar 92 so4-scholen. In totaal hebben er 31 scholen deelname toegezegd, waarvan er 28 daadwerkelijk hebben deelgenomen; 3 scholen hebben zich na aanvankelijke toezegging op het laatste moment weer teruggetrokken. Drie so4-scholen bleken onterecht in de steekproef te zijn opgenomen (niet de juiste leerlingpopulatie). De resterende 58 so-scholen hebben aangegeven niet deel te willen nemen. In Tabel 5.5 wordt de voornaamste reden voor deze weigering weergegeven.

Tabel 5.5 – Voornaamste reden weigering so4-scholen, Peil.Rekenen-wiskunde PO 2023

Voornaamste reden weigering	Aantal scholen (totaal 58)	%
Te belastend voor leerlingen/leerlingen kunnen het niet aan	13	22
Te druk/kost te veel tijd/moeite	12	21
Personeelstekort/ziekteverzuim	8	14
Deelname andere onderzoeken	7	12
Grote veranderingen (zoals verhuizing)	4	7
Onrust in school/al genoeg problemen	3	5
Prioriteiten elders	3	5
Willen focus op regulier 'normaal' onderwijs	3	5
Zien het belang niet in	1	2
Geen duidelijke reden	4	7

5.5 Respons op schoolleider-, leerkracht- en leerlingniveau

De volgende tabellen geven per onderwijssoort een overzicht van de respons op schoolleider-, leerkracht- en leerlingniveau.

Tabel 5.6 – Respons schoolleiders, leerkrachten en leerlingen in bao, Peil.Rekenen-wiskunde PO 2023

Scholen/schoolleiders	Aantal
Aantal deelnemende scholen	106
Benaderde schoolleiders*	98
Schoolvragenlijst ingevuld**	57
Respons schoolvragenlijst	58%
Leerlingen	
Toets gemaakt	3382
Leerlingvragenlijst ingevuld**	3241
Respons leerlingvragenlijst (o.b.v. toetsdeelname)	96%
Leerkrachten/klassen	
Aantal klassen	162
Leerkrachtvragenlijst ingevuld**	104
Respons leerkrachtvragenlijst	64%

* Voor scholen die alleen met groep 8 hebben deelgenomen (n=8) was het niet mogelijk om de TIMSS-schoolvragenlijst voor te leggen; ** Respondent heeft 50% of meer van de vragenlijst ingevuld.

Tabel 5.7 – Respons schoolleiders, leerkrachten en leerlingen in sbo, Peil.Rekenen-wiskunde PO 2023

Scholen/schoolleiders	Aantal
Aantal deelnemende scholen/benaderde schoolleiders	49
Schoolvragenlijst ingevuld*	31
Respons schoolvragenlijst	63%
Leerlingen	
Toets gemaakt	1047
Leerlingvragenlijst ingevuld*	983
Respons leerlingvragenlijst (o.b.v. toetsdeelname)	94%
Leerkrachten/klassen	
Aantal klassen	97
Leerkrachtvragenlijst ingevuld*	56
Respons leerkrachtvragenlijst	58%

*Respondent heeft 50% of meer van de vragenlijst ingevuld

Tabel 5.8 – Respons schoolleiders, leerkrachten en leerlingen in so4, Peil.Rekenen-wiskunde PO 2023

Scholen/schoolleiders	Aantal
Aantal deelnemende scholen/benaderde schoolleiders	28
Schoolvragenlijst ingevuld*	20
Respons schoolvragenlijst	71%
Leerlingen	
Toets gemaakt	525
Leerlingvragenlijst ingevuld*	500
Respons leerlingvragenlijst (o.b.v. toetsdeelname)	95%
Leerkrachten/klassen	
Aantal klassen	64
Leerkrachtvragenlijst ingevuld*	42
Respons leerkrachtvragenlijst	66%

*Respondent heeft 50% of meer van de vragenlijst ingevuld

5.6 Representativiteitsanalyse

Omdat er niet volledig aan de in de call gestelde responseisen kon worden voldaan, is er een representativiteitsanalyse uitgevoerd. Voor bao, sbo en so4 is onderzocht in hoeverre er verschillen zijn tussen de populatie scholen en de responsgroep in schoolkenmerken die een rol zouden kunnen spelen in de uiteindelijke toetsresultaten. Deze schoolkenmerken zijn:

- gemiddeld percentage leerlingen dat niveau 2F (rekenen en taal) niet behaalt, sinds 2015;
- achterstandsscore;
- stedelijkheid van de vestigingsgemeente;
- provincie;
- denominatie;
- geografische spreiding.

Samengevat blijkt dat er geen aanleiding is om te veronderstellen dat de non-respons heeft geresulteerd in een selectieve vertekening in de uiteindelijke (toets)resultaten doordat een relevante groep scholen is onder- of oververtegenwoordigd. Dit geldt zowel voor bao, sbo als so4. Alleen de (gemiddelde) schoolgrootte is in de responsgroep bao wel groter dan in de populatie. Kleine basisscholen zijn ondervertegenwoordigd in responsgroep. Een mogelijke verklaring hiervoor is dat scholen waarvan verwacht werd dat zij in 2023 minder dan 6 leerlingen in groep 6 zouden hebben, vooraf al zijn uitgesloten van de TIMSS-steekproef. Het is de vraag is of de ondervertegenwoordiging van kleine scholen een probleem is op leerlingniveau. Er zijn minder grote scholen dan kleinere scholen, waardoor leerlingen op grote scholen minder kans hebben om deel te nemen aan het onderzoek dan leerlingen op kleinere scholen. De oververtegenwoordiging van grotere scholen kan er dus toe leiden dat de insluitkansen op leerlingniveau juist gelijk zijn geworden. De uitgebreide beschrijving van de representativiteitsanalyse is te vinden in bijlage 1 van dit rapport.

6. Schaling Rekenen-wiskunde

6.1 Inleiding

Het doel van de peiling is om de vaardigheid van de leerlingen op het gebied van Rekenen-wiskunde in het laatste leerjaar van het basisonderwijs in kaart te brengen (zie hoofdstuk 7) en om de relatie tussen deze vaardigheid en achtergrondkenmerken van de leerlingen en vakleerkrachten te onderzoeken (zie ook hoofdstuk 9). Om zeker te zijn dat het domein Rekenen-wiskunde breed wordt afgedekt is een groot aantal items afgenomen. In het eerste deel van dit hoofdstuk bespreken we het design van de afnames en het functioneren van de items.

Een gevolg van het grote aantal items is dat het onmogelijk is om bij elke leerling in de steekproef alle items af te nemen. Dit leidt tot een missende waarde probleem, we weten namelijk niet hoe de leerlingen in de steekproef geantwoord zouden hebben op de items die zij niet voorgelegd hebben gekregen. Om dit probleem op te lossen wordt gebruik gemaakt van Itemrespons Theorie (IRT). In het tweede deel van dit hoofdstuk wordt het proces van het schatten van de IRT item parameters besproken, met speciale aandacht voor de gemaakte keuzes en eventuele gevolgen daarvan voor de uitkomsten in hoofdstuk 7 en 9. Hierbij bespreken we specifiek de koppeling van de data uit de peiling 2023 met die uit de peiling in 2019. Hierbij hebben we te maken met een verschil in afnamemodus. De peiling in 2019 is op papier afgenomen en de peiling in 2023 is digitaal.

In het laatste deel van het hoofdstuk bespreken we de bepaling van de prestatiestandaarden. Deze zijn reeds in de peiling in 2019 (Buisman e.a., 2021) bepaald en worden in dit onderzoek overgezet naar de nieuwe resultaten.

6.2 De toets en afname

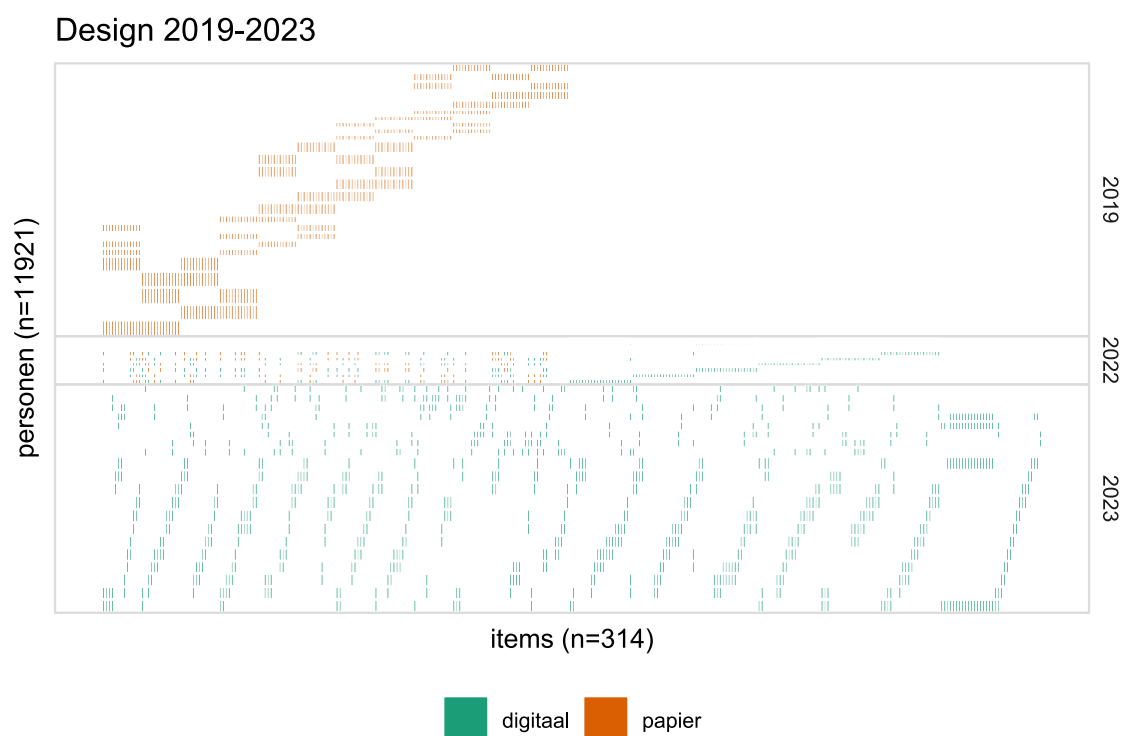
In Peil.Rekenen-wiskunde PO 2023 zijn 304 unieke items afgenomen bij in totaal 4.954 personen. Voor de kalibratie hebben we echter ook te maken met data uit de peiling Rekenen-wiskunde 2019 en daarnaast is er in 2022 een proeftoets afgenomen. Het doel van de proeftoets in 2022 was tweeledig. Het eerste doel was om nieuw ontwikkelde items voor de peiling van 2023 te evalueren, waaronder items voor het nieuw ontwikkelde domein *probleemoplossen*. Het tweede doel was om te onderzoeken wat het effect van de afnamemodus is op de relatieve moeilijkheid van de items. Voor dit doel heeft een aantal leerlingen een gedeelte van de toets op papier gemaakt en een gedeelte digitaal, een zogenoemd *binnen-personen* design. Dit onderzoek is alleen uitgevoerd bij leerlingen in regulier basisonderwijs.

Uit de data van het modusonderzoek zijn 3 leerlingen verwijderd waarbij de afname verstoord werd door computerproblemen en 6 leerlingen die in één van beide afnamemodi niets hadden ingevuld. De overgebleven aantallen leerlingen en scholen zijn gegeven in Tabel 6.1.

Tabel 6.1 – Aantallen scholen en leerlingen in de peilingsonderzoeken, Peil.Rekenen-wiskunde PO 2023

Jaar	Schooltype	Aantal unieke items	Aantal scholen	Aantal leerlingen
2019	bao	156	164	5.040
	sbo	156	43	879
2022	bao	244	30	731
	sbo	124	16	317
2023	bao	268	106	3.382
	so4	148	28	525
	sbo	148	49	1.047

De combinatie van de afnames in 2019, 2022 en 2023 levert een ingewikkeld afnamedesign op. Figuur 6.1 geeft een grafische weergave van het design en de overlap tussen boekjes.



Figuur 6.1 – Design van de peilingsonderzoeken, Peil.Rekenen-wiskunde PO 2023. In de figuur zijn de personen weergegeven op de y-as, gegroepeerd per boekje. De x-as geeft de items aan.

Het aantal afgenomen items in 2023 is bijna twee maal zo groot als in 2019. Daar staat tegenover dat in 2019 het aantal observaties per item groter was. In 2022 en 2023 zijn er per persoon beduidend meer items afgenomen dan in 2019 (Tabel 6.2).

Tabel 6.2 – Aantallen items en gemiddeld aantal observaties in de drie (peilings)onderzoeken, Peil.Rekenen-wiskunde PO 2023*

Jaar	totaal items	afgenomen in jaren					gem. obs. per item	gem. obs. per persoon
		2019	2019-2022	2022	2022-2023	2023		
2019	156	109	47				986	26
2022	244		47	22	175		236	55
2023	304				175	129	682	42

* Papieren en digitaal afgenomen items zijn in deze tabel als afzonderlijke items opgenomen

6.2.1 Klassieke analyse

We hebben ervoor gekozen om personen die minder dan 50% van de items hebben beantwoord te verwijderen uit de analyse; zowel uit de kalibratie als uit het vervolg van het onderzoek. Het percentage ontbrekende antwoorden is bij de digitale afname meer dan twee maal zo groot als bij de papieren afname. De arbitraire keuze voor een percentage ontbrekende antwoorden als grens voor het verwijderen van leerlingen heeft daardoor een verschillend effect op de resultaten in de verschillende jaren (Tabel 6.3). Het lijkt echter aannemelijk dat leerlingen die meer dan 50% van de vragen overslaan, over het algemeen niet geprobeerd hebben de toets zo goed mogelijk te maken.

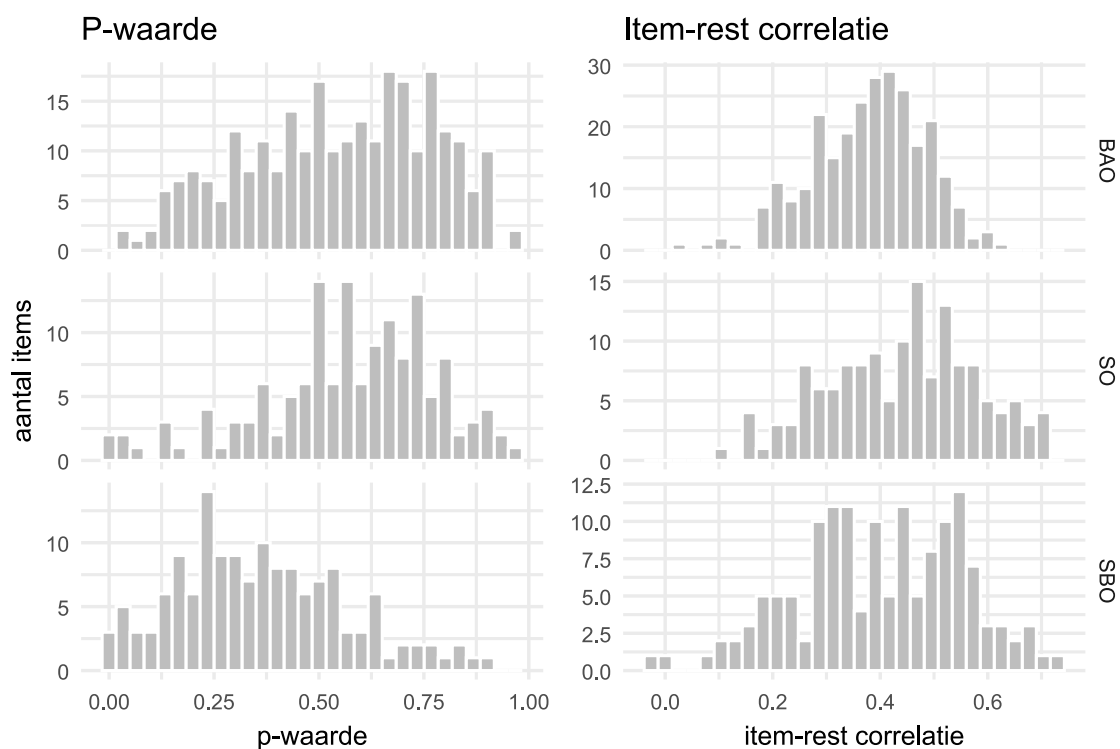
Tabel 6.3 – Percentages missende antwoorden, Peil.Rekenen-wiskunde PO 2023*

Jaar	Aantal personen	Gem. aantal items pp.	Gem. % missing	Personen $\geq 50\%$ missing	Gem. % missing na uitsluiting
2019	5.919	26	3.6%	45 (0.8%)	3.1%
2022	1.048	55	8.4%	35 (3.3%)	6.3%
2023	4.954	42	9.0%	143 (2.9%)	7.3%

* De kolommen geven per jaar respectievelijk: aantal personen, gemiddelde aantal items per persoon, gemiddelde percentage missende antwoorden, aandeel personen met meer dan 50% missende antwoorden en het percentage missende antwoorden na uitsluiting van personen met meer dan 50% missende antwoorden.

Na inspectie van de items, antwoorden en aanvulling van sleutels is besloten om 13 items te verwijderen. Deze items hadden allemaal een item-rest correlatie kleiner dan 0.10 en hoewel er bij de meeste vragen geen onduidelijkheid in formulering of ander gebrek kon worden gevonden, leken de antwoorden voornamelijk gebaseerd op gokken. Er zijn geen items verwijderd op basis van p-waardes.

De gemiddelde p-waarde van de items die zijn afgenomen in 2023 ligt op 0.56. Dit duidt op een vrij goede passing van de itemmoeilijkheid bij de populatie. De gemiddelde item-rest correlatie is 0.37 wat zeker als een acceptabele waarde kan worden gezien. In het sbo zijn de p-waardes lager, met een gemiddelde p-waarde van 0.37. Dit geeft aan dat, ondanks dat sbo-leerlingen aparte toetsversies hebben gekregen, de toets relatief moeilijk was voor sbo-leerlingen. De verdeling van item-rest correlaties en p-waardes is grafisch weergegeven in Figuur 6.2.



Figuur 6.2 – Overzicht item-rest correlaties en p-waardes items, Peil.Rekenen-wiskunde PO 2023

6.3 Schaling

De peiling in 2019 is op papier afgenomen en de peiling in 2023 digitaal. In 2022 vond een modusstudie plaats waarbij leerlingen zowel papieren als digitale items voorgelegd kregen. De digitale items betroffen grotendeels dezelfde vragen als de papieren items, waarbij de laatste met zo min mogelijk aanpassingen geschikt zijn gemaakt voor digitale afname. Om de resultaten uit 2019 te kunnen vergelijken met die uit 2023 moet worden aangetoond dat de items in de verschillende jaren en afnamemodi ongeveer gelijk functioneren. We onderzoeken daarom of er aanwijzingen zijn dat items anders functioneren tussen de verschillende jaren en afnamemodi, zogenaamd Differential Item Functioning (DIF).

Voor dit onderzoek gebruiken we de toets voor Differential Item Pair Functioning (Bechger en Maris 2014). Hierbij wordt het Rasch model apart geschat in twee groepen. Vervolgens wordt het verschil in de relatieve moeilijkheid van alle paren van items vergeleken tussen groepen. De reden om een paarsgewijze vergelijking te gebruiken is dat de moeilijkheid van een item niet is geïdentificeerd op grond van de observaties, maar de verschillen in moeilijkheden (relatieve moeilijkheden) wel. De methode bestaat uit twee delen. Ze begint met een algehele statistische toets waarbij de twee matrices van gestandaardiseerde verschillen in relatieve moeilijkheden worden vergeleken. De gekwadraterde Mahalanobis-afstand volgt een Chi-kwadraat verdeling met het aantal items minus 1 vrijheidsgraden. Deze toets geeft vrijwel altijd een significant resultaat wanneer de groepen worden gebaseerd op een wezenlijk verschil in achtergrondvariabelen. De statistische toets op zichzelf is daarom minder interessant. We zijn daarentegen geïnteresseerd of, gegeven dat er vrijwel altijd DIF bestaat, het gekozen psychometrische model desondanks bruikbaar is voor het doel waarvoor we het willen gebruiken. Hierbij helpt het om te kijken naar het patroon van de afwijkingen tussen

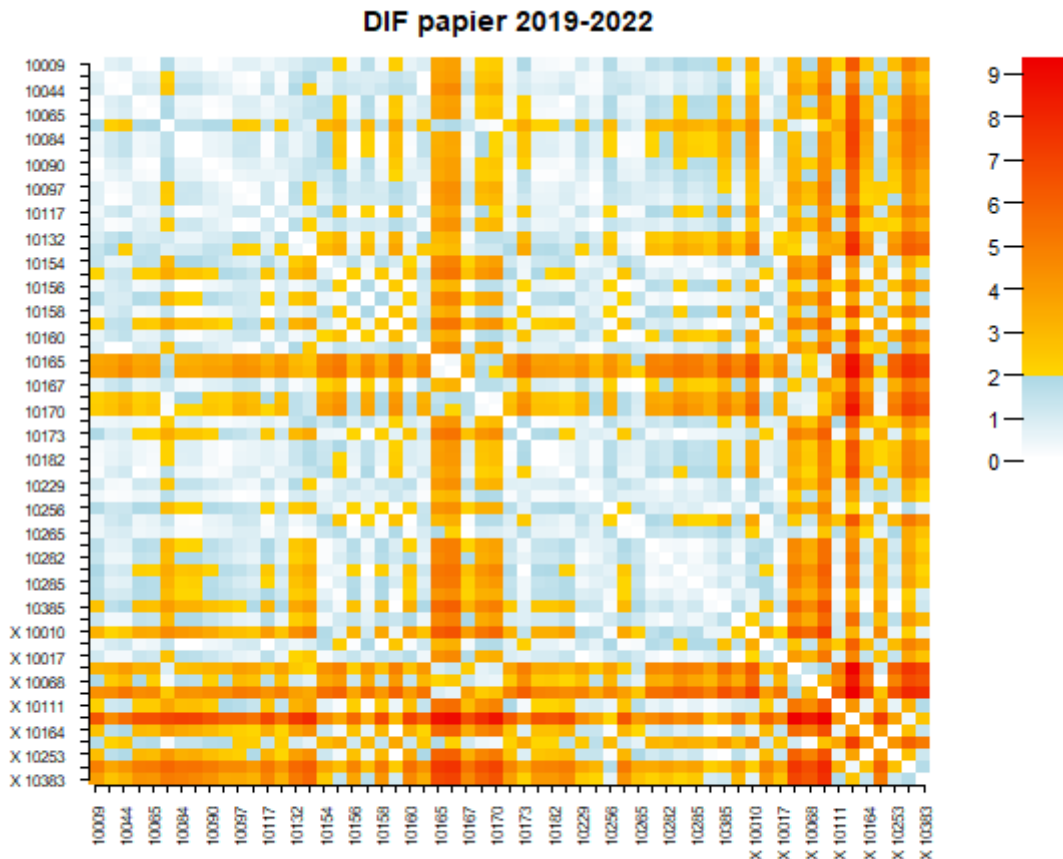
itemparen. Aan de hand van dit patroon kunnen soms problematische items worden geïdentificeerd of groepen items herkend waarbinnen zeer weinig DIF optreedt.

We geven als resultaat van de DIF toets hierna telkens alleen de grafische weergave van de verschillen tussen itemparen (figuren Figuur 6.3, Figuur 6.8, Figuur 6.9, Figuur 6.10). Deze weergave is gebaseerd op absolute gestandaardiseerde verschillen. De kleurovergang tussen wit/blauw en geel/rood vindt plaats bij een absolute z-score van 1.96, oftewel een significantieniveau van 5%. De lengte van de kleurschaal verschilt echter per grafiek. In alle gevallen waar een grafiek wordt weergegeven is de overschrijdingskans van de algehele toets voor DIF tussen de groepen op basis van de Mahalanobis-afstand kleiner dan 0.001.

6.3.1 DIF tussen 2019 en 2022

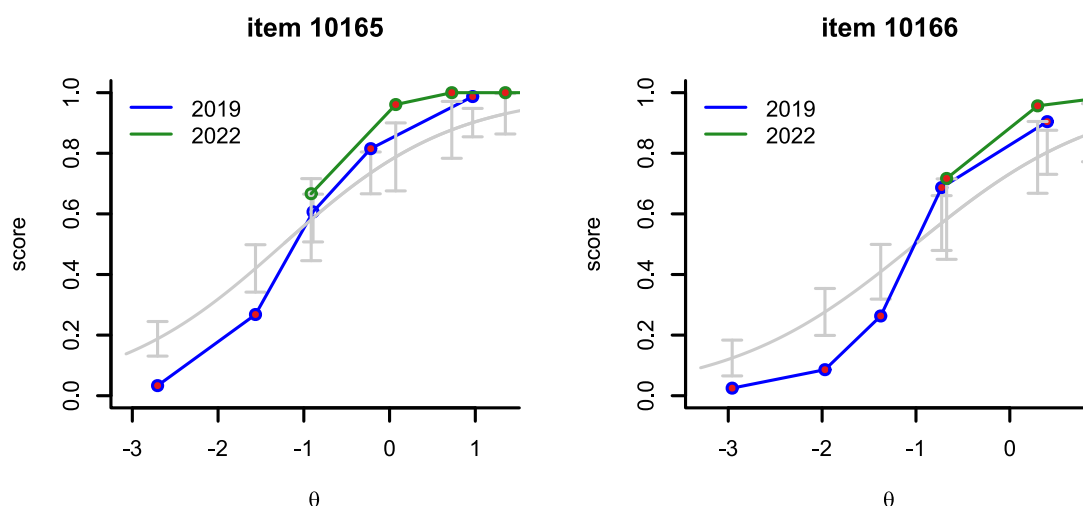
De proeftoets in 2022 bestond uit zowel papieren als digitale items. Als we alleen de papieren afname van het modusonderzoek beschouwen, is het design van de proeftoets niet verbonden. Om het design te verbinden is het noodzakelijk om ook het digitale deel van de afname mee te nemen. We beschouwen hierbij de digitaal afgenomen items als andere items dan de op papier afgenomen items, zodat we mogelijke effecten tussen jaren niet verwarren met effecten van afnamemodus.

Er blijkt in eerste instantie sprake van vrij veel DIF tussen de papieren afnames in 2019 en 2022 (zie Figuur 6.3). Na zorgvuldige inspectie van de items in alle papieren boekjes van de peiling in 2019 en de proeftoets van 2022 blijkt echter dat er voor 18 items inhoudelijke - of presentatieverschillen zijn. Een voorbeeld van een inhoudelijk verschil dat bij een aantal items werd gevonden was een decimaalteken dat wél of niet al was opgenomen in het antwoordveld. Daarnaast waren de referentielijnen in een afbeelding bij één item weggefallen en waren er enkele items waarbij de vraagstelling was aangepast. We hebben besloten om deze items apart te schalen, dat wil zeggen: als verschillende items te beschouwen in de peiling 2019 en de proeftoets in 2022. Het grootste deel van de DIF verdwijnt hierdoor. Hierdoor is het uiteraard zo dat de link tussen 2019 en 2022 kleiner wordt, namelijk 46 in plaats van 59 items.



Figuur 6.3 – DIF toets voor itemparen, Peil.Rekenen-wiskunde PO 2023. Vergelijking papieren afname 2019 – papieren afname 2022. Items waarvoor inhoudelijke- of presentatieverschillen zijn gevonden, zijn gemarkeerd met een X en onderaan in de figuur gezet.

Ook twee items waarvoor we geen inhoudelijke verschillen hebben gevonden, lijken in eerste instantie ook anders te functioneren in de twee groepen. Inspectie van de itemkarakteristieke curves (Figuur 6.4) laat echter zien dat dit een combinatie is van een kleine model misfit en een verschil in vaardigheid tussen de personen in de twee groepen waarbij deze items zijn afgenomen. Deze twee items functioneren bij nader inzien dus bij benadering hetzelfde in beide groepen.

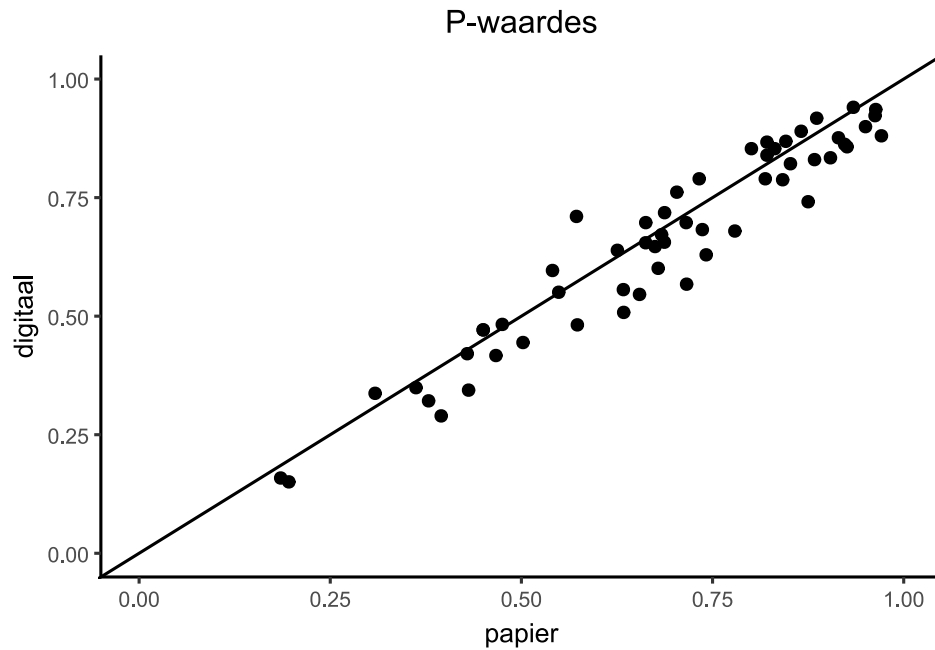


Figuur 6.4 – Twee items met ogenschijnlijke DIF tussen de 2019- en 2022-afname, Peil.Rekenen-wiskunde PO 2023. De blauwe en groene lijnen volgen de gemiddelde scores per vaardigheidsgroep in 2019 en 2022. De grijze lijn geeft de verwachte scores met het bijbehorende 95% betrouwbaarheidsinterval.

6.3.2 Modusonderzoek

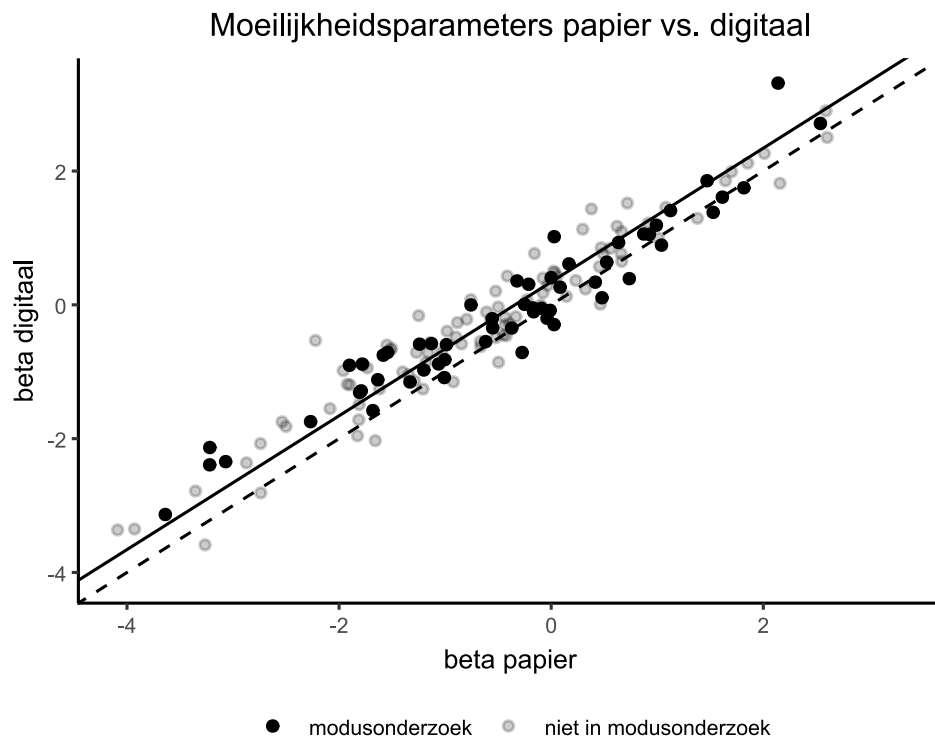
In 2022 zijn 56 items in beide afnamemodi afgenomen bij 729 leerlingen. Deze leerlingen beantwoordden elk ongeveer 20 van deze ankeritems op papier, 20 andere ankeritems digitaal en nog eens 20 nieuwe items digitaal. De latente correlatie tussen Rekenen-wiskunde papier en digitaal is 0.93. Daarom lijkt het in ieder geval gerechtvaardigd om Rekenen-wiskunde in een digitale en een papieren afname als dezelfde vaardigheid te beschouwen.

Een vergelijking van de proportie correct (p-waarde) van de papieren en digitale items in het modusonderzoek laat zien dat de moeilijkheid van de vragen tussen de afnamemodi structureel iets verschilt (zie Figuur 6.5). De gemiddelde p-waarde van de digitaal afgenomen items is 0.66 tegenover 0.68 voor de papieren afname, een verschil van 0.027. Dit geeft aan dat items in een digitale afnamemodus minder vaak goed worden beantwoord dan op papier.



Figuur 6.5 – P-waardes voor de items in het modusonderzoek, Peil.Rekenen-wiskunde PO 2023. Op papier worden 36 van de 56 items (64.3%) beter gemaakt dan digitaal. De p-waardes zijn alleen gebaseerd op de leerlingen die zowel papieren als digitale items hebben gemaakt.

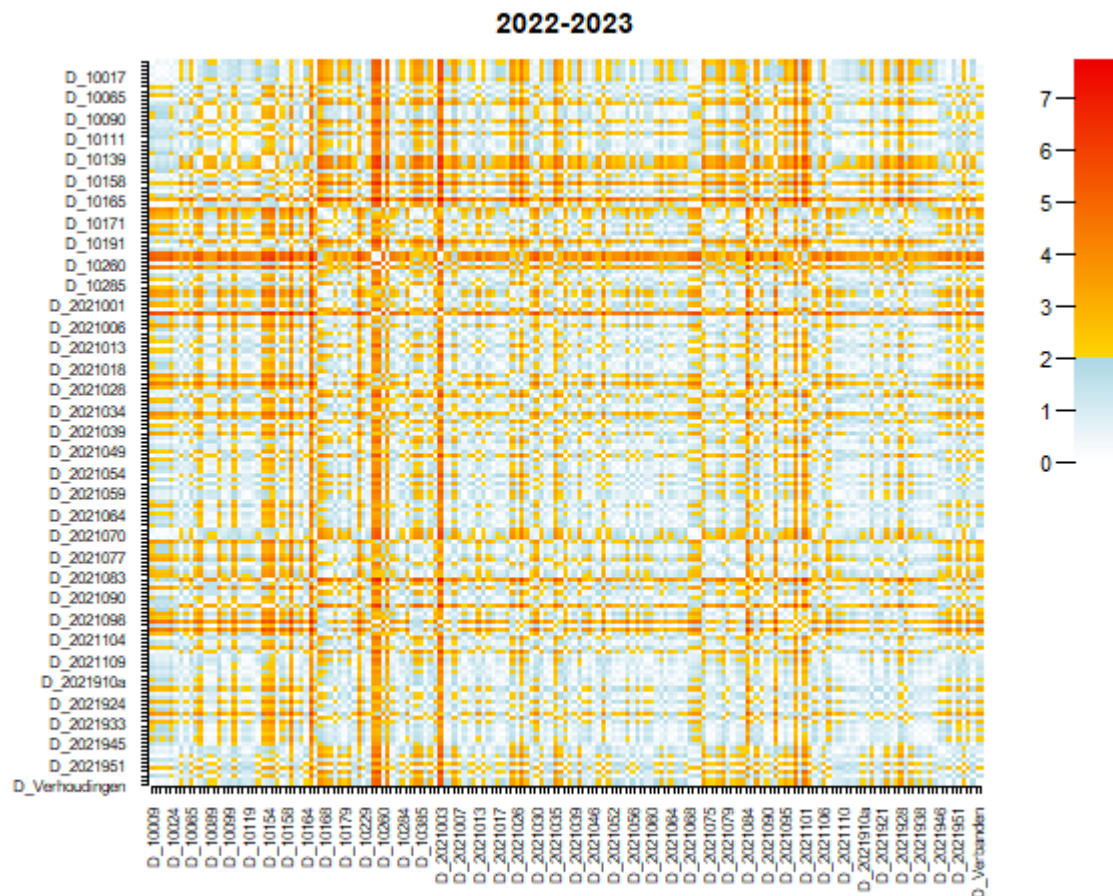
Het *binnen-personen design* maakt het mogelijk om een IRT model te schatten waarbij we de op papier en digitaal afgenomen items als verschillende items beschouwen. Wanneer we de aldus opnieuw geschaalde moeilijkheidsparameters van “dezelfde” items in de twee afnamemodi met elkaar vergelijken zien we dat de digitale afname zorgt voor een gemiddeld hogere moeilijkheid van de items (Figuur 6.6). Dit is in overeenstemming met Figuur 6.5



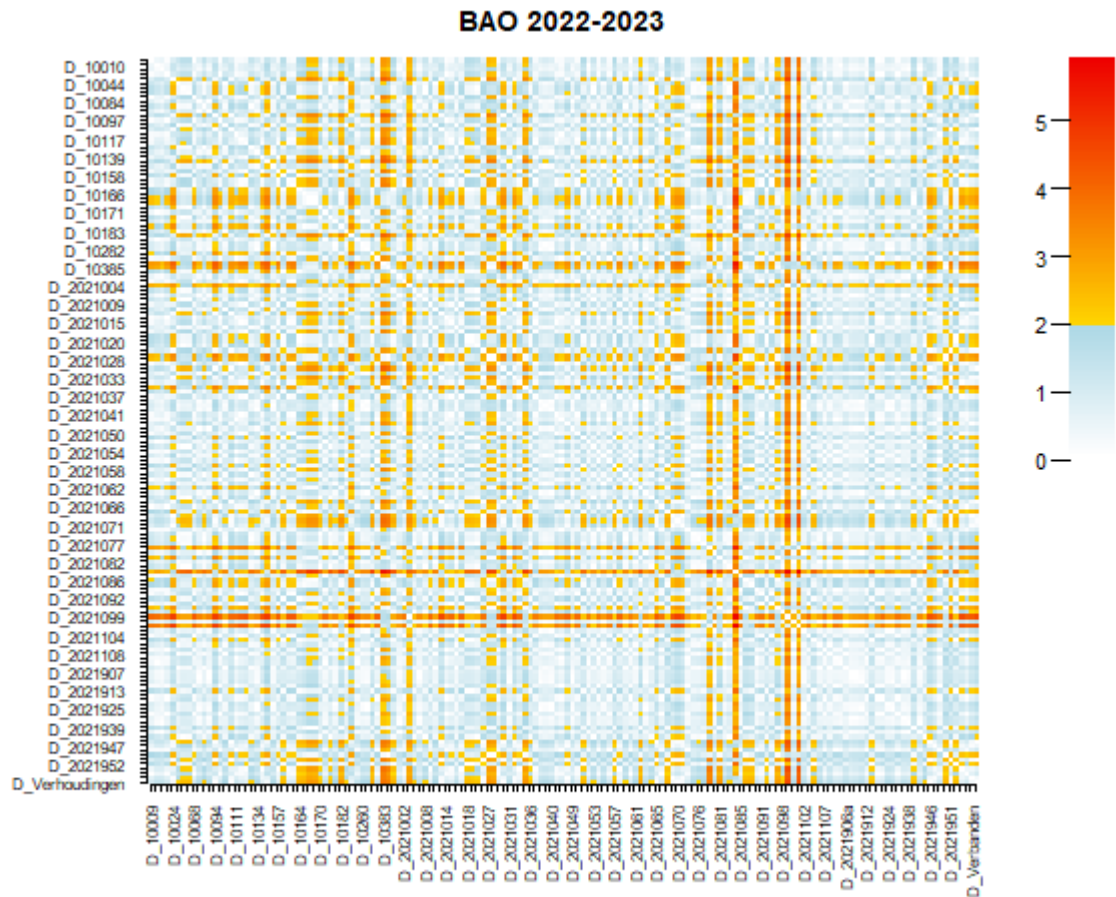
Figuur 6.6 – Moeilijkheidsparameters apart geschat voor papieren en digitale items o.b.v. het binnenpersonen design voor items die zowel op papier als digitaal zijn afgenomen, Peil.Rekenen-wiskunde PO 2023. De onderbroken lijn geeft een gelijke moeilijkheid voor papieren en digitale items aan. De doorgetrokken lijn geeft de gemiddelde verschuiving in moeilijkheid aan als resultaat van de kalibratie.

6.3.3 DIF tussen de proeftoets 2022 en de peiling in 2023

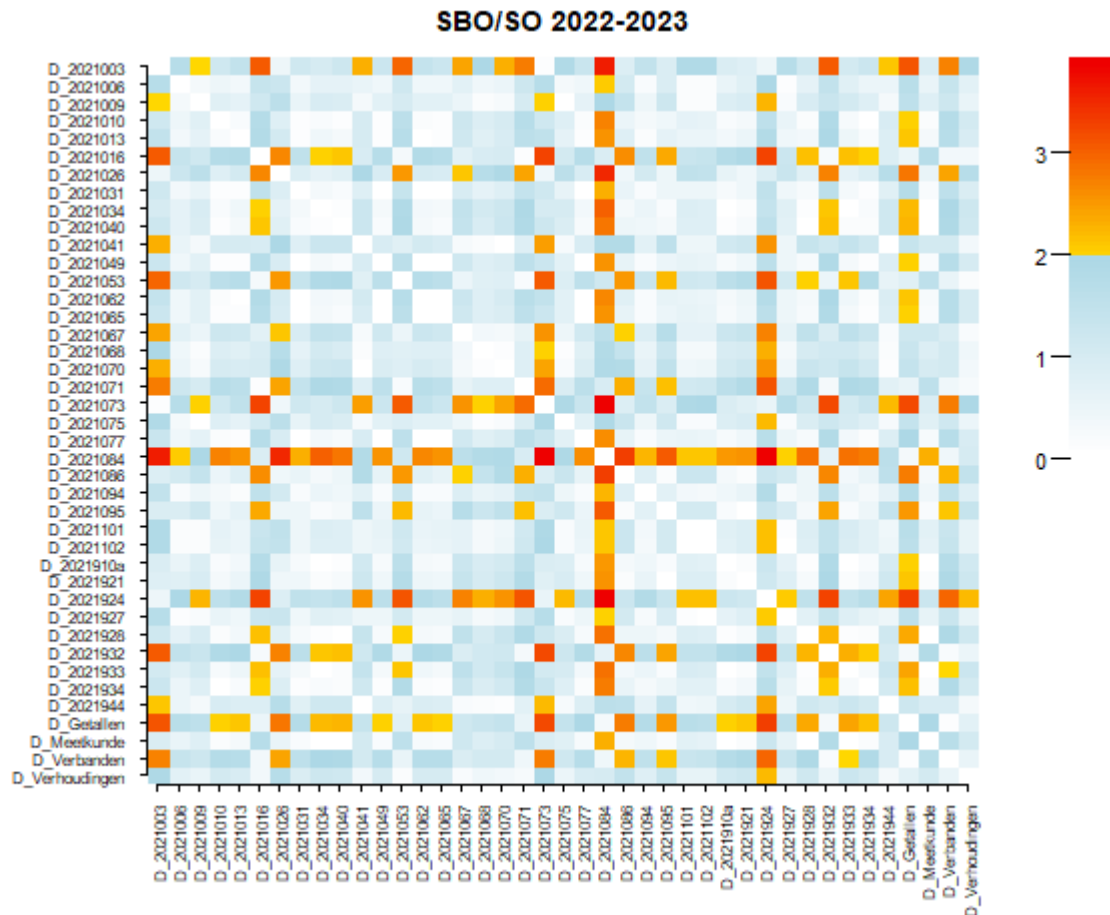
Ook tussen de digitale afnames in 2022 en 2023 is sprake van enige DIF (Figuur 6.7). Hier zijn echter geen inhoudelijke verschillen in de items tussen de afnames gevonden. De DIF blijkt gedeeltelijk een effect te zijn van schooltype. Als we DIF tussen 2022 en 2023 afzonderlijk bepalen voor bao (Figuur 6.8) en voor sbo/so4 (Figuur 6.9) dan valt op dat deze voor een belangrijk deel verdwijnt.



Figuur 6.7 – DIF tussen de digitale afnames van de proeftoets in 2022 en de peiling in 2023 (alle schooltypes), Peil.Rekenen-wiskunde PO 2023.



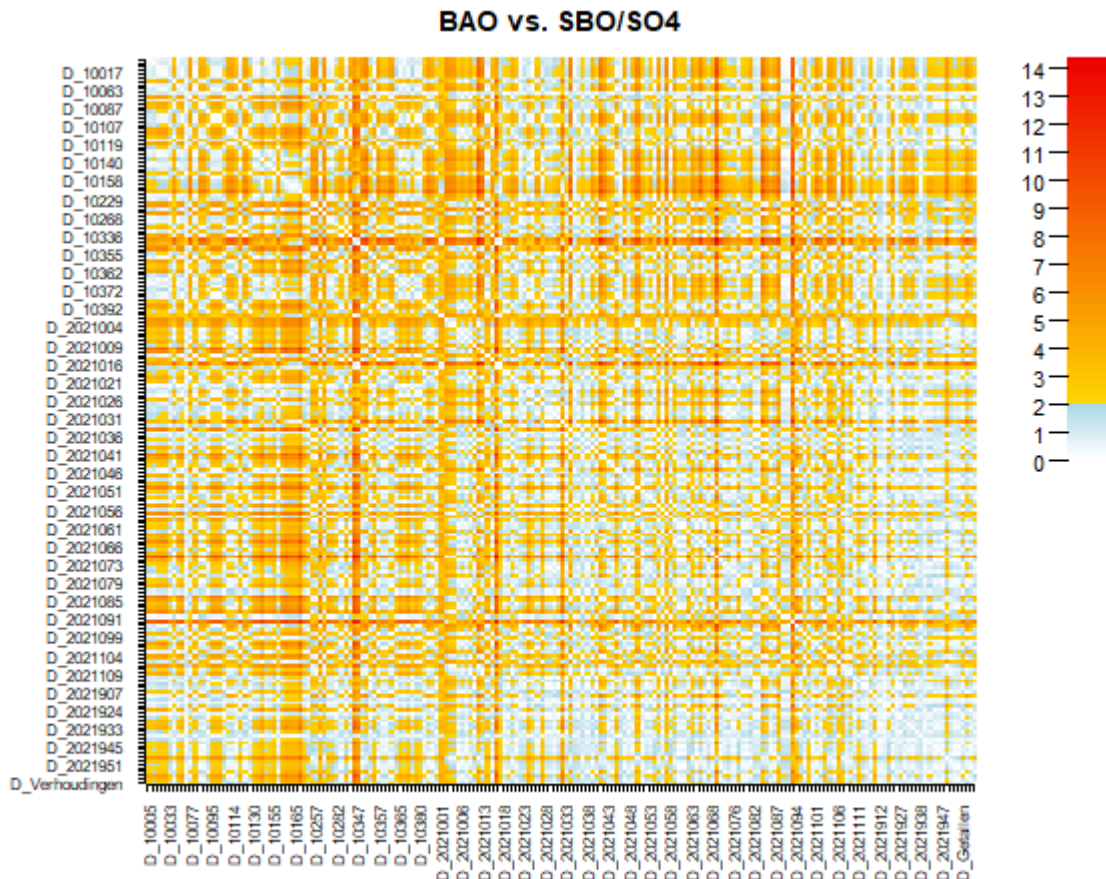
Figuur 6.8 – DIF tussen de digitale afnames van de proeftoets in 2022 en de peiling in 2023 (alleen bao), Peil.Rekenen-wiskunde PO 2023.



Figuur 6.9 – DIF tussen de digitale afnames van de proeftoets in 2022 en de peiling in 2023 in sbo/so4, Peil.Rekenen-wiskunde PO 2023.

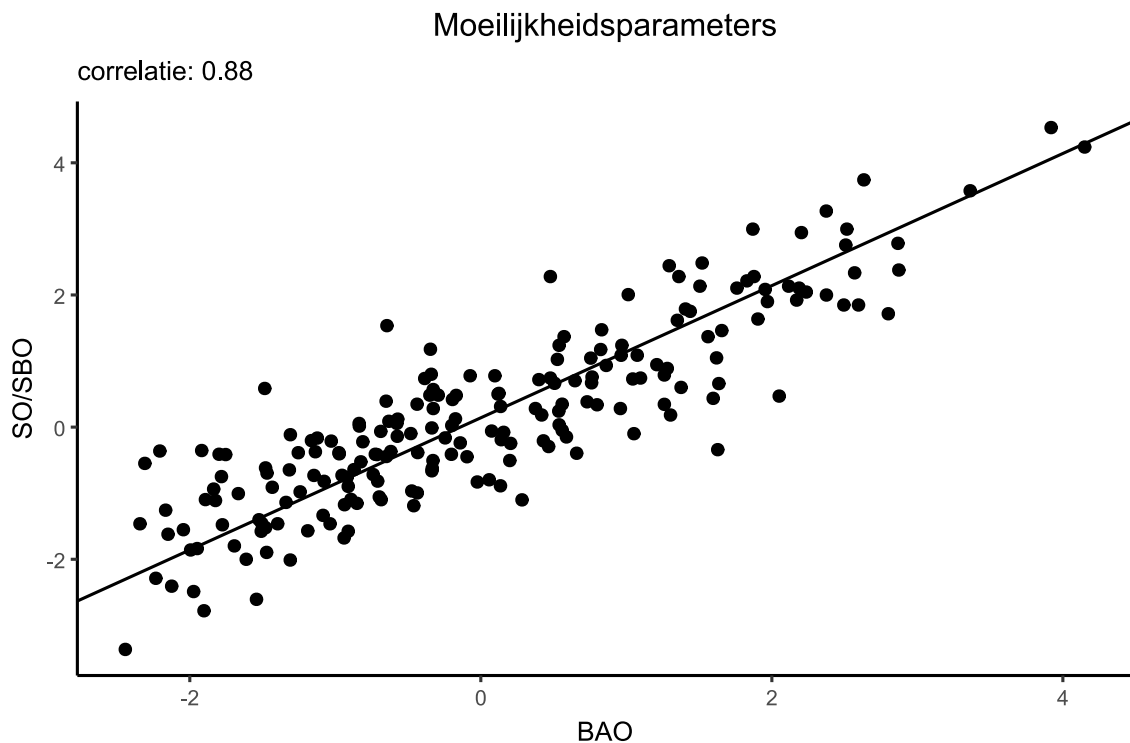
6.3.4 DIF naar schooltype

Een logisch vervolg van de voorgaande analyse is om DIF tussen schooltypes te onderzoeken. Dit doen we alleen op basis van de digitale afnames van 2022 en 2023, omdat in 2019 de afnames in het sbo op zichzelf geen verbonden design vormen. Het blijkt dat de DIF tussen bao-scholen en sbo/so4-scholen inderdaad aanzienlijk is (Figuur 6.10).



Figuur 6.10 – DIF naar schooltype in de digitale afnames, Peil.Rekenen-wiskunde PO 2023. De kleurschaal loopt tot een gestandaardiseerd verschil van 14, wat betekent dat ook de meeste geelgekleurde DIF-paren een vrij grote afwijking vertonen.

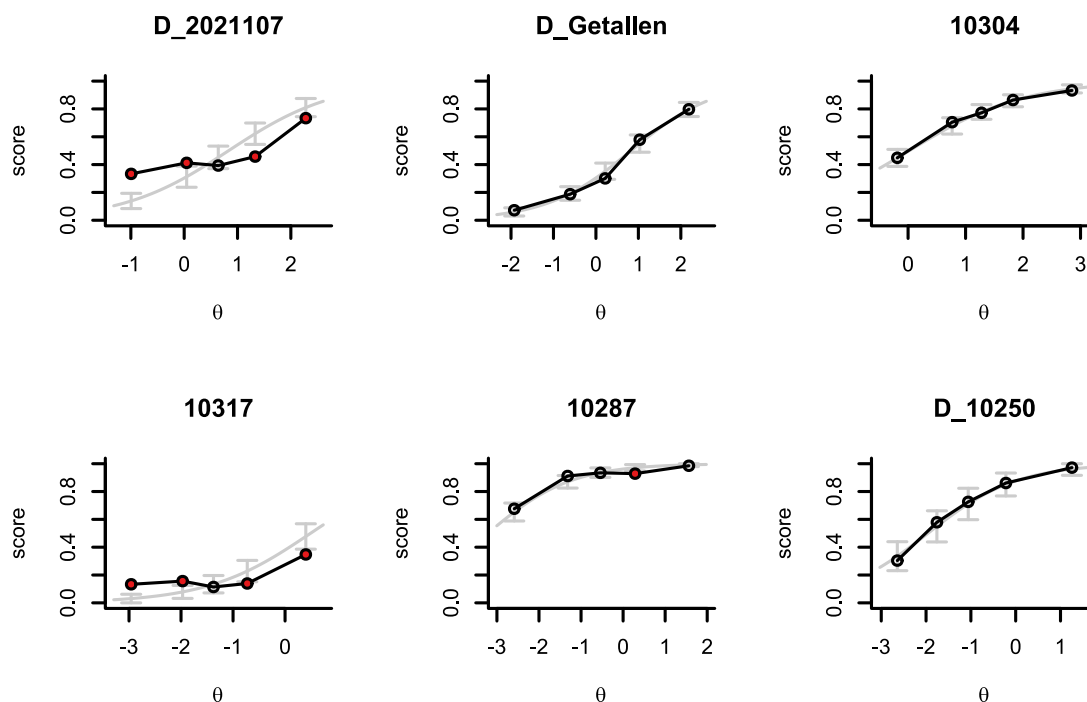
Men kan zich afvragen of het gerechtvaardigd is om bao en sbo/so4 op dezelfde schaal te kalibreren. De correlatie tussen de moeilijkheidsparameters wanneer we separaat binnen het sbo en bao kalibreren is 0.88. Zoals in Figuur 6.11 is te zien, is er enig verschil in (de ordening van) de moeilijkheid van items tussen deze schooltypes. We besluiten echter om deze DIF te negeren en er van uit te gaan dat items binnen de afzonderlijke populaties gelijk functioneren. De overweging hierbij is voornamelijk praktisch: sbo-leerlingen hebben in het proefonderzoek geen digitale items gemaakt, dus wanneer we alleen de so4- en sbo-leerlingen selecteren is het design tussen 2019 en 2023 niet verbonden zodat er geen vergelijking gemaakt kan worden tussen jaren.



Figuur 6.11 – Item moeilijkheidsparameters apart geschat in bao en sbo/so4, Peil.Rekenen-wiskunde PO 2023. De diagonale lijn is ter referentie. Deze heeft richtingscoëfficiënt 1.

6.3.5 Model en itempassing

Gegeven de moduseffecten, DIF tussen populaties en het verschil in vaardigheid in de populaties is het verstandig om te kalibreren met Conditional Maximum Likelihood (CML), aangezien hierbij zeer minimale aannames gedaan hoeven worden in vergelijking met andere schattingsmethoden, zoals het eveneens veel gebruikte Marginal Maximum Likelihood (MML). Ondanks de hierboven beschreven DIF is de passing van de items op het Rasch model overwegend redelijk tot goed. De itemkarakteristieke curves voor een zestal items in Figuur 6.12 geven een indruk van de passing van de items.



Figuur 6.12 – Visuele indicatie van itempassing, Peil.Rekenen-wiskunde PO 2023. De x-as geeft de vaardigheid, de y-as de gemiddelde score (p-waarde). De twee slechtst passende items (links), twee representatieve items (midden) en twee van de best passende items (rechts).

Hoewel een visuele inspectie van de itemkarakteristieke curves een redelijk tot vrij goede passing van de items laat zien, is het alleen op basis van de DIF tussen bao en sbo, in combinatie met het vrij grote aantal leerlingen, al wel te verwachten dat uit een statistische toets zal blijken dat het Rasch model niet past bij de data. Om te onderzoeken of dit praktische implicaties heeft, vergelijken we de belangrijkste uitkomsten van de peiling gebaseerd op het Rasch model met dezelfde uitkomsten wanneer we een beter passend model gebruiken, de zogenaamde praktische significantie van misfit (Köhler en Hartig 2017).

In de peiling Rekenen-wiskunde 2019 is gekozen om gebruik te maken van OPLM (Verhelst, Glas, en Verstralen 1995). Het is daarom logisch om dit model hier ook in de vergelijking op te nemen. Omwille van de consistentie zouden we zelfs kunnen overwegen om in 2023 ook voor OPLM te kiezen. Bij gebruik van OPLM worden de discriminatieparameters geïmputeerd en vervolgens de locatieparameters geschat. Dit imputeren gebeurt gebruikelijk met een Marginal Maximum Likelihood methode (het programma OPCAT) of een Joint Maximum Likelihood methode (het programma OPSUG), waarbij OPCAT het meest wordt gebruikt. In het technisch rapport van de peiling Rekenen-wiskunde 2019 wordt niet vermeld welke methode is gebruikt. Op grond van een heranalyse van de data uit 2019 lijkt het echter vrijwel zeker dat OPCAT is gebruikt. Gebruik van een MML-procedure zoals OPCAT brengt echter een risico met zich mee. In 2019 zijn de toetsboekjes toegewezen aan leerlingen op grond van het vooraf door de school aangeleverde uitstroomprofiel (Buisman e.a., 2021, 28), zodat leerlingen een toets kregen voorgelegd die zo goed mogelijk paste bij hun verwachte vaardigheid. Daarnaast hebben sbo-leerlingen zowel in 2019 als in de proeftoets van 2022 en de peiling in 2023 aparte toetsversies gekregen. Dit staat bekend als een zogenaamd

targeted design. Wanneer bij gebruik van een MML-methode in geval van een *targeted design* géén rekening wordt gehouden met deze selectieve toewijzing van toetsboekjes leidt dit in de analyse tot een vertekening (bias) in de schatting van de itemparameters (Eggen en Verhelst 2011). Het programma OPCAT geeft geen mogelijkheid tot inachtneming van dit design en zou daarom strikt genomen in dit geval niet gebruikt mogen worden.

Naast de OPLM-methode zoals gebruikt in 2019 nemen we ook een 2PL model (Birnbaum, Lord, en Novick 1968) op in de vergelijking. In de schatting van een 2PL model is het wel mogelijk om rekening te houden met het *targeted design*. Dit kan door het specificeren van een aparte normaalverdeling voor elk niveau van de achtergrondvariabele op grond waarvan boekjes zijn toegewezen (Bock en Zimowski 1997).

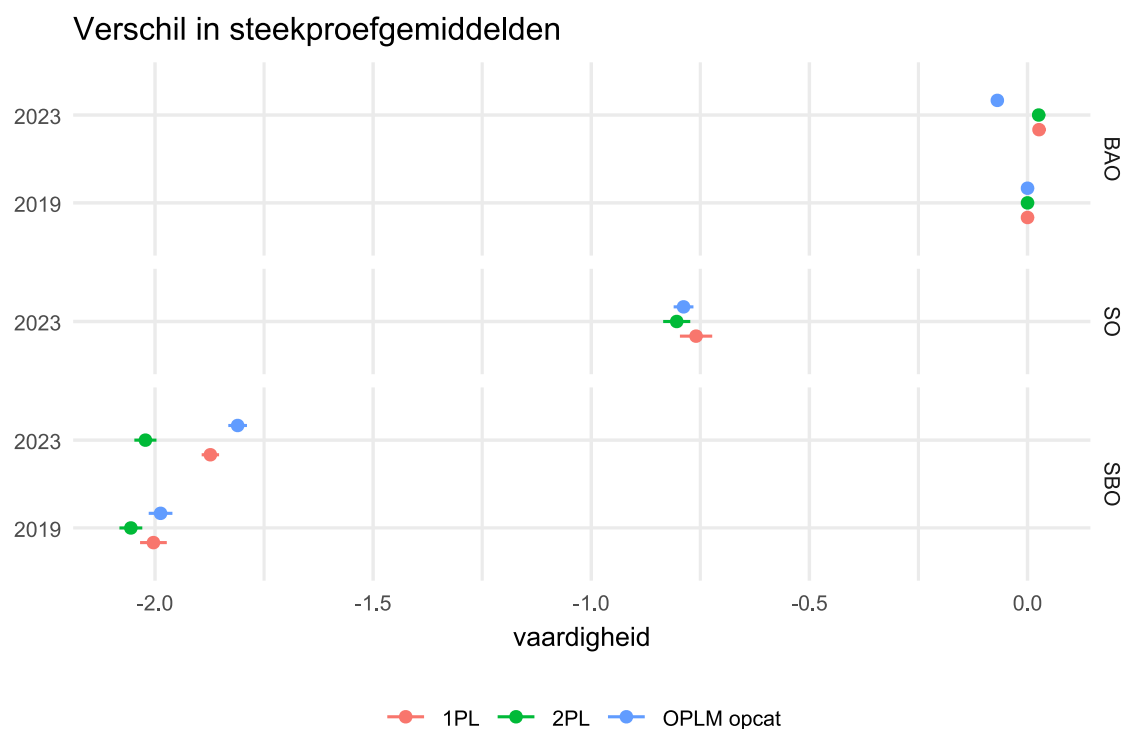
Voor elk van de drie modellen bepalen we eerst de passing met behulp van de zogenaamde S- en R1c-toetsen (Verhelst, Glas, en Verstralen 1995). Vervolgens bepalen we de belangrijkste uitkomsten van de peiling op basis van elk van de modellen. Zoals verwacht laten de analyses met het OPLM/OPCAT model en het 2PL een aanmerkelijk betere passing zien dan die van het 1PL.

Tabel 6.4 – S-toets en R1c toets bij 1PL/Rasch, 2PL en OPLM, Peil.Rekenen-wiskunde PO 2023*

Model	items	S toets			R1c toets			
		p.01	p.05	p.10	statistiek	df	p	Φ
1PL	469	0.382	0.52	0.61	15,510	6,792	0	1.5
OPLM opcat	469	0.085	0.19	0.26	10,084	6,792	0	1.2
2PL	469	0.068	0.13	0.21	9,628	6,792	0	1.2

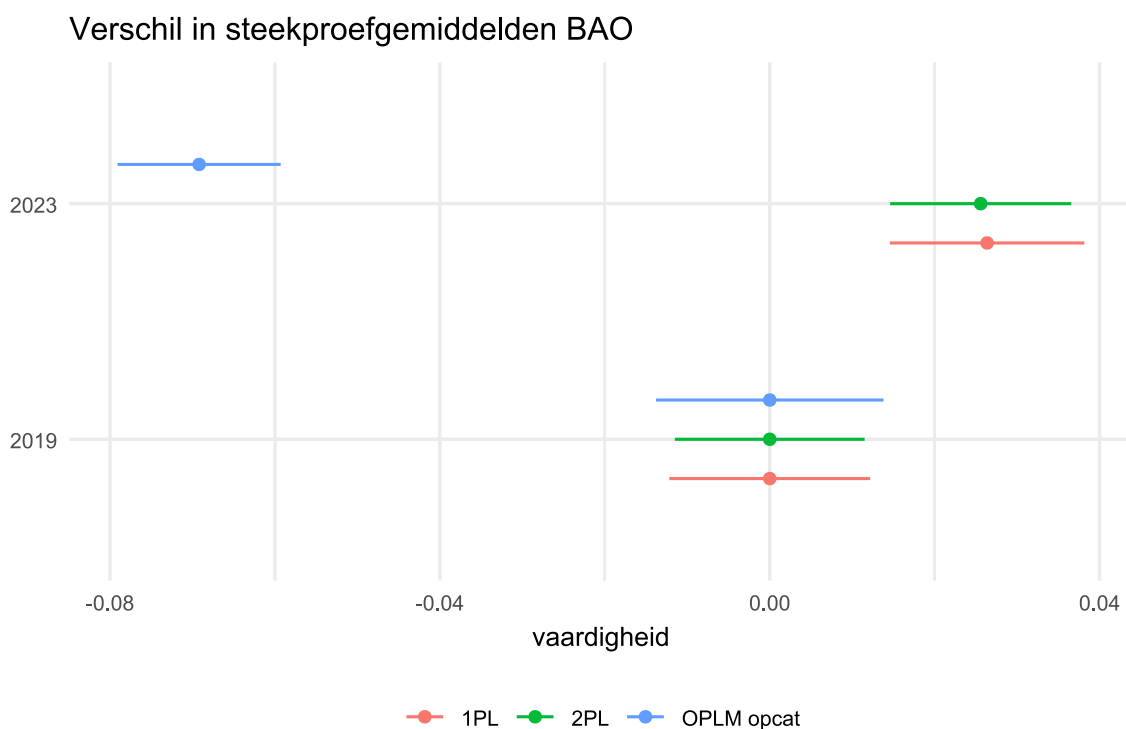
* De kolommen genaamd 'p' duiden overschrijdingskansen aan. p.01, p.05 en p.10 geven de proporties items waarvan de S toets een overschrijdingskans heeft kleiner dan respectievelijk 1%, 5% en 10%.

Om de praktische implicaties van het verschil in passing te bepalen, schatten we met behulp van de drie modellen telkens de vaardigheid van de leerlingen in 2019 en 2023, binnen in de verschillende schooltypes. Het resultaat is weergegeven in Figuur 6.13. De overeenstemming tussen de resultaten op basis van het 1PL en 2PL is zeer groot. De relatieve afstand tussen het sbo en bao is wel enigszins groter op basis van het 2PL dan op basis van het 1PL en de vooruitgang in het sbo tussen 2019 en 2023 die wordt gevonden op basis van het 1PL in het 2PL is relatief iets kleiner. Hierbij moeten we opmerken dat er ook in het 2PL DIF is tussen bao en sbo. De correlatie tussen de discriminatieparameters, afzonderlijk geschat in het so4/sbo en in het bao, is bijvoorbeeld slechts 0.6; de correlatie tussen de locatie parameters is 0.56. Dit houdt in dat een betrouwbare vergelijking tussen het bao en sbo moeilijker te maken is, zowel op basis van een 1PL als een 2PL model.



Figuur 6.13 – Schatting van de gemiddelde vaardigheid in de steekproef op basis van drie IRT modellen, Peil.Rekenen-wiskunde PO 2023. Om de uitkomsten te kunnen vergelijken is voor elk model de gemiddelde vaardigheid van de bao-leerlingen in 2019 op 0 gezet en de gepoolde standaarddeviatie van de vaardigheid in de 5 groepen op 1 gezet. De horizontale lijn geeft het 95% betrouwbaarheidsinterval van het steekproefgemiddelde, op basis van imputatie.

Wanneer we inzoomen op het bao (Figuur 6.14) is duidelijk te zien dat de uitkomsten op basis van het 1PL en 2PL vrijwel identiek zijn. Het resultaat op basis van het OPLM is echter wel verschillend, namelijk een achteruitgang tussen 2019 en 2023 in plaats van een lichte vooruitgang. Dit is waarschijnlijk een gevolg van de eerder besproken bias in de imputatie van de discriminatieparameters.



Figuur 6.14 – Schatting van de vaardigheid in de steekproef op basis van drie IRT modellen, ingezoomd op het bao, Peil.Rekenen-wiskunde PO 2023.

Wat betreft de keuze tussen een 1PL en een 2PL geven wij de voorkeur aan het gebruik van een 1PL boven een 2PL om verschillende redenen. Ten eerste zijn de noodzakelijke aannames bij gebruik van een CML procedure zeer minimaal. Voor het gebruik van een 2PL is het echter van belang om een correcte specificatie van de achtergrondvariabelen te geven. Verder wordt formeel een random steekproef en een gekende vorm van de populatieverdeling vereist. De aanname van een normaal verdeelde populatie per niveau van de achtergrondvariabele lijkt weliswaar zeer redelijk en een MML-procedure is over het algemeen robuust voor kleine schendingen van normaliteit, mits de achtergrondvariabelen correct gespecificeerd zijn. Dit is echter geen zekerheid en de betreffende aannames zijn niet te toetsen.

Daarnaast lijkt het gebruik van een 2PL niet tot een afname in de DIF tussen het so4/sbo en het bao te leiden. In de 2PL schaling krijgen moeilijkere items ook een iets hogere discriminatieparameter; de rangorde correlatie tussen de discriminatie- en locatieparameters in de 2PL schaling is 0.26. Moeilijkere items wegen dus zwaarder. Dit lijkt een onwenselijk artefact van de schaling die de interpretatie van de resultaten verder bemoeilijkt. Op grond van al deze overwegingen lijkt een 1PL de beste optie.

6.3.6 Samenhang tussen onderdelen

De samenhang tussen domeinen en onderdelen bepalen we door middel van latente correlaties. Daarvoor wordt een afzonderlijke schaal geschat voor elk van de domeinen waarna *plausible values* worden getrokken met een multivariaat normale prior. De latente correlaties worden verkregen uit de updates van de covariantiematrix van de prior. De samenhang tussen de inhoudelijke domeinen is zeer sterk, de latente correlaties zijn nagenoeg allemaal groter dan 0.9 (Tabel 6.5). We vinden het

daarom gerechtvaardigd om een univariate schaal te gebruiken. Wel is het zo dat de relatieve moeilijkheid van de domeinen behoorlijk verschilt. Hier wordt in hoofdstuk 7 verder op ingegaan.

Tabel 6.5 – Latente correlaties tussen domeinen Rekenen-wiskunde, Peil.Rekenen-wiskunde PO 2023

	getallen	meten en meetkunde	probleem oplossen	verbanden	verhoudingen
getallen		0.96	0.93	0.95	0.97
meten en meetkunde	0.96		0.91	0.95	0.97
probleemoplossen	0.93	0.91		0.91	0.89
verbanden	0.95	0.95	0.91		0.92
verhoudingen	0.97	0.97	0.89	0.92	

In de peiling 2019 werd voor rekenen mét en rekenen zonder rekenmachine nog een afzonderlijke schaal gebruikt. De latente correlatie tussen rekenen met rekenmachine en rekenen zonder rekenmachine in 2023 is echter 0.85 wat duidt op een redelijk sterke samenhang. De latente correlatie is geschat op een subset van leerlingen die rekenmachine items hebben gemaakt, het merendeel van de leerlingen had dat niet. Hoewel de latente correlatie iets lager is dan die tussen de domeinen zien we hierin onvoldoende reden om dit op een afzonderlijke schaal te kalibreren.

6.4 Betrouwbaarheid

In de peiling baseren we onze conclusies over de vaardigheid op een IRT model. De voor een peiling relevante betrouwbaarheid is daarom de betrouwbaarheid van de vaardigheidsschatting. Deze betrouwbaarheid is gedefinieerd als de proportie ware variantie in de populatie als onderdeel van de variantie in de schatting. Deze kan worden bepaald door de correlatie te nemen tussen uitwisselbare replicaties van de schatting van de vaardigheid (Bechger e.a., 2003). Wij gebruiken hiervoor de gemiddelde correlatie tussen een set van 100 plausible values. Zoals in hoofdstuk 7 verder zal worden besproken zijn *plausible values* onafhankelijke trekkingen uit de posterior verdeling van vaardigheid. Zoals in Tabel 6.6 is te zien is de betrouwbaarheid van de toets iets hoger in 2023 in vergelijking met 2019, dit komt door het grotere aantal items per toetsboekje in 2023.

Tabel 6.6 – IRT betrouwbaarheid van de toets, Peil.Rekenen-wiskunde PO 2023

Jaar	Schooltype	Betrouwbaarheid
2019	bao	0.83
	sbo	0.84
2023	bao	0.89
	sbo	0.87
	so4	0.90

6.5 Referentieniveaus

Voor Rekenen-wiskunde is een landelijke ankerset beschikbaar met prestatiestandaarden voor de referentieniveaus 1F en 1S. Deze zijn in de peiling Rekenen-wiskunde 2019 overgebracht naar de toen gebruikte vaardigheidsschaal. Zoals hierboven reeds besproken, gebruiken we voor de peiling 2023 een Rasch model, waar voor de peiling in 2019 nog een OPLM werd gebruikt. De grenzen op de vaardigheidsschaal die zijn bepaald in 2019 hebben daarom voor de peiling van 2023 geen betekenis. Om toch tot een zo vergelijkbaar mogelijk resultaat te komen hebben we de frequenties van de bepaalde referentieniveaus in de peiling 2019 (Buisman e.a., 2021, Tabel 8.2) gebruikt om grenzen te bepalen op de nieuwe schaal die leidt tot zoveel gelijke percentages referentieniveaus. Dit is gedaan door middel van een equipercntiel methode (Kolen en Brennan, 2014) waarbij we op basis van 100 *plausible value* trekkingen de percentielen behorende bij de percentages referentieniveaus voor het bao in 2019 hebben bepaald. We kiezen hier voor het bao omdat de steekproef bao-leerlingen in 2019 ruim vijfenhalf maal zo groot is als die voor het sbo. Daarnaast liggen voor het bao de percentages behaalde referentieniveaus verder van 0% en 100% af, waardoor deze bepaling veel preciezere resultaten geeft. Omdat er sprake is van DIF tussen schooltypes en we een ander psychometrisch model gebruiken dan in 2019, leidt dit niet tot de exact zelfde percentages referentieniveaus voor het sbo in 2019 als die in de eerdere peiling werden gerapporteerd. Deze verschillen zijn echter klein en liggen binnen de betrouwbaarheidsintervallen van de geschatte percentages (zie hoofdstuk 7).

7. Uitkomsten rekenvaardigheid

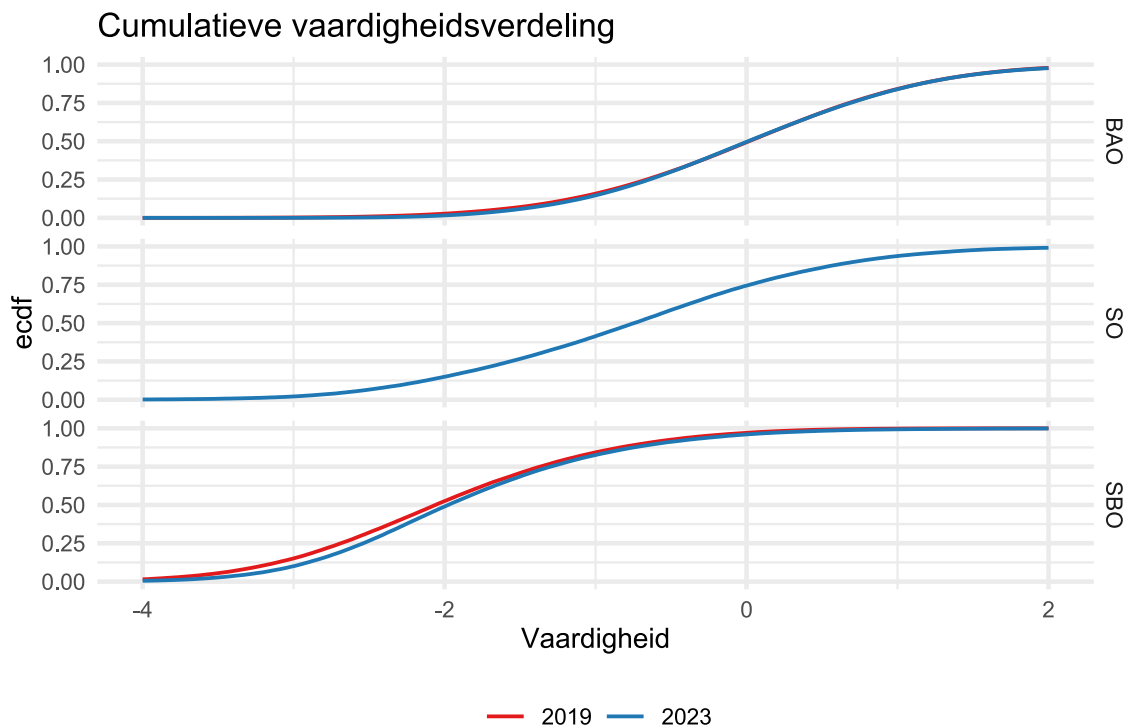
7.1 Vergelijking met 2019

Voor het beschrijven van de vaardigheid maken we gebruik van zogenaamde *plausible values*. Dit zijn willekeurige trekkingen uit de posterior verdeling van de vaardigheid. Door het gebruik van plausible values wordt de onderliggende vaardigheidsverdeling beter beschreven dan bij gebruik van *maximum likelihood* schatters (Marsman e.a., 2016). Daarnaast kan door het gebruiken van meerdere plausible values de onzekerheid over de vaardigheid van de leerlingen eenvoudig meegenomen worden in de bepaling van het betrouwbaarheidsinterval; dit wordt vaak de imputatiefout of meetfout genoemd. De plausible values zijn onafhankelijk getrokken per schooltype. Uiteraard zijn deze wel gebaseerd op dezelfde itemparameters. Binnen elk schooltype is een hiërarchische normale prior gebruikt waarbij de groepen bestaan uit de combinatie van jaar en school (zie voor een technische beschrijving van deze prior: Gelman e.a., 2013, 288–90). We gebruiken 100 plausible value trekkingen per leerling.

Voor het bepalen van de betrouwbaarheid van de populatieschatting moet naast de imputatiefout ook de onzekerheid door het gebruik van een steekproef worden meegenomen. Dit betreft de onzekerheid omtrent uitspraken over de verdeling van de vaardigheid in de populatie die voortkomt uit het feit dat maar een relatief klein deel van de populatie deelneemt aan het onderzoek. Wanneer gebruik wordt gemaakt van een willekeurige steekproef van leerlingen is de steekproeffout voor het bepalen van het gemiddelde gelijk aan de standaarddeviatie van de vaardigheid in de steekproef gedeeld door de wortel van het aantal leerlingen in de steekproef. We hebben in Peil.Rekenen-wiskunde PO 2023 echter niet te maken met een compleet willekeurige steekproef (*simple random sample*), maar met een geclusterde steekproef waarbij eerst scholen zijn benaderd en scholen vervolgens met alle aanwezige klassen met leerlingen die tot de doelpopulatie behoren, deelnamen aan het onderzoek. De spreiding van de vaardigheid van de leerlingen binnen een klas of school is kleiner dan de spreiding van de vaardigheid in een willekeurig getrokken groep leerlingen uit verschillende scholen. Dit houdt in dat de resultaten van een groep leerlingen uit dezelfde school of klas minder informatie geven over de verdeling van de vaardigheid in de populatie dan die van een compleet willekeurige groep leerlingen uit de populatie. Om deze reden is de effectieve steekproefgrootte in een geclusterde steekproef kleiner dan het aantal deelnemende leerlingen. Voor het bepalen van de steekproeffout is een groepering gemaakt per school en niet per klas. De belangrijkste reden hiervoor is dat in de peiling Rekenen-wiskunde in 2019 geen informatie over alle klassen bekend is, maar wel over alle scholen. De grootte van de steekproef in het sbo, so4 en bao is niet representatief voor het aantal leerlingen in deze schooltypes. Om de gemiddelde vaardigheid in de populatie te bepalen zou daarom gebruik moeten worden gemaakt van weging. In deze peiling wordt apart gerapporteerd per schooltype waardoor een verschillende weging gebaseerd op schooltype overbodig is.

Voor de schatting van de gemiddelde en standaarddeviaties in de populatie, en de bijbehorende betrouwbaarheidsintervallen, zijn voor elke plausible value trekking de populatieparameters geschat en de bijbehorende steekproeffouten bepaald. Deze zijn vervolgens gecombineerd over de trekkingen (zie Lumley, 2010 voor een praktische beschrijving van deze methode).

De empirische cumulatieve verdeling van plausible values op basis van de steekproef laat zien dat er niet of nauwelijks waarneembare verschillen zijn in de vaardigheid voor bao in 2023 in vergelijking met 2019. Voor het sbo geldt dat niet. Daar lijkt sprake van een toegenomen vaardigheid waarbij met name de onderkant van de verdeling is opgeschoven (Figuur 7.1).



Figuur 7.1 – Gemiddelde empirische cumulatieve verdeling op basis van 100 plausible value trekkingen, Peil.Rekenen-wiskunde PO 2023. Voor het so4 is geen meting in 2019. Voor het bao en sbo geldt dat waar slechts één lijn zichtbaar is, de vaardigheidsverdeling in 2019 en 2023 visueel niet te onderscheiden is.

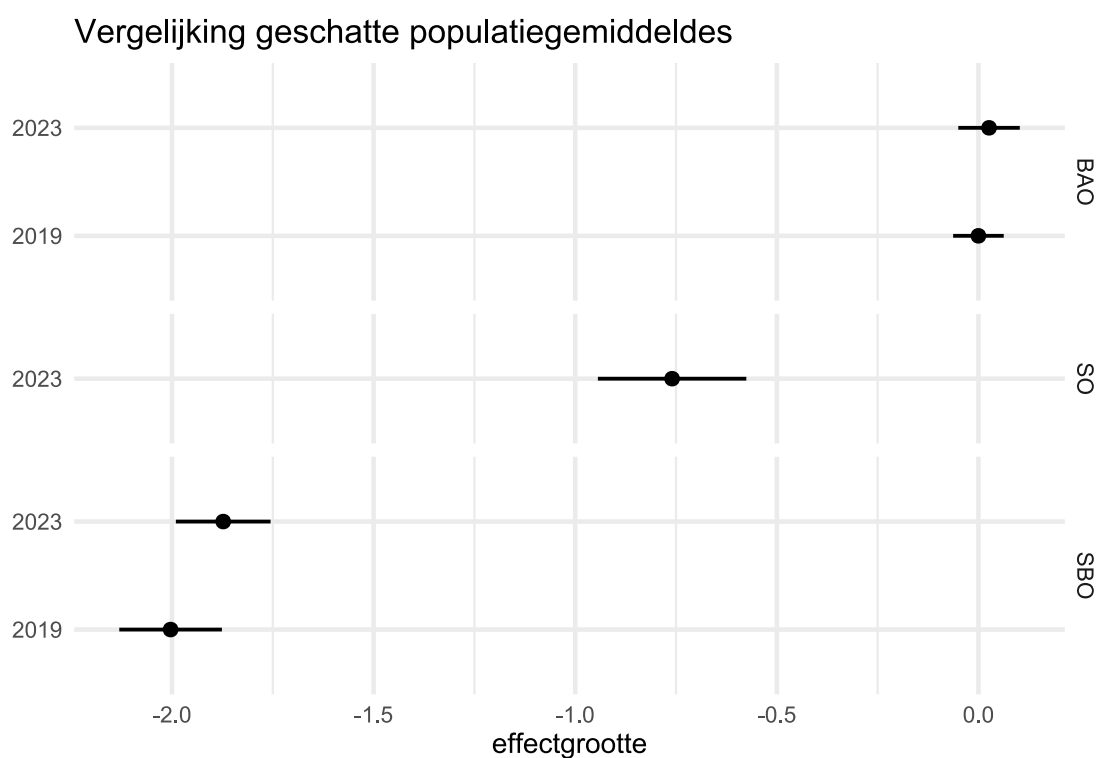
Het nulpunt en de schaal van de latente vaardigheid in een IRT model mogen arbitrair worden gekozen. Om de interpretatie van de resultaten te vergemakkelijken hebben we voor een schaal gekozen waarbij de gemiddelde geschatte vaardigheid in het bao in 2019 gelijk is aan 0 en waarbij de gepoolde standaarddeviatie van de vijf groepen (bao en sbo in 2019 en 2023, so4 in 2023) gelijk is gesteld aan 1. Verschillen op deze vaardigheidsschaal kunnen daarom worden gelezen als effectgroottes. Deze zelfde schaal wordt ook gebruikt in de meerniveau analyses in hoofdstuk 9. In Tabel 7.1 worden de geschatte gemiddelde vaardigheden gegeven. Ook hier valt te zien dat er een groot verschil is tussen de schooltypes en een klein verschil tussen de jaren.

Tabel 7.1 – Schatting van de gemiddelde vaardigheid in de populatie, Peil.Rekenen-wiskunde PO 2023*

Schooltype	Jaar	aantal ll.	gemiddelde	verdeling
bao	2019	5,014	0.00	
	2023	3,293	0.03	
so4	2023	504	-0.76	
sbo	2019	850	-2.00	
	2023	989	-1.87	

* De schaal van de vaardigheid is zo gekozen dat de gemiddelde vaardigheid in bao 2019 gelijk is aan 0 en de gepoolde standaarddeviatie over de vijf groepen gelijk is aan 1. Dientengevolge kunnen verschillen in vaardigheid geïnterpreteerd worden als effectgroottes ten opzichte van bao 2019.

In Figuur 7.2 wordt nog eens ingezoomd op de schatting van de gemiddeldes.



Figuur 7.2 – Geschatte gemiddelde vaardigheid in de populatie als effectgrootte, Peil.Rekenen-wiskunde PO 2023. De horizontale lijn geeft het 95% betrouwbaarheidsinterval met inachtneming van de schattingonzekerheid van de vaardigheid en het steekproefdesign.

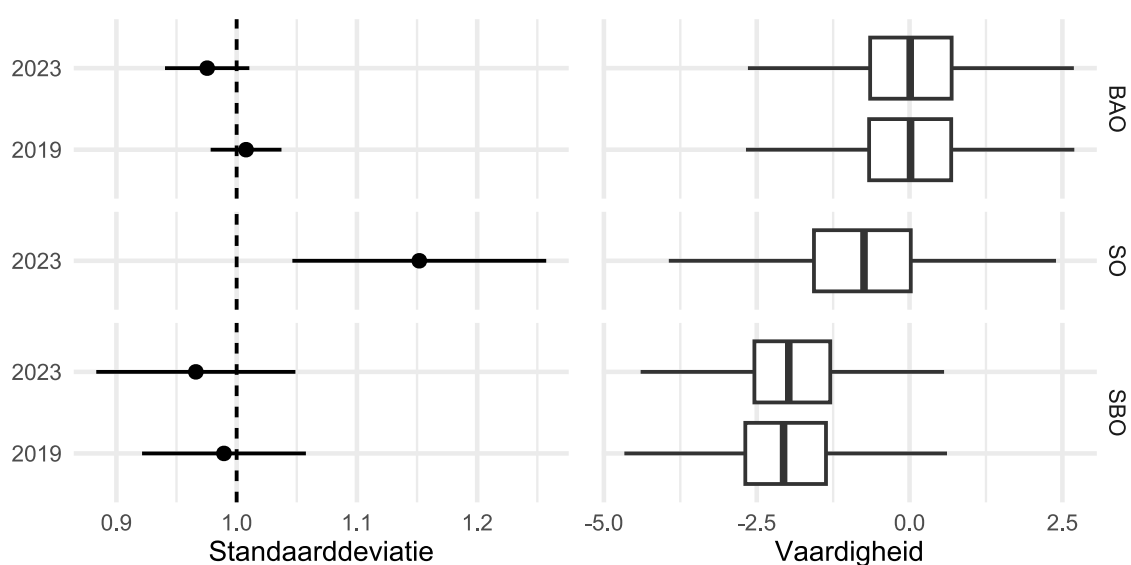
De meest prangende vraag in een peilingsonderzoek is vaak of de gemiddelde vaardigheid in de populatie voor- of achteruit is gegaan. Als we een census van de populatie zouden doen en bij elke leerling een vrij lange toets zouden afnemen zouden we deze vraag met vrij grote zekerheid kunnen beantwoorden. Omdat de effecten relatief klein zijn en de steekproef, hoewel omvangrijk, zeker geen

census is kunnen we die vraag niet met zekerheid beantwoorden. We kunnen echter wel een inschatting maken van de waarschijnlijkheid dat het verschil in gemiddelde vaardigheid binnen een bepaald bereik ligt, zie Tabel 7.2. Het verschil in gemiddelde vaardigheid in het bao is met een redelijk grote waarschijnlijkheid zeer klein, i.e. de kans dat het verschil in gemiddeldes minder dan 0.1 effectgrootte voor of achteruit is gegaan, is groter dan 90%. In het sbo waren de steekproeven aanzienlijk kleiner waardoor er meer onzekerheid is over de trend. De resultaten geven een voorzichtig positief beeld in die zin dat een daling in gemiddelde vaardigheid relatief onwaarschijnlijk lijkt.

Tabel 7.2 – Geschatte effectgrootte en waarschijnlijkheden van ranges van effectgroottes van de voor- of achteruitgang van de gemiddelde vaardigheid per schooltype tussen 2019 en 2023, Peil.Rekenen-wiskunde PO 2023

Schooltype	effect	se	daling		stijging			
			> 0.1	0.1 ↔ 0	0 ↔ 0.1	0.1 ↔ 0.2	0.2 ↔ 0.3	> 0.3
bao	0.03	0.05	0.6%	29.4%	62.8%	7.2%	0.0%	0.0%
sbo	0.13	0.09	0.5%	6.5%	29.5%	41.9%	18.8%	2.7%

De spreiding van de vaardigheid lijkt duidelijk wat groter te zijn in het so4 dan in de andere schooltypes (Figuur 7.3). In het so4 is de vaardigheid dus wat meer gevarieerd.

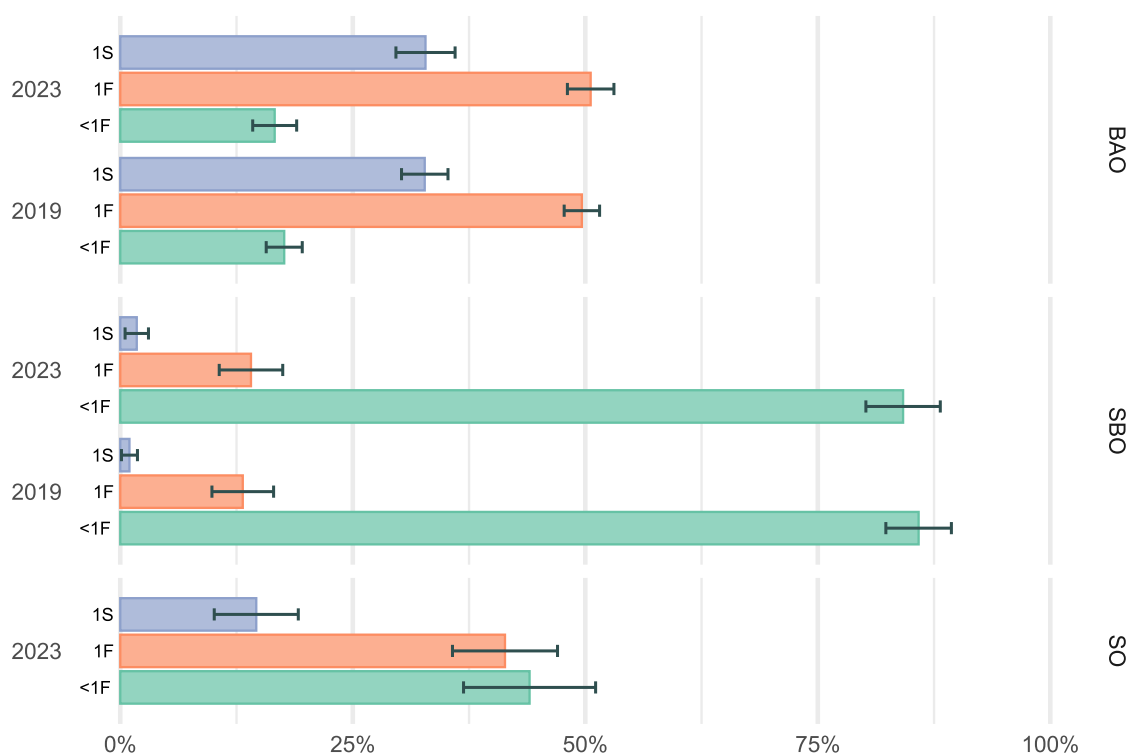


Figuur 7.3 – Spreiding van de vaardigheid, Peil.Rekenen-wiskunde PO 2023. Links de geschatte standaarddeviatie in de populatie inclusief het 95% betrouwbaarheidsinterval. De verticale onderbroken lijn geeft de gepoolde standaarddeviatie aan. Rechts de verdeling van de vaardigheid middels een boxplot.

7.2 Referentieniveaus

In de peiling Rekenen-wiskunde was één van de doelen om de prestaties van leerlingen te relateren aan de vastgestelde referentieniveaus. Voor de bepaling van de prestatiestandaarden hebben we de in 2019 gerapporteerde percentages referentieniveaus in het bao in 2019 vertaald naar grenzen op de vaardigheidsschaal, zoals beschreven in hoofdstuk 6.

De indeling in referentieniveaus gebeurt op grond van de geschatte vaardigheid. Omdat de vaardigheid van een leerling niet met zekerheid kan worden bepaald, geldt dit uiteraard ook voor het referentieniveau. Door het referentieniveau voor een leerling meerdere malen te bepalen, eens voor elke plausible value, zijn ook de referentieniveaus meervoudig geïmputeerd. In Figuur 7.4 en Tabel 7.3 is te zien dat de verschillen tussen 2019 en 2023 met grote waarschijnlijkheid gering zijn. In de steekproef voor het sbo uit 2023 zijn er minder leerlingen met referentieniveau onder 1F, maar de schatting van dit percentage in de populatie heeft een relatief grote onzekerheid.



Figuur 7.4 – Geschatte referentieniveaus in de populatie inclusief 95% betrouwbaarheidsinterval, Peil.Rekenen-wiskunde PO 2023.

In het sbo heeft zowel in 2019 als 2023 het grootste deel van de leerlingen niveau 1F nog niet behaald. In 2023 heeft circa 14% niveau 1F behaald en is het percentage leerlingen dat 1S behaalt gering (ongeveer 2%). In so4 behaalt het grootste deel van de leerlingen 1F of hoger (circa 41% 1F, 15% 1S). In het bao haalt ongeveer 17% van de leerlingen niveau 1F niet. Ongeveer 51% haalt 1F en 33% 1S.

Tabel 7.3 – Percentages referentieniveau per jaar en schooltype incl. 95% betrouwbaarheidsinterval, Peil.Rekenen-wiskunde PO 2023

Schooltype	Jaar	<1F		1F		1S	
bao	2019	17.6%	15.7-19.6	49.6%	47.7-51.5	32.7%	30.2-35.2
	2023	16.6%	14.2-19.0	50.6%	48.1-53.1	32.8%	29.6-36.0
sbo	2019	85.8%	82.3-89.4	13.2%	9.9-16.5	1.0%	0.1-1.8
	2023	84.2%	80.2-88.2	14.1%	10.6-17.5	1.8%	0.5-3.0
so4	2023	44.0%	36.9-51.1	41.4%	35.7-47.0	14.6%	10.1-19.1

7.3 Prestaties op domeinen en onderdelen

In Peil.Rekenen-wiskunde PO 2023 zijn items afgenomen binnen de bekende domeinen *getallen*, *verhoudingen*, *meten en meetkunde* en *verbanden*. Daarnaast zijn er voor het eerst ook items afgenomen in het domein *probleemoplossen*. Ook wordt er een onderscheid gemaakt tussen items met context en zogenoemde kale items. Bij een klein aantal items is het toegestaan een rekenmachine te gebruiken. Voor deze soorten en types items bestaan geen verschillende referentieniveaus dus het referentieniveau van de leerlingen kan alleen worden bepaald op de volledige schaal.

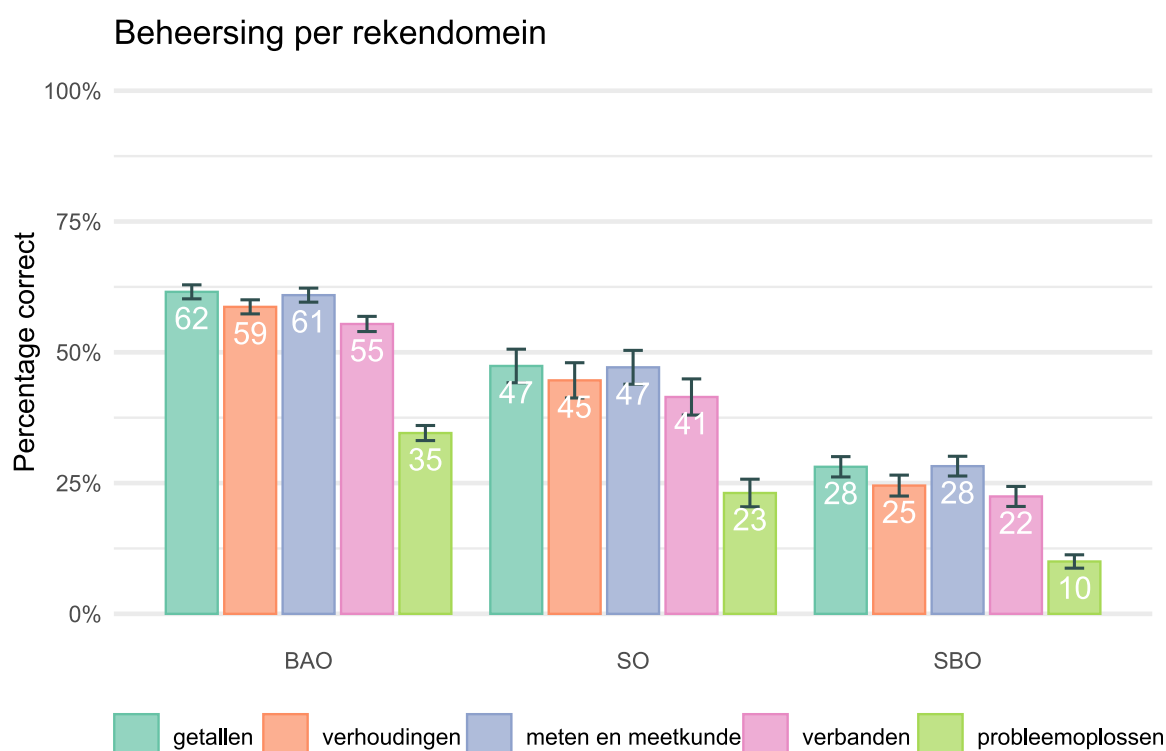
Om een indruk te krijgen van de prestaties van de leerlingen op de verschillende domeinen en onderdelen kijken we naar de percentages correcte antwoorden. Aangezien elke leerling maar een beperkt deel van de items heeft beantwoord (circa 14%), moeten we de ontbrekende antwoorden van de leerlingen imputeren. Dit imputeren gebeurt op basis van de plausible values en de relatieve moeilijkheid van de items die niet zijn beantwoord, met andere woorden: op basis van het geschatte psychometrische model. Dit kan enige bias met zich meebrengen. Het is bijvoorbeeld aannemelijk dat leerlingen bij items die veel te moeilijk voor hen zijn, eerder geneigd zullen zijn om het antwoord te raden. Hoeveel kans zij hebben om een antwoord goed te raden hangt af van het type item, de inhoud van het item en mogelijk andere zaken. De kans om een antwoord goed te raden is echter niet opgenomen in het model. Dit zou ervoor kunnen zorgen dat de kans dat leerlingen hele moeilijke items goed beantwoorden zal worden onderschat.

7.3.1 Rekendomeinen

Figuur 7.5 maakt in één oogopslag duidelijk dat de items binnen het domein *probleemoplossen* erg moeilijk zijn voor de meeste leerlingen in alle schooltypes. Voor het sbo geldt dat ook de items in de overige domeinen voor de meeste leerlingen erg moeilijk zijn met slechts ongeveer 25% correcte antwoorden. In het so4 zijn de items, met uitzondering van *probleemoplossen* uitdagend, maar voor een redelijk deel van de leerlingen wel te beantwoorden. De percentages correcte antwoorden liggen daar tussen de 40 en 50%. In het bao wordt, met uitzondering van *probleemoplossen*, rond de 60% van de vragen correct beantwoord.

Hierbij moet wel worden opgemerkt dat de items die zijn afgenomen bij de leerlingen in het sbo makkelijker waren dan de items die zijn afgenomen in het bao. Door de overige antwoorden te imputeren maken we een schatting van het percentage correct op de gehele itembank, maar dit is

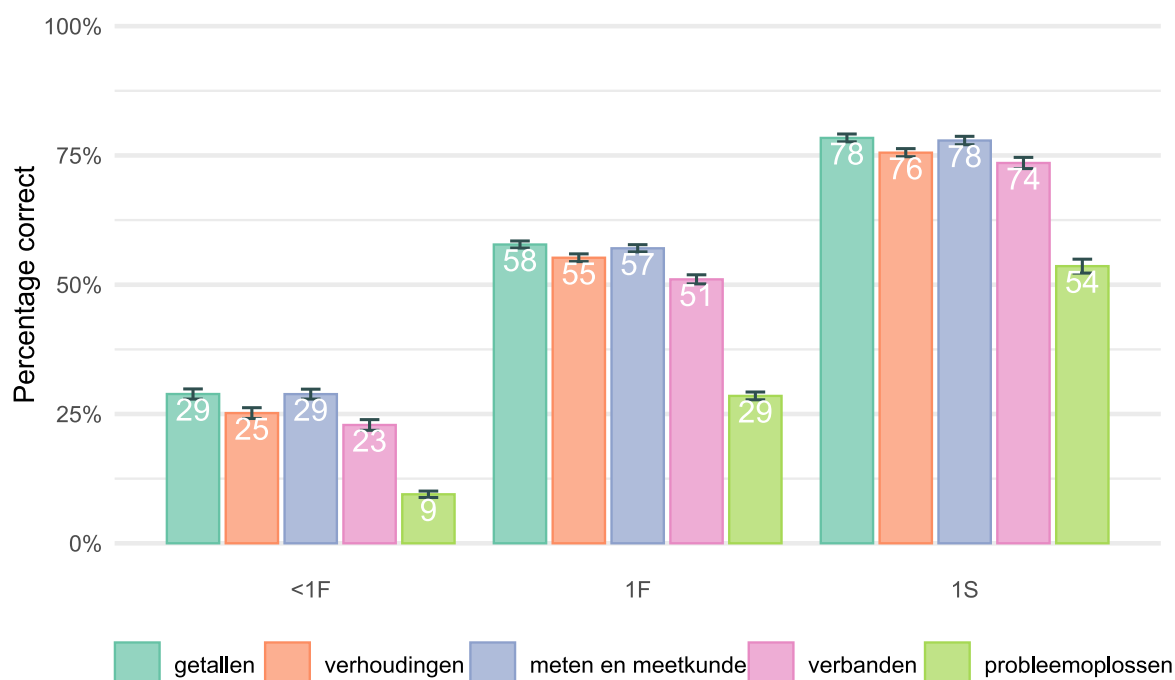
niet representatief voor de moeilijkheid van de toets zoals ervaren door de leerling. De relatieve moeilijkheid van de toetsen met het oog op de ervaren moeilijkheid wordt beschreven in hoofdstuk 6.



*Figuur 7.5 – Domeinbeheersing in 2023 naar schooltype incl. 95% betrouwbaarheidsinterval;
Peil.Rekenen-wiskunde PO 2023*

Op de schaal van de referentieniveaus zien we dat voor de leerlingen met referentieniveau 1S het merendeel van de items relatief eenvoudig te beantwoorden is, met een percentage correct rond de 75%. Voor probleemoplossen ligt dat percentage lager. Voor de leerlingen met niveau 1F liggen de percentages tussen 50% en 60% wat toets-technisch gezien vrij ideaal is. Probleemoplossen is voor deze categorie relatief moeilijk met een percentage correct van ongeveer 29%. Voor leerlingen met niveau onder 1F zijn de meeste items erg moeilijk.

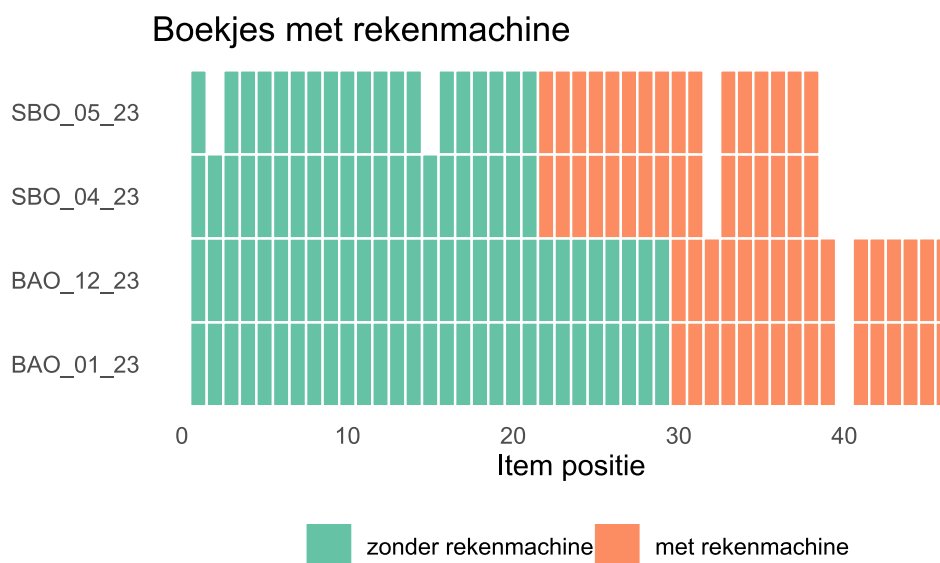
Beheersing per rekendomein



Figuur 7.6 – Domeinbeheersing in 2023 naar referentieniveau incl. 95% betrouwbaarheidsinterval;
Peil.Rekenen-wiskunde PO 2023

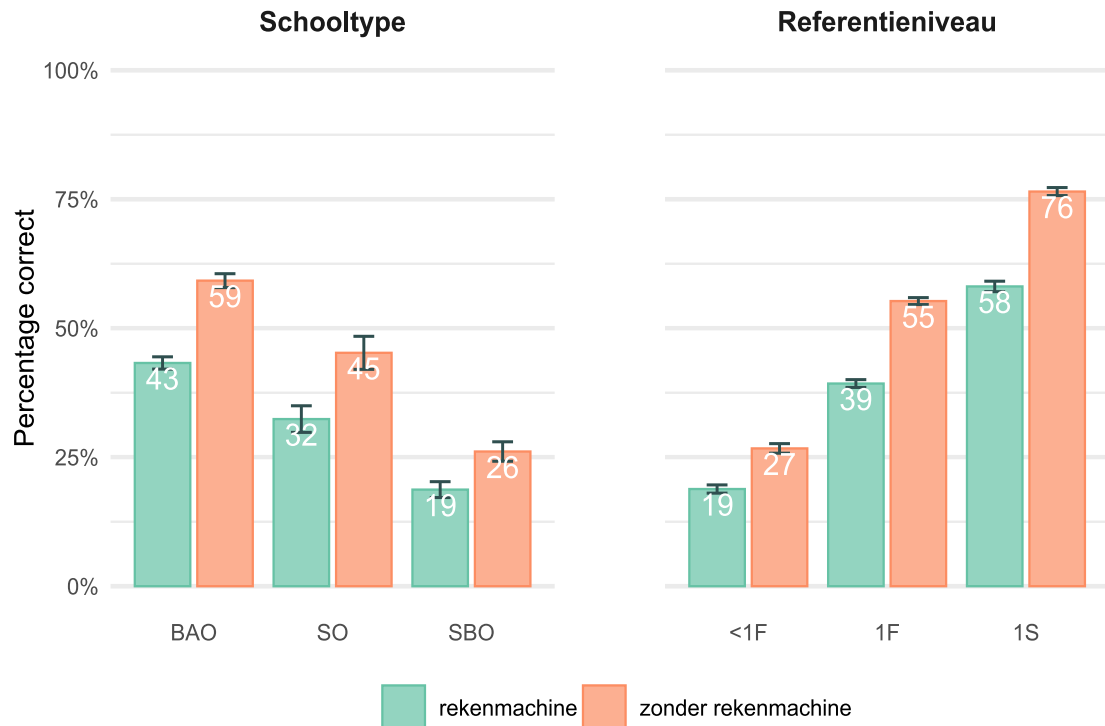
7.3.2 Rekenmachine

Er zijn 18 items met een rekenmachine in de analyse (één item is verwijderd). Deze items zijn verdeeld over vier van de in totaal 32 boekjes met elk 16 rekenmachine-items (Figuur 7.7). Het merendeel van de leerlingen heeft geen items met rekenmachine gemaakt en de leerlingen in het so4 en sbo hebben grotendeels dezelfde rekenmachine-items gemaakt als die in het bao. Rekenmachine-items zijn aan het einde van een boekje afgenomen.



Figuur 7.7 – Design van de vier boekjes met rekenmachine-items, Peil.Rekenen-wiskunde PO 2023. De 'SBO'-boekjes zijn ook afgenomen in so4. De x-as geeft de itempositie aan. Door het incomplete design staan op dezelfde positie in de verschillende boekjes andere items. Lege cellen geven items aan die uit de analyse zijn verwijderd.

Omdat de vaardigheid van het deel van de leerlingen dat rekenmachine-items heeft gemaakt niet representatief hoeft te zijn voor de gehele steekproef of populatie, zijn ook voor leerlingen die geen rekenmachine-items hebben gemaakt scores op de rekenmachine-items geïmputeerd. De leerlingen maken items zonder rekenmachine beter dan items met rekenmachine (Figuur 7.8). In so4 en vooral in sbo zijn de rekenmachine-items zeer moeilijk voor de leerlingen. Het is overigens niet te verwachten dat dit vaak leidt tot correct raden van het antwoord aangezien 15 van de 18 rekenmachine-items open items zijn.



Figuur 7.8 – Beheersing items met/zonder rekenmachine in 2023 naar schooltype en referentieniveau inclusief 95% betrouwbaarheidsinterval, Peil.Rekenen-wiskunde PO 2023

Er is zoals gezegd een zeer klein aantal *gesloten* items met rekenmachine. Die zijn wel opvallend veel makkelijker dan de open items (Tabel 7.4). Omdat het slechts om drie items gaat is het echter lastig om te zeggen wat de oorzaak van dit verschil is. Het kan evengoed zo zijn dat deze drie items toevallig of om een andere inhoudelijke reden relatief makkelijk zijn.

Tabel 7.4 – Percentages correcte antwoorden in de steekproef op de items met rekenmachine onderverdeeld naar open en gesloten items, Peil.Rekenen-wiskunde PO 2023

Schooltype	aantal leerlingen	item soort	aantal items	percentage correct
bao	560	gesloten	3	65.2%
		open	15	41.7%
sbo	248	gesloten	3	37.2%
		open	15	17.3%
so4	123	gesloten	3	56.1%
		open	15	30.8%

7.3.3 Kaal/context

De focus bij de kaal/context vergelijking ligt op het uitvoeren van de bewerkingen *optellen*, *aftrekken*, *vermenigvuldigen* en *delen*, kaal of in een beschreven situatie (context). Om deze reden worden alleen opgaven uit de domeinen *getallen* en *verhoudingen* gebruikt voor de kaal/context indeling. Ook opgaven waarbij een rekenmachine mag worden gebruikt worden uit de analyse gelaten.

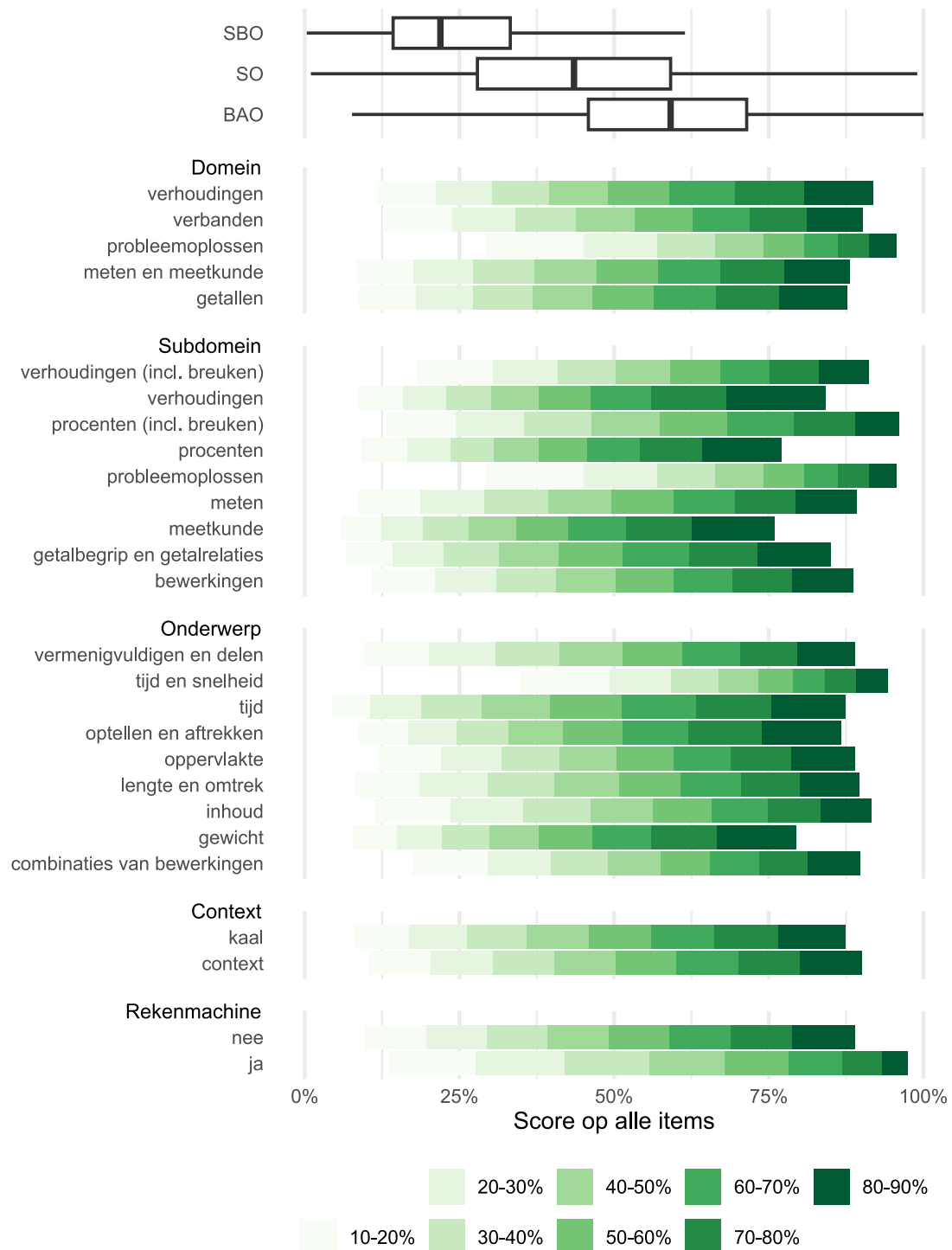
Leerlingen in alle schooltypes maken items die worden geclassificeerd als kaal ongeveer even goed als items die zijn geclassificeerd als context (Figuur 7.9).



Figuur 7.9 – Beheersing kaal/context-items naar schooltype en referentieniveau, Peil.Rekenen-wiskunde PO 2023

7.4 Beschrijving van de schaal

In een univariaat model kunnen we met behulp van de item-moeilijkheid de verschillende onderdelen op de rekenschaal leggen om zo de relatieve moeilijkheid te relateren aan de populaties. Figuur 7.10 laat de relatieve moeilijkheden zien van de domeinen en onderdelen gerelateerd aan het percentage correct op de verzameling items die is afgenomen in de peiling.



Figuur 7.10 – Alle domeinen en onderdelen op de rekenschaal, Peil.Rekenen-wiskunde PO 2023. De x-as geeft het percentage correct op alle items in de bank. De gekleurde balken geven de decielen van verwachte correct scores gegeven de bankscore. Ter referentie zijn de vaardigheidsverdelingen van de drie onderwijstypes toegevoegd.

8. Uitkomsten contextvragenlijsten

8.1 Inleiding

In dit hoofdstuk worden de uitkomsten vergeleken tussen bao, sbo en so4 van de achtergrond- en domeinspecifieke (onderwijs)kenmerken van leerlingen, leerkrachten en scholen. Paragraaf 8.2, 8.3 en 8.4 beschrijven de uitkomsten van respectievelijk de leerling-, leerkracht- en schoolvragenlijsten.

De leerlingvragenlijsten voor bao, sbo en so4 zijn grotendeels identiek. De versies bestemd voor de sbo- en so4-leerlingen hadden alleen een uitgebreidere invulinstructie dan de versie voor groep 8. De leerkrachtvragenlijsten zijn voor alle drie de onderwijssoorten eveneens gelijk, met uitzondering van de vraag naar de onderrichtervaring van de leerkracht. Omdat in het reguliere basisonderwijs de schoolvragenlijst van TIMSS is afgenomen (mede gericht op groep 6 en op natuurwetenschappen), zijn de schoolvragenlijsten voor sbo en so4 wel verschillend van de bao-versie. In paragraaf 8.4 wordt over een beperkt aantal schoolkenmerken gerapporteerd die in alle drie de schoolvragenlijsten aan bod zijn gekomen.

Een groot aantal van de domeinspecifieke (onderwijs)kenmerken van leerlingen, leerkrachten en scholen zijn in de vragenlijsten gemeten met een construct bestaande uit een set items, veelal in de vorm van stellingen die aan de respondent zijn voorgelegd. Er is in de vragenlijsten vooral gebruik gemaakt van bestaande constructen die eerder voor TIMSS en/of Peil.Rekenen-wiskunde PO 2019 zijn ontwikkeld. Voor deze constructen is nagegaan of zij een vergelijkbare betrouwbaarheid (Lambda-2) laten zien in de onderzoekspopulaties (bao, so4 en sbo) van Peil.Rekenen-wiskunde PO 2023.

Meestal geldt voor de waarden in de tabellen: hoe hoger de score, hoe positiever. De oorspronkelijke antwoordcategorieën in de vragenlijsten lopen veelal van positief naar negatief (bijvoorbeeld: 1=zeer mee eens, 4=zeer mee oneens). Deze antwoordcategorieën zijn gehercodeerd zodat zij van negatief naar positief lopen, behalve als het item of de stelling negatief geformuleerd was. Onder de tabellen wordt aangegeven waar de laagste en hoogste schaalscore voor staan. Voor tabellen met meerdere items (bijvoorbeeld stellingen) kan het aantal respondenten dat antwoord heeft gegeven per item verschillen. Voor de leesbaarheid is niet per stelling het aantal gegeven antwoorden weergegeven, maar het minimale en maximale aantal respondenten aangegeven (bijvoorbeeld 'n=3064-3235' in Tabel 8.5).

8.2 Leerlingvragenlijst

8.2.1 Deelname en ontbrekende waarden

Groep 8 regulier basisonderwijs

In totaal hebben 3.382 leerlingen in groep 8 deelgenomen aan het peilingsonderzoek. De leerlingvragenlijst is door 3.262 leerlingen ingevuld, de overige leerlingen hebben alleen de toets gemaakt. Leerlingen die meer dan de helft van de vragen in de leerlingvragenlijst niet hebben beantwoord, zijn uit het vragenlijstbestand verwijderd. Dit waren 21 leerlingen (0,6%). Dit betekent dat 3241 groep 8 leerlingen de vragenlijst helemaal of grotendeels ingevuld (Tabel 8.1).

Tabel 8.1 – Ontbrekende waarden in leerlingvragenlijst baso, Peil.Rekenen-wiskunde PO 2023

Aantal ontbrekende waarden	Aantal leerlingen
0	2016
1	954
2	167
3	32
4	18
5	4
6	5
7	5
8	9
9	13
≥ 10	18
Totaal	3241

Schoolverlaters speciaal basisonderwijs

In totaal hebben 1047 sbo-leerlingen deelgenomen aan het peilingsonderzoek. De leerlingvragenlijst is door 1025 leerlingen ingevuld, de overige leerlingen hebben alleen de toets gemaakt. Leerlingen die meer dan de helft van de vragen in de leerlingvragenlijst niet hebben beantwoord, zijn uit het bestand verwijderd. Dit waren 42 leerlingen (4,1%). In totaal hebben 983 sbo-leerlingen de vragenlijst helemaal of grotendeels ingevuld (Tabel 8.2).

Tabel 8.2 – Ontbrekende waarden in leerlingvragenlijst sbo, Peil.Rekenen-wiskunde PO 2023

Aantal ontbrekende waarden	Aantal leerlingen
0	529
1	276
2	67
3	33
4	15
5	6
6	5
7	9
8	6
9	5
≥ 10	32
Totaal	983

Schoolverlaters speciaal onderwijs

In totaal hebben 525 leerlingen in so4 deelgenomen aan het peilingsonderzoek. De leerlingvragenlijst is door 513 leerlingen ingevuld, de overige leerlingen hebben alleen de toets gemaakt. Leerlingen die meer dan de helft van de vragen niet hebben beantwoord, zijn uit het bestand verwijderd. Dit waren 13 leerlingen (2,5%). In totaal hebben 500 so4-leerlingen de vragenlijst helemaal of grotendeels ingevuld (Tabel 8.3).

Tabel 8.3 – Ontbrekende waarden in leerlingvragenlijst so4, Peil.Rekenen-wiskunde PO 2023

Aantal ontbrekende waarden	Aantal leerlingen
0	297
1	146
2	38
3	4
4	4
5	1
6	1
7	1
8	2
9	2
≥ 10	4
Totaal	500

8.2.2 Achtergrondkenmerken leerlingen

In Tabel 8.4 worden de achtergrondkenmerken van de leerlingen in percentages gepresenteerd.

Tabel 8.4 – Achtergrondkenmerken leerlingen, in percentages, Peil.Rekenen-wiskunde PO 2023

Kenmerk	% bao	% sbo	% so4
<i>Geslacht*</i>	<i>n=3232</i>	<i>n=977</i>	<i>n=499</i>
meisje	48,4	34,6	13,8
jongen	50,9	62,7	83,6
anders	0,7	2,7	2,6
<i>Frequentie thuis Nederlands gesproken*</i>	<i>n=3233</i>	<i>n=975</i>	<i>n=498</i>
nooit	2,2	3,6	2,8
soms	14,9	25,3	24,9
bijna altijd	14,1	17,0	22,9
altijd	68,8	54,1	49,4

Kenmerk	% bao	% sbo	% so4
<i>Aantal boeken thuis*</i>	<i>n=3212</i>	<i>n=971</i>	<i>n=493</i>
geen of weinig (0–10 boeken)	14,5	37,3	31,6
een plank vol (11–25 boeken)	31,8	35,8	33,1
een boekenkast vol (26–100 boeken)	34,1	18,0	22,7
twee boekenkasten vol (101–200 boeken)	11,7	3,6	5,1
drie of meer boekenkasten vol (meer dan 200 boeken)	7,9	5,3	7,5
<i>VO advies/uitstroomprofiel**</i>	<i>n=3163</i>	<i>n=924</i>	<i>n=485</i>
pro	0,6	44,8	13,8
vmbo-b	4,9	29,7	5,2
vmbo-b/k	3,9	5,0	3,7
vmbo-k	9,9	4,7	4,9
vmbo-k/gt	5,5	1,3	2,5
vmbo-gt	15,8	2,7	7,2
vmbo-gt/havo	10,3	1,3	2,1
havo	17,2	0,2	2,9
havo/vwo	10,2	0,2	2,9
vwo	21,2	0,4	1,9
Vso vmbo b/k	0,2	7,9	32,4
Vso vmbo g/t	0,1	1,5	11,3
Vso havo/vwo	0,1	0,3	9,3
<i>Vertraagd***</i>	<i>n=3228</i>	<i>n=967</i>	<i>n=500</i>
ja	11,5	44,8	34,6
nee	88,5	55,2	65,4

* Vraag 1.1, 1.3, 1.4 leerlingvragenlijst

** Op basis van op leerlingadministratielijsten van alle getoetste leerlingen voor zover deze door de leerkracht zijn ingevuld.

***Op basis van geboortedatum en '1 oktober-grens'. Er is vanuit gegaan dat een leerling geboren voor 1 oktober 2010 'vertraagd' is. Geboortedatum is gebaseerd op leerlingadministratielijst en waar nodig aangevuld met gegevens uit leerlingvragenlijst (vraag 1.2a en vraag 1.2b).

8.2.3 Domeinspecifieke leerlingkenmerken

In deze paragraaf worden de domeinspecifieke kenmerken (houdingen ten opzichte van Rekenen-wiskunde) van de leerlingen gepresenteerd. Deze kenmerken zijn gemeten aan de hand van een aantal uitspraken die aan de leerlingen zijn voorgelegd. Deze uitspraken gaan over hun rekenplezier, zelfvertrouwen in rekenen, waaraan zij goede rekenprestaties toewijzen (attributies) en de mate waarin leerlingen vinden dat zij zelf invloed kunnen hebben op deze prestaties (*growth mindset*). In de volgende tabellen worden de gemiddelden en standaarddeviaties weergegeven.

Tabel 8.5 – Plezier in rekenen volgens leerlingen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Stellingen	bao <i>n=3064-3235</i>	sbo <i>n=832-973</i>	so4 <i>n=455-498</i>
a. Ik vind het leuk om te leren rekenen	2,8 (0,9)	3,0 (1,0)	2,8 (1,0)
b. Ik zou willen dat ik geen rekenen hoefde te leren	2,7 (1,0)	2,7 (1,1)	2,5 (1,2)
c. Rekenen is saai	2,5 (1,0)	2,5 (1,1)	2,4 (1,1)
d. Ik leer veel interessante dingen bij rekenen	3,0 (0,9)	3,0 (1,0)	2,9 (1,0)
e. Ik vind rekenen leuk	2,6 (1,0)	2,7 (1,1)	2,6 (1,1)
f. Ik vind alle schooltaken leuk waarbij ik met getallen moet werken	2,4 (0,9)	2,5 (0,9)	2,4 (0,9)
g. Ik vind het leuk om rekensommen op te lossen	2,6 (1,0)	2,7 (1,1)	2,6 (1,0)
h. Ik verheug mij op de rekenles	2,1 (0,9)	2,4 (1,1)	2,2 (1,0)
i. Rekenen is één van mijn favoriete vakken	2,3 (1,2)	2,4 (1,2)	2,4 (1,2)
Construct Plezier in rekenen	2,6 (0,7)	2,7 (0,8)	2,6 (0,8)

Vraag 2.1

a, d, e, f, g, h, i: 1= zeer mee oneens, 4= zeer mee eens

b, c: 1= zeer mee eens, 4= zeer mee oneens

Tabel 8.6 – Zelfvertrouwen in rekenen volgens leerlingen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Stellingen	bao <i>n=3113-3224</i>	sbo <i>n=896-972</i>	so4 <i>n=479-498</i>
a. Meestal ben ik goed in rekenen	3,2 (0,9)	3,2 (0,8)	3,2 (0,8)
b. Rekenen is voor mij moeilijker dan voor veel van mijn klasgenootjes	2,9 (1,0)	2,7 (1,1)	2,7 (1,1)
c. Ik ben gewoon niet goed in rekenen	3,1 (1,0)	3,0 (1,1)	3,1 (1,1)
d. Rekenen is makkelijk voor mij	2,7 (0,9)	2,8 (1,0)	2,8 (1,0)
e. Ik ben goed in het oplossen van moeilijke rekensommen	2,7 (1,0)	2,7 (1,0)	2,7 (1,0)
f. Ik kan rekenen goed uitleggen aan anderen	2,7 (0,9)	2,7 (1,0)	2,7 (1,0)
g. Rekenen is voor mij moeilijker dan alle andere vakken op school	2,9 (1,1)	2,8 (1,1)	2,8 (1,1)
h. Ik vind rekenen moeilijk te snappen	3,0 (1,0)	2,8 (1,0)	2,8 (1,0)
Construct Zelfvertrouwen in rekenen	2,9 (0,8)	2,8 (0,7)	2,8 (0,7)

Vraag 2.4

a, d, e, f: 1= zeer mee oneens, 4= zeer mee eens

b, c, g, h: 1= zeer mee eens, 4= zeer mee oneens

Tabel 8.7 – *Attributies goede rekenprestaties volgens leerlingen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023*

Stellingen	bao <i>n=3185-3215</i>	sbo <i>n=949-975</i>	so4 <i>n=486-497</i>
a. Als ik een rekenopdracht goed maak komt dat vooral omdat ik goed mijn best heb gedaan	3,4 (0,7)	3,6 (0,7)	3,4 (0,8)
c. Als ik een rekenopdracht goed maak komt dat vooral omdat ik hard gewerkt heb	3,3 (0,8)	3,4 (0,7)	3,3 (0,8)
f. Als ik een rekenopdracht goed maak komt dat vooral omdat ik goed heb opgelet tijdens de les	3,1 (0,8)	3,3 (0,8)	3,1 (0,8)
Construct Attributies rekenprestaties	3,3 (0,6)	3,5 (0,6)	3,3 (0,6)

Vraag 2.5

a, c, f: 1= zeer mee oneens, 4= zeer mee eens

Tabel 8.8 – *Growth mindset in rekenen volgens leerlingen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023*

Stellingen	bao <i>n=3112-3203</i>	sbo <i>n=906-957</i>	so4 <i>n=480-495</i>
b. Hoe goed je bent in rekenen ligt vast en je kunt niet veel doen om dat te veranderen	2,8 (1,0)	2,3 (1,0)	2,5 (1,1)
d. Het maakt niet uit hoe slim je bent, als je wil kun je altijd goed worden in rekenen	3,4 (0,8)	3,5 (0,8)	3,3 (0,9)
e. Je kunt niet veranderen hoe slim je in rekenen bent	3,0 (1,0)	2,5 (1,2)	2,8 (1,2)
g. Goed kunnen rekenen is iets waar je mee geboren wordt, je kunt niet veel doen om jezelf beter in rekenen te maken	3,3 (0,9)	2,9 (1,1)	3,1 (1,1)
h. Iedereen kan goed leren rekenen	3,4 (0,8)	3,4 (0,9)	3,3 (0,9)
i. Je kunt zelf veranderen hoe goed je kunt rekenen.	3,4 (0,8)	3,4 (0,8)	3,3 (0,9)
Construct Growth mindset in rekenen	3,2 (0,6)	3,0 (0,6)	3,1 (0,6)

Vraag 2.5

d, h, i: 1 = zeer mee oneens, 4= zeer mee eens

b, e, g: 1= zeer mee eens, 4= zeer mee oneens

8.2.4 Domeinspecifieke onderwijskenmerken

Deze paragraaf geeft de uitkomsten weer van een aantal domeinspecifieke onderwijskenmerken (ofwel kenmerken van het onderwijsleerproces) vanuit de perceptie van de leerling. De uitkomsten worden weergegeven in percentages of waar van toepassing, in gemiddelden met de bijbehorende standaarddeviaties.

Tabel 8.9 – Zelfstandig werken tijdens de rekenles volgens leerlingen, in percentages, Peil.Rekenen-wiskunde PO 2023

Frequentie	% bao n=3231	% sbo n=979	% so4 n=499
Nooit	0,5	1,5	1,2
Sommige lessen	14,8	21,0	15,6
Helpt van de lessen	27,4	25,2	24,4
Elke les	57,3	52,2	58,7

Vraag 2.3

Tabel 8.10 – Helderheid rekeninstructie volgens leerlingen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Stellingen	bao n=3167-3229	sbo n=946-973	so4 n=485-499
a. Ik weet precies wat ik van mijn juf of meester moet doen	3,4 (0,7)	3,3 (0,7)	3,2 (0,7)
b. Ik kan mijn juf of meester goed begrijpen	3,5 (0,6)	3,4 (0,7)	3,3 (0,7)
c. Mijn juf of meester geeft duidelijk antwoord op mijn vragen	3,5 (0,7)	3,5 (0,7)	3,3 (0,8)
d. Mijn juf of meester kan rekenen goed uitleggen	3,7 (0,6)	3,6 (0,7)	3,5 (0,7)
e. Mijn juf of meester doet verschillende dingen om ons te helpen met leren	3,6 (0,7)	3,6 (0,7)	3,5 (0,8)
f. Als wij iets niet snappen dan legt mijn juf of meester het onderwerp nog een keer uit	3,6 (0,6)	3,6 (0,7)	3,5 (0,7)
Construct Helderheid rekeninstructie	3,5 (0,5)	3,5 (0,5)	3,4 (0,5)

Vraag 2.2

a t/m f: 1 = zeer mee oneens, 4= zeer mee eens

Tabel 8.11 – Overige stellingen rekeninstructie volgens leerlingen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Stellingen	bao n=3197-3218	sbo n=949-957	so4 n=498-499
g. Mijn juf of meester geeft commentaar op mijn rekenwerk waar ik wat aan heb	3,0 (0,9)	2,9 (1,1)	2,8 (1,0)
h. Mijn juf of meester vraagt me wel eens om te laten zien wat ik geleerd heb	2,6 (1,0)	2,9 (1,0)	2,7 (1,1)
i. Mijn juf of meester vraagt me wel eens om mijn antwoorden uit te leggen	3,2 (0,8)	3,2 (0,9)	3,1 (1,0)

Vraag 2.2

g, h, i: 1 = zeer mee oneens, 4= zeer mee eens

Tabel 8.12 – Ordelijkheid rekenles volgens leerlingen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Frequentie	bao n=3158-3212	sbo n=928-963	so4 n=487-496
a. De leerlingen luisteren niet naar wat de juf of meester zegt	2,8 (0,8)	2,7 (0,9)	2,4 (0,9)
b. Het is te onrustig voor de leerlingen om te kunnen werken	2,8 (0,8)	2,6 (0,9)	2,5 (0,9)
c. Mijn juf of meester moet een lange tijd wachten voordat de leerlingen stil zijn	2,5 (0,9)	2,4 (1,0)	2,0 (0,9)
d. Leerlingen onderbreken de juf of meester	2,7 (0,9)	2,4 (1,0)	2,1 (0,9)
e. Leerlingen houden zich niet aan de regels in de klas	3,0 (0,8)	2,7 (0,9)	2,4 (0,9)
f. Het gedrag van andere leerlingen in de klas maakt het voor mij moeilijk om me te kunnen concentreren.	2,7 (1,0)	2,5 (1,1)	2,2 (0,9)
Construct Ordelijkheid rekenles	2,7 (0,6)	2,6 (0,7)	2,2 (0,7)

Vraag 2.6

1= elke les of bijna elke les, 2= sommige lessen, 3= ongeveer de helft van de lessen, 4=nooit

8.2.5 Overzicht constructen leerlingvragenlijst

Tabel 8.13 geeft een overzicht van alle constructen die met de leerlingvragenlijst gemeten zijn voor de drie onderwijssoorten.

Tabel 8.13 – Overzicht constructen leerlingvragenlijst, Peil.Rekenen-wiskunde PO 2023

Construct	Items	Schaal-type	Label construct	Onderwijs-soort	n	Gemiddelde (sd)	Lambda-2
Plezier in rekenen	v2.1a t/m v2.1i	4-punts	plezier	bao	3064	2,6 (0,7)	0,92
				sbo	832	2,7 (0,8)	0,91
				so4	455	2,6 (0,8)	0,92
Zelfvertrouwen in rekenen	v2.4a t/m v2.4h	4-punts	zelfvertrouwen	bao	3113	2,9 (0,8)	0,92
				sbo	896	2,8 (0,7)	0,85
				so4	479	2,8 (0,7)	0,88
Attributies positieve rekenprestaties eigen inzet	v2.5a, v2.5c, v2.5f	4-punts	attributies	bao	3185	3,3 (0,6)	0,76
				sbo	949	3,5 (0,6)	0,71
				so4	486	3,3 (0,6)	0,76
Fixed/growth mindset	v2.5b, v2.5d, v2.5e, v2.5g, v2.5h, v2.5i	4-punts	mindset	bao	3112	3,2 (0,6)	0,73
				sbo	906	3,0 (0,6)	0,64
				so4	480	3,1 (0,6)	0,70
Helderheid rekeninstructie	v2.2a t/m v2.2f	4-punts	helder	bao	3167	3,5 (0,5)	0,82
				sbo	946	3,5 (0,5)	0,84
				so4	485	3,4 (0,5)	0,84
Ordelijk klasklimaat	v2.6a t/m v2.6f	4-punts	klasklimaat	bao	3158	2,7 (0,6)	0,81
				sbo	928	2,6 (0,7)	0,83
				so4	487	2,2 (0,7)	0,83

Sommige items die wel bij de betreffende hoofdvraag zijn voorgelegd, zijn geen onderdeel van het construct en zijn apart gerapporteerd (zie bijvoorbeeld Tabel 8.10 en 8.11). Dit zijn meestal items die wel relevant waren (en deels op verzoek van de Inspectie zijn toegevoegd), maar oorspronkelijk niet tot het construct behoorden. Afhankelijk van de inhoud van deze items is met factor- en betrouwbaarheidsanalyse nagegaan welke items samen de meest consistente schaal of schalen vormden.

8.3 Leerkrachtvragenlijst

8.3.1 Deelname en ontbrekende waarden

Groep 8 regulier basisonderwijs

De vragenlijst voor de leerkracht van groep 8 is door 110 leerkrachten geheel of gedeeltelijk ingevuld. Leerkrachten die meer dan de helft van de vragen niet hebben beantwoord, zijn uit het bestand verwijderd. Dit waren 6 respondenten (5,4%). In totaal hebben 104 groep 8-leerkrachten de vragenlijst voor de helft of meer beantwoord.

Tabel 8.14 – Ontbrekende waarden in leerkrachtvragenlijst baso, Peil.Rekenen-wiskunde PO 2023

Aantal ontbrekende waarden	Aantal leerkrachten
0	43
1	9
2	3
3	0
4	1
5	0
6	4
7	6
8	22
9	6
≥ 10	10
Totaal	104

Leerkrachten speciaal basisonderwijs

De leerkrachtvragenlijst voor het sbo is door 62 leerkrachten geheel of gedeeltelijk ingevuld. Leerkrachten die meer dan de helft van de vragen niet hebben beantwoord, zijn uit het bestand verwijderd. Dit waren 6 respondenten (9,6%). In totaal hebben 56 sbo-leerkrachten de vragenlijst voor de helft of meer beantwoord.

Tabel 8.15 – Ontbrekende waarden in leerkrachtvragenlijst sbo, Peil.Rekenen-wiskunde PO 2023

Aantal ontbrekende waarden	Aantal leerkrachten
0	9
1	9
2	6
3	2
4	1
5	0
6	1
7	5
8	5
9	6
≥ 10	12
Totaal	56

Leerkrachten speciaal onderwijs

De leerkrachtvragenlijst voor het so4 is door 43 leerkrachten geheel of gedeeltelijk ingevuld. Eén leerkracht heeft meer dan de helft van de vragen niet beantwoord en is uit het bestand verwijderd. In totaal hebben 42 so4-leerkrachten de vragenlijst voor de helft of meer beantwoord.

Tabel 8.16 – Ontbrekende waarden in leerkrachtvragenlijst so4, Peil.Rekenen-wiskunde PO 2023

Aantal ontbrekende waarden	Aantal leerkrachten
0	9
1	7
2	4
3	0
4	0
5	0
6	0
7	2
8	6
9	3
≥ 10	11
Totaal	42

8.3.2 Achtergrondkenmerken leerkrachten

In de volgende tabellen worden de achtergrondkenmerken van de leerkrachten van de getoetste klassen gepresenteerd. De uitkomsten worden weergegeven in percentages of in gemiddelden. Voor

deze en alle daaropvolgende tabellen geldt dat het met name het aantal sbo- en so4- leerkrachten beperkt is en daardoor de uitkomsten met voorzichtigheid geïnterpreteerd moeten worden.

Tabel 8.17 – Achtergrondkenmerken leerkrachten, in percentages, Peil.Rekenen-wiskunde PO 2023

Kenmerk	% bao	% sbo	% so4
<i>Geslacht</i>	<i>n=103</i>	<i>n=55</i>	<i>n=42</i>
Vrouw	65,0	80,0	61,9
Man	35,0	20,0	38,1
<i>Leeftijd</i>	<i>n=103</i>	<i>n=54</i>	<i>n=42</i>
Jonger dan 25	6,8	1,9	7,1
25-29	18,4	7,4	11,9
30-39	36,9	25,9	33,3
40-49	20,4	40,7	28,6
50-59	9,7	18,5	19,0
60 jaar of ouder	7,8	5,6	7,1
<i>Hoogste opleiding met diploma afgerond</i>	<i>n=103</i>	<i>n=54</i>	<i>n=42</i>
Voortgezet onderwijs of middelbaar beroepsonderwijs	1,0	0,0	4,2
Hoger beroepsonderwijs van 2-3 jaar	3,9	0,0	0,0
Hoger beroepsonderwijs (HBO) - bachelor	68,9	40,7	59,5
Hoger beroepsonderwijs (HBO) - master	21,4	53,7	26,2
Universiteit (bachelor, master of gepromoveerd)	4,9	5,6	9,5
<i>Lerarenopleiding(en) met diploma afgerond</i>	<i>n=103</i>	<i>n=54</i>	<i>n=42</i>
Nog niet afgerond/stage	8,7	1,9	16,7
PABO (4 jaar)	76,7	85,2	69,0
Academische PABO	2,9	1,9	7,1
Zij-instroom lerarenopleiding primair onderwijs (verkorte PABO)	8,7	9,3	14,3
PA, kweekschool of kleuteropleiding	5,8	1,9	0
Lerarenopleiding voortgezet onderwijs	1,0	1,9	2,4
Master Special Educational Needs	7,8	40,7	23,8
Opleiding voor onderwijsassistent	14,6	16,7	16,7
Andere lerarenopleiding	9,7	14,8	9,5
<i>Specialisatie rekenen-wiskunde tijdens opleiding</i>	<i>n=103</i>	<i>n=54</i>	<i>n=42</i>
Ja	6,8	5,6	4,8
Nee	93,2	94,4	95,2
<i>Post-HBO opleiding voor rekencoördinator</i>	<i>n=103</i>	<i>n=55</i>	<i>n=42</i>
Ja	10,7	9,1	2,4
Nee	89,3	90,9	97,6

Vraag 1.2 tot 1.7

Tabel 8.18 – Aantal jaar aan het eind van dit schooljaar lesgegeven in betreffende onderwijssoort, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

	Bao n=103	Sbo n=53	so4 n=31
Aantal jaar basisonderwijs	14,0 (10,1)		
Aantal jaar speciaal basisonderwijs		11,1 (8,6)	
Aantal jaar speciaal onderwijs			8,8 (6,9)

Vraag 1.1

Tabel 8.19 – Minimaal 1 jaar leservaring s(b)o leerkrachten in bao, sbo en/of so, in percentages, Peil.Rekenen-wiskunde PO 2023

	% sbo n=55	% so4 n=42
Basisonderwijs	61,8	45,2
Speciaal basisonderwijs	96,4	52,4
Speciaal onderwijs	9,1	71,4

Vraag 1.1

8.3.3 Schoolkenmerken (perspectief leerkracht)

Aan de leerkrachten is een aantal vragen voorgelegd over de school, zoals onderwijsconcept, klimaat en werkdruk. De uitkomsten hiervan worden in deze paragraaf getoond. De uitkomsten worden weergegeven in percentages of waar van toepassing, in gemiddelden.

Tabel 8.20 – Onderwijsconcept, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

Onderwijsconcept	% bao n=103	% sbo n=54	% so4 n=42
Traditioneel	73,8	50,0	52,4
Ontwikkelingsgericht onderwijs	7,8	22,2	14,3
Dalton	4,9	0,0	0,0
Montessori	2,9	1,9	0,0
Nieuw leren/natuurlijk leren/competentiegericht leren	1,9	0,0	2,4
Jenaplan	1,0	0,0	0,0
Vrije school	0,0	1,9	0,0
Ander onderwijsconcept	7,8	24,2	31,0

Vraag 2.1

Tabel 8.21 – Prestatiegerichtheid school volgens leerkrachten, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Typeringen eigen school	bao n=100-103	sbo n=52-55	so4 n=42
<i>Leerkrachten</i>			
a. De mate waarin leerkrachten de onderwijsdoelen van de school kennen	3,7 (0,6)	3,6 (0,7)	3,5 (0,7)
b. De mate waarin leerkrachten succesvol zijn in de realisatie van het curriculum van de school	3,7 (0,6)	3,6 (0,6)	3,3 (0,7)
c. Verwachtingen van leerkrachten wat betreft de leerresultaten van de leerlingen	3,8 (0,6)	3,5 (0,6)	3,2 (0,8)
d. De mate waarin leerkrachten in staat zijn leerlingen te inspireren	3,7 (0,6)	3,7 (0,6)	3,8 (0,6)
<i>Ouders</i>			
e. Betrokkenheid van ouders bij schoolactiviteiten	3,0 (0,7)	2,7 (0,9)	2,2 (0,8)
f. De mate waarin ouders ervoor zorgen dat hun kind in staat is deel te nemen aan het leerproces (bijv. voldoende slaap, ontbijt)	3,3 (0,9)	2,6 (0,6)	2,4 (0,8)
g. Verwachtingen van ouders wat betreft de leerresultaten van hun kind	3,8 (0,8)	3,2 (0,6)	3,0 (0,7)
h. Ondersteuning door ouders bij het leren van hun kind	3,2 (0,7)	2,4 (0,6)	2,2 (0,7)
<i>Leerlingen</i>			
i. De wil van leerlingen om goed te presteren op school	3,5 (0,7)	3,1 (0,6)	2,8 (0,7)
j. De mate waarin leerlingen in staat zijn om de leerdoelen van de school te halen	3,5 (0,7)	3,2 (0,6)	2,9 (0,7)
k. Respect van leerlingen voor medeleerlingen die zeer goed presteren	3,6 (0,7)	3,4 (0,7)	3,0 (0,7)
Construct Prestatiegerichtheid school (alle items)	3,5 (0,4)	3,2 (0,3)	2,9 (0,4)
Construct Prestatiegerichtheid school_leerkracht	3,7 (0,4)	3,6 (0,4)	3,4 (0,5)
Construct Prestatiegerichtheid school_ouders	3,3 (0,6)	2,7 (0,4)	2,4 (0,5)
Construct Prestatiegerichtheid school_leerlingen	3,5 (0,6)	3,3 (0,5)	2,9 (0,6)

Vraag 2.2

1= erg laag, 5= erg hoog

Tabel 8.22 – Aanwezigheid rekencoördinator, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

	% bao n=103	% sbo n=55	% so4 n=42
Rekencoördinator aanwezig	79,6	72,7	59,5

Vraag 2.3

Tabel 8.23 – Door leerkracht gepercipieerde ondersteuning van de rekencoördinator bij vormgeven van het rekenonderwijs, in percentage en gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

	bao n=82	sbo n=40	so4 n=25
Mate van ondersteuning	3,1 (0,8)	3,3 (0,8)	3,2 (0,7)

Vraag 2.4

1= erg laag 5= erg hoog

Tabel 8.24 – Huidige werkdruk volgens leerkrachten, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Stellingen	bao n=104	sbo n=54-55	so4 n=39-42
a. De klassen zijn te groot	2,5 (0,9)	2,7 (0,9)	2,6 (1,0)
b. Ik moet in mijn lessen te veel onderwerpen behandelen	2,6 (0,8)	2,6 (0,9)	2,4 (0,7)
c. Ik heb teveel lesuren	3,1 (0,7)	3,0 (0,8)	2,9 (0,8)
d. Ik heb meer tijd nodig voor lesvoorbereiding	2,6 (0,9)	2,4 (0,9)	2,4 (0,9)
e. Ik heb meer tijd nodig om individuele leerlingen te kunnen helpen	1,8 (0,7)	2,1 (0,8)	1,8 (0,6)
f. Ik ervaar te veel druk van ouders	2,8 (0,9)	3,2 (0,7)	3,0 (0,7)
g. Ik vind het moeilijk om alle veranderingen in het curriculum bij te houden	3,0 (0,7)	2,8 (0,8)	2,6 (0,6)
h. Ik heb teveel administratieve taken.	2,2 (0,7)	2,0 (0,8)	1,9 (0,7)
i. Ik heb te weinig tijd voor bijscholing (zoals workshops, cursussen, congressen/beurzen, bijhouden vakliteratuur)	2,5 (0,8)	2,3 (0,8)	2,2 (0,8)
j. Ik heb meer lesuren dan ik zou willen vanwege een tekort aan beschikbare leerkrachten op onze school	3,3 (0,8)	3,0 (0,8)	3,0 (0,7)
k. Bij ons op school vallen er wel eens lessen uit omdat we geen vervanging kunnen vinden als een leerkracht uitvalt	2,9 (1,0)	2,4 (0,9)	2,0 (0,9)
l. Ik heb te weinig tijd om mijn onderwijs aan te passen aan de leervorderingen van elke individuele leerling	2,2 (0,8)	2,5 (0,8)	2,2 (0,7)
Construct Werkdruk	2,6 (0,5)	2,6 (0,5)	2,4 (0,4)

Vraag 3.1

1= zeer mee eens, 4= zeer mee oneens (hoe hoger de score hoe lager de werkdruk)

8.3.4 Domeinspecifieke leerkrachtkenmerken

Deze paragraaf beschrijft de houdingen en opvattingen van de leerkracht over rekenen en wiskunde. Zelfvertrouwen en 'fixed/growth mindset' zijn gemeten aan de hand van een aantal uitspraken die aan de leerkrachten zijn voorgelegd en samen een construct vormen. De uitkomsten worden weergegeven in percentages of waar van toepassing, in gemiddelden.

Tabel 8.25 – Oordeel gebruikte reken-wiskundemethode van de leerkracht, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

In hoeverre passend bij..	bao <i>n=104</i>	sbo <i>n=54</i>	so4 <i>n=42</i>
a. uw eigen instructiemethode?	4,1 (0,6)	4,0 (0,5)	4,0 (0,6)
b. uw eigen kennis en vaardigheden?	4,2 (0,6)	4,2 (0,5)	4,2 (0,5)
c. uw overtuigingen over rekenonderwijs?	3,9 (0,9)	3,8 (0,7)	3,9 (0,5)
d. de leerlingen in uw klas?	3,9 (0,7)	3,6 (0,7)	3,8 (0,7)
Construct Oordeel geschiktheid reken-wiskundemethode	4,0 (0,6)	3,9 (0,4)	4,0 (0,4)

Vraag 4.3

1= zeer slecht, 5= zeer goed

Tabel 8.26 – Afgelopen twee jaar (bij)scholing op gebied van Rekenen-wiskunde gevolgd, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

	% bao <i>n=104</i>	% sbo <i>n=54</i>	% so4 <i>n=41</i>
a. De inhoud van het rekenonderwijs	42,3	38,9	29,3
b. Pedagogiek/didactiek van het rekenonderwijs	41,3	37,0	36,6
c. Het reken- en wiskundecurriculum (bijv. referentieniveaus)	34,6	27,8	14,6
d. Integratie van techniek in het rekenonderwijs	9,6	13,0	7,3
e. Het verbeteren van kritisch denken en de oplossingsvaardigheden van leerlingen	20,2	20,4	14,6
f. Het meten en beoordelen van leervorderingen in rekenen	20,2	29,6	17,1
g. Inspelen op de behoeften van individuele leerlingen bij rekenen	29,8	35,2	19,5

Vraag 7.1

Tabel 8.27 – Toekomstige behoefte aan (bij)scholing op gebied van Rekenen-wiskunde, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

	% bao n=104	% sbo n=51	% so4 n=40
a. De inhoud van het rekenonderwijs	39,4	43,1	30,0
b. Pedagogiek/didactiek van het rekenonderwijs	41,3	41,1	40,0
c. Het reken- en wiskundecurriculum (bijv. referentieniveaus)	33,7	29,4	27,5
d. Integratie van techniek in het rekenonderwijs	35,6	45,1	47,5
e. Het verbeteren van kritisch denken en de oplossingsvaardigheden van leerlingen	55,8	47,1	30,0
f. Het meten en beoordelen van leervorderingen in rekenen	37,5	31,4	42,5
g. Inspelen op de behoeften van individuele leerlingen bij rekenen	55,8	41,2	50,0

Vraag 7.1

Tabel 8.28 – Tijdsbesteding afgelopen twee jaar aan formele (bij)scholing op het gebied van rekenen-wiskunde, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

Tijdsbesteding	% bao n=103	% sbo n=54	% so4 n=41
Geen	25,0	25,9	39,0
Minder dan 6 uur	24,0	27,8	29,3
6 -15 uur	25,0	31,5	24,4
16-35 uur	12,5	1,9	0,0
Meer dan 35 uur	13,5	13,0	7,3

Vraag 7.2

Tabel 8.29 – Zelfvertrouwen vakdidactische vaardigheden Rekenen-wiskunde leerkracht, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Vaardigheden	bao n=101-104	sbo n=50-54	so4 n=41
a. Leerlingen inspireren om rekenen te leren	3,0 (0,6)	3,1 (0,7)	2,9 (0,7)
b. Leerlingen verschillende oplossingsstrategieën aanreiken	3,0 (0,7)	2,6 (0,8)	2,8 (0,7)
c. Uitdagende rekentaken geven aan excellente leerlingen	2,8 (0,8)	2,3 (0,8)	2,2 (0,9)
d. Mijn instructie aanpassen om de betrokkenheid van leerlingen bij de les te vergroten	3,0 (0,7)	3,2 (0,7)	3,0 (0,8)
e. Leerlingen het belang van rekenen te laten inzien	3,0 (0,7)	3,1 (0,7)	3,0 (0,6)
f. Het evalueren van het begrip van leerlingen in rekenen	2,8 (0,7)	2,6 (0,7)	2,7 (0,7)
g. Het verbeteren van het begrip van zwakke rekenaars	2,9 (0,7)	2,9 (0,8)	2,9 (0,8)
h. Leerlingen het nut van rekenen laten inzien	2,9 (0,8)	3,0 (0,7)	2,6 (0,7)
i. Het ontwikkelen van hogere denkvaardigheden bij leerlingen	2,6 (0,7)	2,5 (0,8)	2,5 (0,8)
Construct Zelfvertrouwen vakdidactische vaardigheden	2,9 (0,5)	2,8 (0,5)	2,7 (0,6)

Vraag 7.3

1= laag, 4= zeer hoog

Tabel 8.30 – Fixed/growth mindset ten opzichte van Rekenen-wiskunde van de leerkracht, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Stellingen	bao n=102	sbo n=53	so4 n=40
a. Het maakt niet uit hoe intelligent een leerling is, als leerlingen willen kunnen ze altijd goed worden in rekenen.	2,9 (0,7)	2,5 (0,8)	2,7 (0,8)
b. Hoe goed een leerling is in rekenen ligt vast en je kunt niet veel doen om dat te veranderen.	3,4 (0,7)	3,4 (0,6)	3,3 (0,6)
c. Je kunt leerlingen wel nieuwe dingen leren over rekenen, maar je kunt niet veranderen hoe slim ze in rekenen zijn.	2,8 (0,8)	2,6 (0,8)	2,6 (0,8)
d. Elke leerling kan goed leren rekenen.	2,8 (0,7)	2,3 (0,7)	2,7 (0,7)
e. Een leerling kan zelf veranderen hoe goed hij/zij kan rekenen.	2,8 (0,7)	2,6 (0,7)	2,7 (0,8)
f. Aanleg voor rekenen is iets waar een leerling mee geboren wordt, je kunt niet veel doen om ze beter in rekenen te maken dan het niveau van de aangeboren aanleg.	3,1 (0,8)	3,1 (0,7)	2,9 (0,7)
Construct Fixed/growth mindset	3,0 (0,5)	2,8 (0,4)	2,8 (0,5)

Vraag 8.4 (van fixed naar growth mindset)

a,d,e: 1= zeer mee oneens, 4= zeer mee eens

b,c,f: 1= zeer mee eens, 4= zeer mee oneens

8.3.5 Domeinspecifieke onderwijskenmerken

Deze paragraaf beschrijft kenmerken van het onderwijsleerproces van de getoetste klas. Een aantal van deze kenmerken zijn gemeten aan de hand van een aantal uitspraken die aan de leerkrachten zijn voorgelegd en samen een construct vormen. De uitkomsten worden weergegeven in percentages of waar van toepassing, in gemiddelden.

Tabel 8.31 – Minuten rekenonderwijs gemiddeld per week in de getoetste groep, Peil.Rekenen-wiskunde PO 2023

Aantal minuten per week	bao n=102	sbo n=54	so4 n=42
Gemiddelde (sd)	303,4 (37,9)	282,0 (38,7)	265,3 (37,7)
Minimum	225	165	180
Maximum	450	400	300
Modus	300	300	300

Vraag 4.1

Tabel 8.32 – Hoofdzakelijk gebruikte methode voor instructie en verwerking in het rekenonderwijs, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

Methode	% bao n=104	% sbo n=55	% so4 n=42
De Wereld in Getallen	45,2	30,9	31,0
Pluspunt	20,2	3,6	7,1
Snappet-leerlijnen	12,5	3,6	7,1
Getal en Ruimte Junior	11,5	20,0	35,7
Wizwijs	3,8	9,1	4,8
RekenRijk	2,9	0,0	2,4
Reken Zeker	1,9	7,3	9,5
Gynzy	1,0	10,9	2,4
Wis en Reken	0,0	0,0	0,0
Alles Telt /Alles TeltQ	0,0	7,3	0,0
Rekenkracht	0,0	0,0	0,0
Math Exova	0,0	0,0	0,0
Met Sprongen vooruit	0,0	0,0	0,0
Numo/Muiswerk	0,0	0,0	0,0
Montessori materiaal	0,0	0,0	0,0
Andere rekenmethode	1,0	7,3	0,0

Vraag 4.2

Tabel 8.33 – Werkvormen tijdens de rekenlessen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Werkvormen	bao n=101-104	sbo n=53-56	so4 n=42
a. Luisteren naar mijn uitleg over nieuwe leerstof voor rekenen	3,6 (0,6)	3,5 (0,7)	3,5 (0,6)
b. Luisteren naar mijn uitleg over hoe je opgaven kunt oplossen	3,6 (0,6)	3,7 (0,6)	3,7 (0,5)
c. Uit het hoofd leren van regels, procedures en feiten (zoals automatiseren en strategieën voor onder elkaar optellen)	3,2 (0,8)	3,1 (0,9)	3,1 (0,8)
d. Het zelfstandig oefenen van procedures	3,8 (0,4)	3,7 (0,5)	3,7 (0,5)
e. Het zelfstandig toepassen van het geleerde bij nieuwe probleemsituaties	3,6 (0,6)	3,4 (0,7)	3,3 (0,8)
f. Klassikaal aan opgaven werken onder mijn directe begeleiding	3,1 (0,9)	3,3 (0,8)	3,4 (0,8)
g. Werken in heterogene vaardigheidsgroepen	2,8 (0,9)	2,7 (1,1)	2,4 (1,1)
h. Werken in homogene vaardigheidsgroepen	2,7 (0,9)	2,9 (1,1)	2,4 (1,0)
i. Uitwerkingen opschrijven op het bord	3,2 (1,0)	3,4 (0,8)	3,4 (0,9)
j. Uitleggen hoe ze op een gegeven antwoord zijn gekomen	3,7 (0,5)	3,5 (0,6)	3,6 (0,6)
k. Klassikale bespreking van verschillende oplossingsstrategieën	3,2 (0,8)	2,9 (0,9)	3,3 (0,8)
Construct Leerkracht-gestuurde klassikale instructie (zonder g en h)	3,5 (0,4)	3,4 (0,4)	3,4 (0,4)

Vraag 4.4

1= (bijna) nooit, 2= sommige lessen, 3= ongeveer de helft van de lessen, 4= (bijna) alle lessen

Tabel 8.34 – Frequentie gebruik instructiemodel rekenlessen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Instructiemode	bao n=102	sbo n=55	so4 n=42
Directe instructiemodel	3,7 (0,6)	3,7 (0,5)	3,4 (0,5)
Ontdekkend of onderzoekend leren	2,0 (0,6)	2,1 (0,7)	2,0 (0,8)

Vraag 4.5

1= (bijna) nooit, 2= sommige lessen, 3= ongeveer de helft van de lessen, 4= (bijna) alle lessen

Tabel 8.35 – Aanbod oplossingsstrategieën voor rekenen, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

Aanbod	% bao n=103	% sbo n=56	% so4 n=42
ik bied meestal één oplossingsstrategie aan	17,5	42,9	28,6
Ik bied meestal meerdere rekenstrategieën aan voor:			
alleen <u>goed</u> presterende leerlingen in rekenen	16,5	14,3	19,0
alleen <u>gemiddeld</u> en <u>goed</u> presterende leerlingen in rekenen	35,9	16,1	19,0
alleen <u>zwak</u> en <u>gemiddeld</u> presterende leerlingen in rekenen	0,0	5,4	0,0
alleen <u>zwak</u> presterende leerlingen in rekenen	0,0	0,0	0,0
alle leerlingen	30,1	21,4	33,3

Vraag 4.6

Tabel 8.36 – Behandelde strategieën voor optellen, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

Strategie voor optellen - lang	% bao n=100	% sbo n=56	% so4 n=41
Ja, voor alle leerlingen	13,0	42,9	19,5
Ja, voor meeste leerlingen	3,0	8,9	7,3
Ja, voor sommige leerlingen	23,0	25,0	26,8
Nee	61,0	23,2	46,3
Strategie voor optellen - kort	n=104	n=56	n=42
Ja, voor alle leerlingen	80,8	48,2	61,9
Ja, voor meeste leerlingen	11,5	21,4	26,2
Ja, voor sommige leerlingen	5,8	17,9	11,9
Nee	1,9	12,5	0,0

Vraag 4.7

Tabel 8.37 – Behandelde strategieën voor aftrekken, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

	% bao <i>n</i> =98	% sbo <i>n</i> =56	% so4 <i>n</i> =41
Strategie voor aftrekken - lang			
Ja, voor alle leerlingen	11,2	41,1	17,1
Ja, voor meeste leerlingen	3,1	10,7	9,8
Ja, voor sommige leerlingen	18,4	19,6	29,3
Nee	67,3	28,6	43,9
Strategie voor aftrekken - kort	<i>n</i> =104	<i>n</i> =56	<i>n</i> =41
Ja, voor alle leerlingen	80,8	48,2	61,9
Ja, voor meeste leerlingen	11,5	17,9	28,6
Ja, voor sommige leerlingen	6,7	19,6	9,5
Nee	1,0	14,3	0,0

Vraag 4.7

Tabel 8.38 – Behandelde strategieën voor vermenigvuldigen, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

	% bao <i>n</i> =99	% sbo <i>n</i> =56	% so4 <i>n</i> =41
Strategie voor vermenigvuldigen - lang			
Ja, voor alle leerlingen	16,2	32,1	17,1
Ja, voor meeste leerlingen	6,1	21,4	17,1
Ja, voor sommige leerlingen	28,3	19,6	24,4
Nee	49,5	26,8	41,5
Strategie voor vermenigvuldigen - kort	<i>n</i> =104	<i>n</i> =56	<i>n</i> =42
Ja, voor alle leerlingen	81,7	35,7	57,1
Ja, voor meeste leerlingen	14,4	14,3	21,4
Ja, voor sommige leerlingen	3,8	23,2	19,0
Nee	0,0	26,8	2,4

Vraag 4.7

Tabel 8.39 – Behandelde strategieën voor delen, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

	% bao n=102	% sbo n=56	% so4 n=41
Strategie voor delen - lang			
Ja, voor alle leerlingen	58,8	41,1	51,2
Ja, voor meeste leerlingen	13,7	14,3	14,6
Ja, voor sommige leerlingen	11,8	21,4	19,5
Nee	15,7	23,2	14,6
Strategie voor delen - kort	n=102	n=56	n=42
Ja, voor alle leerlingen	43,1	19,6	28,6
Ja, voor meeste leerlingen	18,6	3,6	16,7
Ja, voor sommige leerlingen	20,6	21,4	19,0
Nee	17,6	55,4	35,7

Vraag 4.7

Tabel 8.40 – Behandelde strategieën rekenen lang of kort, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

	bao n=97-104	sbo n=56	so4 n=41
Strategieën voor optellen, aftrekken en vermenigvuldigen			
Construct Strategie Lang (kolom)	1,7 (1,0)	2,6 (1,1)	2,0 (1,1)
Construct Strategie Kort (cijferend)	3,7 (0,5)	2,9 (1,1)	3,5 (0,7)

1= nee, 4= alle leerlingen. De items over 'delen' vielen uit deze constructen o.b.v factor- en betrouwbaarheidsanalyse. De tabellen laten ook zien dat de 'lange strategie' voor delen veel vaker wordt toegepast dan de 'lange strategie' voor de overige bewerkingen. Het omgekeerde geldt voor de korte strategie.

Tabel 8.41 – Gebruik digitale apparaten in de rekenles, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

	% bao n=104	% sbo n=56	% so4 n=42
Gebruik digitale apparaten	90,4%	91,1%	83,3
Toegang	n=92-93	n=51	n=35
Elke leerling heeft een computer van de school in de klas	79,6	80,4	65,7
De klas heeft computers die de leerlingen kunnen delen	33,3	22,9	48,5
De school heeft computers die uw klas soms kan gebruiken	25,0	20,8	42,4
Leerlingen brengen hun eigen computer mee	3,2	0,0	0,0

Vraag 5.1

Vraag 5.2 aantallen gebaseerd op 'ja' vraag 5.1

Tabel 8.42 – Frequentie gebruik softwareprogramma's en/of rekenwebsites tijdens de rekenlessen, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

Frequentie	% bao n=104	% sbo n=56	% so4 n=42
nooit of bijna nooit	15,4	10,7	21,4
sommige lessen	17,3	23,2	31,0
ongeveer de helft van de lessen	6,7	21,4	16,7
alle of bijna alle lessen	60,6	44,6	31,0

Vraag 5.3a: inclusief 'nee' vraag 5.1 ('nee' omgezet naar '(bijna) nooit')

Tabel 8.43 – Softwareprogramma's en/of rekenwebsites gebruikt voor instructie en verwerking in het rekenonderwijs, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

Software/rekenwebsite	% bao n=83-88	% sbo n=45-48	% so4 n=29-32
Software behorende bij de rekenmethode	77,3	72,9	81,3
Bareka/Rekensprint	45,3	25,0	6,7
Gynzy	35,3	57,4	64,5
Snappet	33,3	17,8	10,0
Squla	31,0	56,5	61,3
Bingel	26,5	8,9	10,0
Ambrasoft /Junior Cloud Trainer	17,9	32,6	40,6
Rekentuin (Oefenweb)	14,0	14,9	24,1
Andere software of rekenwebsites	58,1	70,8	56,7

Vraag 5.3b: aantallen gebaseerd op 'ja' vraag 5.1 en exclusief (bijna) nooit vraag 5.3a

Tabel 8.44 – Leeractiviteiten waarbij digitale apparaten wordt gebruikt, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Leeractiviteiten	bao n=104	sbo n=54-55	so4 n=42
a. Het oefenen van sommen en procedures	3,6 (0,9)	3,5 (1,0)	3,1 (1,3)
b. Het oplossen van complexe problemen of problemen uit het dagelijks leven	3,2 (1,1)	2,6 (1,3)	2,4 (1,3)
c. Het maken van grafieken, tabellen of andere manieren om data te presenteren	2,4 (1,2)	1,8 (1,1)	1,8 (1,0)
d. Het spelen van games waarin rekensommen of concepten worden gebruikt	2,9 (1,1)	3,0 (1,1)	2,9 (1,2)
e. Het lezen van instructie of het kijken van instructie-video's	2,5 (1,2)	2,3 (1,2)	2,1 (1,2)
f. Het maken van rekentoetsen	2,1 (1,1)	1,9 (1,1)	1,6 (0,9)
Construct Leeractiviteiten met digitale apparaten	2,8 (0,8)	2,5 (0,8)	2,3 (0,8)

Vraag 5.4

1= (bijna) nooit, 2= een paar keer per jaar, 3= één of twee keer per maand, 4= tenminste één keer per week

T.b.v. schaalconstructie inclusief 'nee' vraag 5.1 ('nee omgezet naar '(bijna) nooit').

Tabel 8.45 – Gebruik rekenmachine tijdens de rekenles, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

	% bao n=104	% sbo n=55	% so4 n=42
Ja, is altijd toegestaan	0,0	0,0	2,4
Ja, is beperkt toegestaan	90,4	67,3	81,0
Nee	9,6	32,7	16,7

Vraag 5.5

Tabel 8.46 – Leeractiviteiten waarbij rekenmachine wordt gebruikt, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Leeractiviteiten	bao n=104	sbo n=55	so4 n=42
a. Het controleren van antwoorden	1,7 (0,6)	1,4 (0,5)	1,4 (0,6)
b. Het uitvoeren van routineberekeningen	1,7 (0,7)	1,4 (0,6)	1,8 (0,8)
c. Het oplossen van complexe vraagstukken	2,0 (0,6)	1,7 (0,8)	2,0 (0,8)
d. Het ontwikkelen van getalbegrip	1,6 (0,7)	1,3 (0,5)	1,5 (0,6)
Construct Leeractiviteiten met rekenmachine	1,7 (0,5)	1,5 (0,5)	1,7 (0,5)

Vraag 5.6

1=nooit, 2= sommige lessen, 3=ongeveer de helft van de lessen, 4=(bijna) alle lessen

Tbv schaalconstructie inclusief 'nee' vraag 5.5 (omgezet naar 'nooit').

Tabel 8.47 – Frequentie huiswerk voor rekenen, in percentages leerkrachten, Peil.Rekenen-wiskunde PO 2023

Frequentie	% bao n=104	% sbo n=55	% so4 n=42
Ik geef nooit rekenhuiswerk op	24,0	52,7	40,5
Minder dan één keer per week	15,4	27,3	28,6
1 of 2 keer per week	60,6	20,0	28,6
3 of 4 keer per week	0,0	0,0	2,4
Elke dag	0,0	0,0	0,0

Vraag 6.1

Tabel 8.48 – Feedback huiswerk, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Vormen van feedback	bao n=104	sbo n=55	so4 n=42
a. Het huiswerk corrigeren en feedback geven aan de leerlingen	2,2 (0,9)	1,7 (0,8)	1,9 (1,0)
b. Het huiswerk bespreken in de klas	2,2 (0,9)	1,6 (0,7)	1,6 (0,8)
c. Controleren of leerlingen het huiswerk gemaakt hebben	2,5 (0,9)	1,9 (1,0)	2,2 (1,0)
d. Ik laat leerlingen zelf hun huiswerk nakijken	1,8 (0,9)	1,2 (0,5)	1,3 (0,7)
Construct Feedback huiswerk	2,3 (0,8)	1,7 (0,8)	1,9 (0,8)

Vraag 6.2

1=(bijna) nooit, 2= soms, 3=(bijna) altijd

Tbv schaalconstructie inclusief 'nooit' vraag 6.1 (omgezet naar '(bijna) nooit').

Tabel 8.49 – Mate waarin vormen van evalueren van rekenprestaties en leerproces van toepassing zijn voor eigen onderwijs, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Stelling	bao n=102-103	sbo n=53	so4 n=40
a. Ik gebruik de scores en/of antwoorden op lvs-rekentoetsen om te evalueren of de leerdoelen bereikt zijn	2,7 (0,5)	2,7 (0,5)	2,6 (0,6)
b. Ik analyseer de scores en/of antwoorden op methodegebonden rekentoetsen om te evalueren of de leerdoelen van dat blok bereikt zijn	2,7 (0,5)	2,8 (0,5)	2,9 (0,4)
c. Ik evalueer regelmatig of alle leerlingen de lesdoelen bereikt hebben op basis van hun dagelijkse rekenwerk	2,8 (0,4)	2,8 (0,4)	2,6 (0,6)
d. Ik evalueer of alle leerlingen de lesdoelen bereikt hebben op basis van (informele) observaties tijdens de rekenles	2,7 (0,5)	2,7 (0,5)	2,6 (0,5)
e. Ik voer diagnostische gesprekken om te evalueren of specifieke leerlingen de lesdoelen bereikt hebben	2,1 (0,6)	2,0 (0,7)	1,8 (0,7)
f. Ik evalueer of de door mij gekozen manieren van instructie en verwerking effectief waren voor de meerderheid van de leerlingen in de klas	2,6 (0,5)	2,6 (0,5)	2,4 (0,7)
g. Ik evalueer of een specifieke manier van instructie effectief was voor specifieke leerlingen	2,4 (0,6)	2,4 (0,7)	2,2 (0,6)
Construct Evaluatie	2,6 (0,3)	2,6 (0,3)	2,4 (0,3)

Vraag 8.1

1= geen, 2= enigszins, 3= veel

Tabel 8.50 – Belang beoordelingsstrategieën voor rekenen, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Strategie	bao n=103	sbo n=53	so4 n=39-40
a. Het observeren van leerlingen als zij aan het werk zijn	2,9 (0,3)	2,9 (0,3)	2,9 (0,4)
b. Leerlingen vragen laten beantwoorden tijdens de les	2,8 (0,4)	2,8 (0,4)	2,9 (0,4)
c. Het doornemen van hun dagelijkse rekenwerk en/of huiswerk	2,6 (0,6)	2,6 (0,6)	2,5 (0,6)
d. Korte, tussentijdse schriftelijke toetsen	2,0 (0,8)	2,1 (0,6)	2,1 (0,6)
e. Uitgebreide toetsen aan het eind van een periode of blok (bijv methodegebonden toetsen of LVS-toets)	2,6 (0,6)	2,6 (0,5)	2,6 (0,6)

Vraag 8.2

1= geen, 2= enigszins, 3= veel

Tabel 8.51 – Mate waarin vormen van differentiatie van toepassing zijn voor eigen onderwijs, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Stelling	bao n=100-102	sbo n=53	so4 n=40
a. Ik hanteer verschillende doelen voor de leerlingen, afhankelijk van hun niveau	2,5 (0,6)	2,7 (0,6)	2,7 (0,5)
b. Ik pas mijn instructie aan aan de behoefte(n) van de leerlingen (niveau, tempo etc.)	2,9 (0,3)	2,9 (0,2)	2,9 (0,3)
c. Ik stem verschillende vormen van verwerking af op de behoeften van de verschillende kinderen in de klas	2,5 (0,6)	2,6 (0,5)	2,5 (0,5)
d. Ik benut de mogelijkheden die de methode biedt voor differentiatie voor zwakke rekenaars	2,7 (0,5)	2,5 (0,5)	2,6 (0,5)
e. Ik selecteer de meest belangrijke verwerkingsstof voor zeer zwakke rekenaars	2,6 (0,6)	2,7 (0,5)	2,5 (0,6)
f. Ik benut de mogelijkheden die de methode biedt voor differentiatie voor sterke rekenaars	2,7 (0,6)	2,4 (0,6)	2,2 (0,7)
g. Ik geef rekensterke kinderen regelmatig instructie of begeleiding op hun niveau, in groepsverband of individueel	2,4 (0,6)	2,5 (0,6)	2,3 (0,6)
h. Ik maak gebruik van de compacting werkwijze voor sterke rekenaars (overslaan van overbodige herhalings- en oefenstof).	2,6 (0,7)	2,3 (0,7)	2,2 (0,8)
i. Ik bied sterke rekenaars verrijkingsopdrachten	2,8 (0,4)	2,3 (0,6)	2,1 (0,7)
Construct Differentiatie	2,6 (0,3)	2,5 (0,3)	2,4 (0,3)

Vraag 8.3

1= niet, 2= enigszins, 3= zeer

8.3.6 Overzicht constructen leerkrachtvragenlijst

Een groot aantal van de (onderwijs)kenmerken is in de leerkrachtvragenlijst gemeten met een construct (samengestelde variabele) bestaande uit een set items, veelal in de vorm van stellingen die aan de respondent zijn voorgelegd. Voor de ontwikkeling van de vragenlijsten is vooral gebruikt gemaakt van bestaande constructen die eerder voor TIMSS (2015, 2019) en Peil.Rekenen-wiskunde PO 2019 zijn ontwikkeld. Voor deze constructen is nagegaan of zij een vergelijkbare betrouwbaarheid (Lambda-2) laten zien in onderzoekspopulaties (bao, so4 en sbo) van Peil.Rekenen-wiskunde PO-2023. Vanwege het beperkte aantal respondenten zijn voor de constructie van de leerkrachtschalen voor sbo en so4 de data van deze leerkrachten samengenomen.

Sommige items die wel bij de betreffende hoofdvraag zijn voorgelegd, zijn geen onderdeel het construct en zijn apart gerapporteerd (zie bijvoorbeeld Tabel 8.34). Dit zijn meestal items die wel relevant waren en daarom aan een vraag zijn toegevoegd, maar oorspronkelijk niet tot het construct behoorden. Afhankelijk van de inhoud deze items is met factor- en betrouwbaarheidsanalyse nagegaan welke items samen de meest consistente schaal of schalen vormden.

Tabel 8.52 – Overzicht constructen leerkrachtvragenlijst*, Peil.Rekenen-wiskunde PO 2023

Construct	Items	Schaal- type	Label construct	Onderwijs- soort	n	Gemiddelde (sd)	Lambda-2
Prestatiegerichtheid school (alle)	A2.2_1 t/m A2.2_11	5-punts	prestatie_sch	bao	100	3,5 (0,4)	0,86
				sbo+so4	94	3,1 (0,4)	0,78
Prestatiegerichtheid school: leerkracht	A2.2_1 t/m A2.2_4	5-punts	prestatie_lkr	bao	103	3,7 (0,4)	0,73
				sbo+so4	97	3,5 (0,5)	0,71
Prestatiegerichtheid school: ouders	A2.2_5 t/m A2.2_8	5-punts	prestatie_oud	bao	100	3,3 (0,6)	0,81
				sbo+so4	96	2,6 (0,5)	0,60
Prestatiegerichtheid school: leerling	A2.2_9 t/m A2.2_11	5-punts	prestatie_ll	bao	103	3,5 (0,6)	0,80
				sbo+so4	94	3,1 (0,6)	0,74
Werkdruk	A3.1_1 t/m A3.1_12	4-punts	werkdruk	bao	104	2,6 (0,5)	0,83
				sbo+so4	93	2,5 (0,5)	0,84
Oordeel geschiktheid reken-wiskundemethode	A4.3_1 t/m A4.3_4	5-punts	methode	bao	104	4,0 (0,6)	0,88
				sbo+so4	96	3,9 (0,4)	0,72
Werkvormen: Leerkracht- gestuurde klassikale instructie	A4.4_1 t/m A4.4_6 A4.4_9 t/m A4.4_11	4-punts	lkr gestuurd	bao	101	3,5 (0,4)	0,70
				sbo+so4	95	3,4 (0,4)	0,69
Behandelde strategieën dit schooljaar: lang (kolom)	A4.7_1, A4.7_3, A4.7_5	4-punts	strat_lang	bao	97	1,7 (1,0)	0,92
				sbo+so4	97	2,4 (1,1)	0,92
Behandelde strategieën dit schooljaar: kort (cijferend)	A4.7_2, A4.7_4, A4.7_6	4-punts	strat_kort	bao	104	3,7 (0,5)	0,91
				sbo+so4	98	3,1 (1,0)	0,94
Leeractiviteiten met digitale apparaten	A5.4_1 t/m A5.4_6	4-punts	ictleeract	bao	104	2,8 (0,8)	0,86
				sbo+so4	96	2,2 (0,8)	0,82
Leeractiviteiten met rekenmachine	A5.6_1 t/m A5.6_4	4-punts	rekenact	bao	104	1,7 (0,5)	0,82
				sbo+so4	97	1,6 (0,5)	0,70
				sbo+so4	97	1,8 (0,8)	0,91
Feedback huiswerk	A6.2_1 A6.2_2 A6.2_3	3-punts	huiswerk	bao	104	2,3 (0,8)	0,87
Vormen van evaluatie	A8.1_1 to A8.1_7	3-punts	evaluatie	bao	102	2,6 (0,3)	0,58
				sbo+so	93	2,5 (0,3)	0,72
Zelfvertrouwen vakdidac- tische vaardigheden	A7.3_1 t/m A7.3_9	4-punts	zelfvertrouwen	bao	101	2,9 (0,5)	0,84
				sbo+so4	91	2,6 (0,8)	0,91
Differentiatie	A8.3_1 t/m A8.3_9	3-punts	differentiatie	bao	100	2,6 (0,3)	0,71
				sbo+so4	92	2,5 (0,3)	0,70
Fixed/growth mindset **	A8.4_1 t/m A8.4_6	4-punts	mindset	bao	102	3,0 (0,5)	0,71
				sbo+so4	93	2,8 (0,4)	0,67

* Construct is overgenomen van TIMSS-2019/Peil RW-2019

** The Implicit Theories of Intelligence Scale for Children (Dweck, 2000)

8.4 Schoolvragenlijst

8.4.1 Deelname en ontbrekende waarden

In deze peiling is in het reguliere basisonderwijs de schoolvragenlijst van TIMSS-2023 afgenomen. Dit betekent dat een deel van deze vragenlijst betrekking heeft op groep 6 en op de natuurwetenschappelijke vakken. De sbo- en so4-versie van de schoolvragenlijst is gebaseerd op de schoolvragenlijst van TIMSS, maar uiteraard zonder vragen over natuuronderwijs en groep 6. In overleg met de Inspectie van het Onderwijs is een selectie gemaakt van vragen die het meest relevant zijn voor dit peilingsonderzoek voor zowel bao als sbo en so4.

De uitkomsten van deze vragen worden meestal weergegeven in percentages of in gemiddelden, afhankelijk van de antwoordcategorieën. Vanwege de lage respons en het lage aantal ingevulde schoolvragenlijsten moeten deze uitkomsten met de nodige voorzichtigheid gelezen worden.

Schoolvragenlijst regulier basisonderwijs

De data van de schoolvragenlijsten van de scholen die alleen met groep 6 hebben deelgenomen, zijn voor deze peiling buiten beschouwing gelaten. De schoolvragenlijst is door 61 schoolleiders (die dus deelgenomen hebben met tenminste groep 8) geheel of gedeeltelijk ingevuld. Schoolleiders die meer dan de helft van de vragen niet hebben beantwoord, zijn uit het bestand verwijderd. Dit is berekend op basis van alle gestelde vragen, dus niet op alleen de geselecteerde vragen waarover hier wordt gerapporteerd. In totaal hebben 57 respondenten de bao-schoolvragenlijst voor de helft of meer beantwoord.

Tabel 8.53 – *Ontbrekende waarden in de schoolvragenlijst bao, Peil.Rekenen-wiskunde PO 2023*

Aantal ontbrekende waarden	Aantal
0	42
1	11
2	3
≥ 3	1
Totaal	57

Schoolvragenlijst speciaal basisonderwijs

De schoolvragenlijst voor het sbo is door 33 schoolleiders geheel of gedeeltelijk ingevuld. Twee schoolleiders die (ruim) meer dan de helft van alle vragen niet hadden beantwoord, zijn uit het bestand verwijderd. In totaal hebben 31 respondenten de sbo-schoolvragenlijst voor de helft of meer beantwoord.

Tabel 8.54 – Ontbrekende waarden in de schoolvragenlijst sbo, Peil.Rekenen-wiskunde PO 2023

Aantal ontbrekende waarden	Aantal
0	15
1	12
2	2
3	0
4	1
≥ 5	1
Totaal	31

Schoolvragenlijst speciaal onderwijs

De schoolvragenlijst voor so4 is door 23 schoolleiders geheel of gedeeltelijk ingevuld. Drie schoolleiders die (ruim) meer dan de helft van alle vragen niet hadden beantwoord, zijn uit het bestand verwijderd. In totaal hebben 20 respondenten de so4-schoolvragenlijst voor de helft of meer beantwoord.

Tabel 8.55 – Ontbrekende waarden in de schoolvragenlijst so4, Peil.Rekenen-wiskunde PO 2023

Aantal ontbrekende waarden	Aantal
0	6
1	11
2	1
≥ 3	2
Totaal	20

8.4.2 Achtergrondkenmerken schoolleiders

In de volgende twee tabellen worden enkele achtergrondkenmerken van de schoolleiders weergegeven.

Tabel 8.56 – Aantal jaar functie schoolleider vervuld aan het eind van dit schooljaar, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

	bao <i>n=(55-56)</i>	sbo <i>n=30</i>	so4 <i>n=18</i>
Aantal jaar op deze school	5,5	6,1 (5,3)	3,5 (1,9)
Aantal jaar in totaal	11,3	10,1 (6,8)	6,0 (3,5)

Bao: vraag 18 en 17, s(b)o vraag 14 en 13

Tabel 8.57 – Opleiding schoolleider, in percentages, Peil.Rekenen-wiskunde PO 2023

Kenmerk	% bao	% sbo	% so4
<i>Hoogste opleiding met diploma afgerond</i>	<i>n=56</i>	<i>n=30</i>	<i>n=18</i>
Hoger beroepsonderwijs van 2-3 jaar	7,1	3,3	3,5
Hoger beroepsonderwijs of universiteit kandidaats/bachelor	67,9	73,3	50,0
Universiteit doctoraal of master	25,0	23,3	44,4
<i>Diploma/certificaat onderwijskundig leiderschap</i>	<i>n=56</i>	<i>n=30</i>	<i>n=18</i>
Ja	85,7	76,7	88,9
Nee	14,3	23,3	11,1

Bao: vraag 19 en 20, s(b)o vraag 15 en 16

8.4.3 Schoolkenmerken

Tabel 8.58 – Belemmeringen in het onderwijs volgens schoolleiders, in gemiddelden, Peil.Rekenen-wiskunde PO 2023

Mate van belemmering door gebrek aan of onvolkomenheid van:	bao <i>n=57</i>	sbo <i>n=31</i>	so4 <i>n=20</i>
<i>Algemeen</i>			
Digitale apparaten voor gebruik door leerlingen	1,4 (0,6)	1,3 (0,7)	1,4 (0,6)
Gekwalificeerde leerkrachten	1,6 (0,9)	1,8 (1,0)	2,3 (1,1)
<i>Rekenonderwijs</i>			
Leerkrachten met een specialisatie in rekenen-wiskunde	1,8 (0,9)	2,2 (0,9)	2,0 (1,0)
Software/apps voor rekenonderwijs	1,4 (0,8)	1,5 (0,8)	1,8 (0,9)
Bibliotheekmaterialen die gebruikt kunnen worden bij rekenonderwijs	1,8 (0,8)	2,1 (1,0)	2,5 (0,8)
Rekenmachines voor rekenonderwijs	1,3 (0,6)	1,5 (0,8)	1,6 (0,9)
Concrete voorwerpen of materialen die leerlingen helpen hoeveelheden en procedures te begrijpen	1,6 (0,7)	1,6 (0,8)	1,9 (0,9)

Bao: vraag 11 Ah, Aj en vraag 11 Ba tot Be, s(b)o: vraag 9 Ah, Aj en vraag 9 Ba tot Be
1= niet, 2= weinig 3= enigszins, 4= veel

Tabel 8.59 – Prestatiegerichtheid school volgens schoolleiders, in gemiddelden (sd), Peil.Rekenen-wiskunde PO 2023

Typeringen eigen school	bao <i>n=57</i>	sbo <i>n=31</i>	so4 <i>n=20</i>
<i>Leerkrachten</i>			
a. De mate waarin leerkrachten de onderwijsdoelen van de school kennen	3,6 (0,7)	3,5 (0,7)	3,6 (0,8)
b. De mate waarin leerkrachten succesvol zijn in de realisatie van het curriculum van de school	3,6 (0,7)	3,4 (0,6)	3,4 (0,6)
c. Verwachtingen van leerkrachten wat betreft de leerresultaten van de leerlingen	3,6 (0,7)	3,6 (0,7)	3,3 (0,7)
d. De mate waarin leerkrachten in staat zijn leerlingen te inspireren	3,7 (0,6)	3,7 (0,7)	3,6 (0,5)
<i>Ouders</i>			
e. Betrokkenheid van ouders bij schoolactiviteiten	3,0 (0,9)	2,4 (0,8)	2,2 (0,7)
f. De mate waarin ouders ervoor zorgen dat hun kind in staat is deel te nemen aan het leerproces (bijv. voldoende slaap, ontbijt)	3,2 (0,9)	2,7 (0,8)	2,3 (0,7)
g. Verwachtingen van ouders wat betreft de leerresultaten van hun kind	3,5 (0,8)	3,0 (0,7)	2,9 (0,6)
h. Ondersteuning door ouders bij het leren van hun kind	2,8 (0,8)	2,3 (0,7)	1,8 (0,4)
<i>Leerlingen</i>			
i. De wil van leerlingen om goed te presteren op school	3,5 (1,0)	3,2 (0,6)	2,7 (0,8)
j. De mate waarin leerlingen in staat zijn om de leerdoelen van de school te halen	3,5 (1,1)	3,0 (0,7)	2,7 (0,5)
k. Respect van leerlingen voor medeleerlingen die zeer goed presteren	3,5 (0,7)	3,3 (0,5)	2,8 (0,8)
Construct Prestatiegerichtheid school (alle items)	3,4 (0,5)	3,1 (0,4)	2,8 (0,4)
Construct Prestatiegerichtheid school_leerkracht	3,6 (0,5)	3,5 (0,6)	3,5 (0,5)
Construct Prestatiegerichtheid school_ouders	3,1 (0,7)	2,6 (0,6)	2,3 (0,5)
Construct Prestatiegerichtheid school_leerlingen	3,5 (0,7)	3,2 (0,5)	2,7 (0,5)

Bao vraag 12, s(b)o vraag 10

1= erg laag, 5= erg hoog

Tabel 8.60 – Volledige scholensluiting tijdens COVID-19 pandemie, bao, opgave schoolleider, in percentages, Peil.Rekenen-wiskunde PO 2023 (n=54-56)

Schooljaar	% niet volledig gesloten geweest	% < 1 maand	% 1 t/m 3 maanden	Meer dan 3 maanden
2019-2020	20,4	18,5	55,6	5,6%
2020-2021	17,9	28,6	46,4	7,2%
2021-2022	37,5	44,6	16,1	1,8
2022-2023	98,2	0,0	0,0	1,8

Vraag 21

Tabel 8.61 – Volledige scholensluiting tijdens COVID-19 pandemie, sbo, opgave schoolleider, in percentages, Peil.Rekenen-wiskunde PO 2023 (n=30-31)

Schooljaar	% niet volledig gesloten geweest	% < 1 maand	% 1 t/m 3 maanden
2019-2020	30,0	20,0	50,0
2020-2021	45,2	29,0	25,8
2021-2022	61,3	29,0	9,7
2022-2023	90,0	10,0	0,0

Vraag 17, > 3 maanden kwam niet voor

Tabel 8.62 – Volledige scholensluiting tijdens COVID-19 pandemie, so4, opgave schoolleider, in percentages, Peil.Rekenen-wiskunde PO 2023 (n=18)

Schooljaar	% niet volledig gesloten geweest	% < 1 maand	% 1 t/m 3 maanden
2019-2020	38,9	33,3	27,8
2020-2021	44,4	27,8	27,8
2021-2022	61,1	27,8	11,1
2022-2023	100	0,0	0,0

Vraag 17, > 3 maanden kwam niet voor

8.4.4 Domeinspecifieke schoolkenmerken

Tabel 8.63 – Aanwezigheid rekencoördinator/-specialist in percentages, Peil.Rekenen-wiskunde PO 2023*

	% bao (n=57)	% sbo (n=31)	% so4 (n=18)
	%	%	%
Rekencoördinator of rekenspecialist aanwezig	84,2	58,1	38,9

Bao vraag 23 (afgeleid)⁴, S(b)o vraag 18 en 19

Tabel 8.64 – Taakuren per jaar rekencoördinator/-specialist sbo, in percentages/gemiddeld (sd), Peil.Rekenen-wiskunde PO 2023*

Aantal taakuren	sbo (n=18)
Gemiddelde (sd)	44,4 (37,7)
Minimum	0
Maximum	150
Modus	40

vraag 19

* geen gegevens bao (geen data) en so4 (n=5)

4 In de TIMSS-versie van de schoolvragenlijst is de vraag naar wel /geen rekencoördinator niet gesteld. Als alle genoemde taken met 'nee' zijn beantwoord is er van uitgegaan dat er geen rekencoördinator is.

Tabel 8.65 – Taken rekencoördinator, in aantallen, Peil.Rekenen-wiskunde PO 2023

Taak	bao (n=48)	sbo (n=18)	so (n=7)
<i>Kwaliteitsbewaking van het reken-wiskundeonderwijs op schoolniveau</i>			
a. Begeleiden en aansturen aanschaf van nieuwe rekenmaterialen	44	18	7
b. Beheer en verdeling van de rekenmaterialen	36	14	6
c. Bijdragen aan de ontwikkeling van rijke leeromgevingen voor rekenen-wiskunde op school	45	18	7
d. Regelmatige evaluatie van het gebruik en de opbrengsten van de rekenmethode	43	14	6
e. Bijhouden ontwikkelingen in het reken-wiskundeonderwijs (vakliteratuur, conferenties, studiedagen, etc.)	43	18	5
f. Nieuwe ontwikkelingen reken-wiskundeonderwijs overbrengen aan het team	46	18	5
g. Onderhouden van externe contacten en participeren in netwerken	36	11	2
h. Verstrekken van informatie aan ouders over de invulling van het rekenonderwijs	20	5	1
<i>Zorgniveau</i>			
a. Analyseren toetsgegevens LVS	20	6	3
b. Hulp aan collega's bij het opstellen van handelingsplannen	20	10	4
c. Rekenrapportages van onderwijsinspectie analyseren	15	5	2
<i>Leerkrachtniveau</i>			
a. Collegiale consultatie	36	11	7
b. Klasbezoek (zoals lesobservaties) en nabespreking met leerkracht	32	11	3
c. Begeleiden van nieuwe leerkrachten	22	8	3

Bao vraag 23, s(b)o vraag 20

8.4.5 Overzicht constructen schoolvragenlijst

*Tabel 8.66 - Construct prestatiegerichtheid, volgens schoolleiders, Peil.Rekenen-wiskunde PO 2023**

Construct	Items	Schaal-type	Label construct	n	Gemiddelde (sd)	Lambda-2
Prestatiegerichtheid school (alle items)	11	5-punts	prestatie_sch	108	3,2 (0,5)	0,85
Prestatiegerichtheid school: leerkracht	4	5-punts	prestatie_lkr	108	3,6 (0,5)	0,83
Prestatiegerichtheid school: ouders	4	5-punts	prestatie_oud	108	2,8 (0,7)	0,87
Prestatiegerichtheid school: leerling	3	5-punts	prestatie_ll	108	3,2 (0,6)	0,63

* Vanwege de lage aantallen berekend op gecombineerde data van bao, sbo en so

9. Samenhang kenmerken leerlingen, leerkrachten, scholen en onderwijsleerproces met leerprestaties

9.1 Inleiding

Peilingsonderzoek biedt de mogelijkheid tot het uitvoeren van relationele analyses omdat de gegevens van leerlingen, leerkrachten en scholen aan elkaar gekoppeld zijn. In deze analyses proberen we verschillen in toetsresultaten te verklaren aan de hand van enerzijds algemene achtergrondkenmerken en anderzijds onderwijsprocesvariabelen (domeinspecifieke kenmerken). Met behulp van regressieanalyses wordt onderzocht in hoeverre de rekenprestaties van de leerlingen (de afhankelijke variabele) te verklaren zijn vanuit de kenmerken (voorspellers) op leerling-, klas- en schoolniveau.

Leerlingen zijn genest binnen klassen en scholen, wat tot een additionele afhankelijkheid van de waarnemingen leidt. Het negeren van deze onderlinge afhankelijkheid van de waarnemingen zou tot een onderschatting van de standaardfouten van de populatieschattingen kunnen leiden. Een multilevel analyse houdt rekening met deze clustering van de data en daarom maken we gebruik van multilevel-regressie analyse (Hox, Moerbeek, en Van de Schoot 2017) in deze relationele analyses. De analyses worden apart uitgevoerd voor de drie betrokken onderwijstypen (bao, sbo en so4). Op deze manier kunnen we de relaties tussen de voorspellers en de afhankelijke variabele apart bepalen binnen elk schooltype.

Voor de multilevel analyse zijn de ontbrekende waarden in de achtergrond variabelen geïmputeerd door middel van *Multivariate Imputations by Chained Equations* (Van Buuren en Groothuis-Oudshoorn 2011). Als afhankelijke variabele zijn *plausible values* gebruikt. Van zowel de imputaties als de plausible values zijn honderd trekkingen gebruikt.

In dit hoofdstuk gaan we in op de spreiding van de toetsresultaten en meer specifiek op het deel van de spreiding dat toe te schrijven is aan de leerlingen en aan de klassen/scholen. Daarna gaan we achtereenvolgens in op de leerling-, leerkracht- (inclusief onderwijs-) en schoolkenmerken en hoe deze zich verhouden tot de toetsresultaten.

9.2 Interpretatie van effecten in multilevel modellen

De multilevel-modellen worden afzonderlijk geschat en gepresenteerd voor elk schooltype. De effecten⁵ worden telkens gegeven op dezelfde vaardigheidsschaal die ook in hoofdstuk 7 wordt gebruikt voor de beschrijving van de vaardigheid. Een gegeven effect in, bijvoorbeeld, het so4 is daardoor direct te vergelijken met een effect in het sbo of bao. In andere peilingen is het soms gebruikelijk om de afhankelijke variabele in elke multilevel analyse afzonderlijk te standaardiseren.

5 De uitkomsten van de multilevel analyses worden in het algemeen beschreven in termen van 'heeft een effect op'. Hiermee wordt echter geen (direct) causaal verband verondersteld; het geeft alleen weer dat beide variabelen (bijvoorbeeld zelfvertrouwen in rekenen en leerlingprestatie) samenhangen. De richting van deze samenhang en/of mogelijke wederkerigheid kan met de cross-sectionele data niet worden onderzocht.

Wij doen dit hier echter niet omdat effecten daardoor een andere betekenis zouden krijgen binnen elk schooltype.

Voor categorische variabelen worden de effecten per categorie gegeven. Een effect van bijvoorbeeld een bepaald *onderwijsconcept* is daardoor te interpreteren als het verschil op de vaardigheidsschaal tussen dat onderwijsconcept en de referentiecategorie, onder de aanname dat alle andere variabelen gelijk blijven.

Voor continue of semi-continue variabelen (dit zijn overwegend de constructen op basis van de vragenlijst) is de oorspronkelijke schaal aangehouden. Dit betekent dat de gegeven effecten gebaseerd zijn op een score van één punt extra op de gegeven schalen. In slechts enkele gevallen is een onafhankelijke variabele gestandaardiseerd, als deze in bereik te ver van de overige schalen ligt dan wel geen nulpunt heeft dat realistisch gezien kan voorkomen, zoals in het geval van *schoolweging* en *onderwijstijd*.

Er is met andere woorden gekozen voor een weergave van de effecten die zo goed mogelijk te interpreteren en onderling vergelijkbaar is.⁶

9.3 Spreiding toetsresultaten

9.3.1 Hoeveel variabiliteit in leerlingprestaties is toe te schrijven aan scholen/klassen en leerlingen?

Om inzicht te krijgen in welk deel van de spreiding in leerlingprestatie toe te schrijven is aan leerlingen en welk deel van deze prestaties toe te schrijven is aan klassen (indien sprake is van meerdere groepen 8) of scholen, maken we gebruik van multilevel analyses (Hox, Moerbeek, en Van de Schoot 2017). In een leeg multilevel model kunnen we de intra-klasse correlatiecoëfficiënt (ICC) schatten.

De intra-klasse correlatie drukt uit in hoeverre de resultaten van leerlingen binnen een school of klas dichter bij elkaar liggen dan resultaten van leerlingen die niet bij elkaar op school zitten⁷. Grofweg betekent een intra-klasse correlatie die neigt naar 0, dat resultaten van leerlingen niet structureel samenhangen met schoolkeuze en/ of schooleffectiviteit. Bij een intra-klasse correlatie die neigt naar 1 hangen leerlingresultaten volledig samen met schoolkeuze en/ of schooleffectiviteit.

De ICC kan in formulevorm uitgedrukt worden als:

$$ICC = \frac{\sigma_u^2}{(\sigma_u^2 + \sigma_e^2)}$$

waar σ_e^2 de variantie is in leerlingresultaten en σ_u^2 de variantie in gemiddelde leerlingresultaten over scholen (Koch, 2004).

6 Hier moet echter wel de kanttekening bij worden geplaatst dat het om geïdealiseerde afzonderlijke effecten gaat. Men kan een effect interpreteren als het gemiddelde verschil op de vaardigheidsschaal tussen twee groepen die exact 1 van elkaar verschillen in de onafhankelijke variabele maar die identiek zijn wat betreft alle andere variabelen in het model. In de praktijk hebben de verklarende variabelen onderling echter vaak sterke relaties. Een verandering in één van de onafhankelijke variabelen zal daardoor realistisch gezien niet in isolatie plaatsvinden. Zie voor een uitgebreidere beschrijving (Gelman, Hill, en Vehtari 2020, 133–34).

7 Meer in algemene zin is een ICC in principe te interpreteren als een correlatie tussen twee willekeurig getrokken leerlingen binnen een eenheid op hoger niveau.

Heeft een school relatief veel leerlingen met hoge resultaten en de andere school relatief veel leerlingen met lage resultaten, dan neemt σ_u^2 relatief toe en gaat de intra-klasse correlatie omhoog. De intra-klasse correlatie kan daardoor gebruikt worden om de vraag te beantwoorden of er verschillen zijn tussen scholen op het gebied van leerlingresultaten. Er zijn echter een aantal belangrijke beperkingen aan de toepassing van de intra-klasse coëfficiënt als methode om op een valide manier school effectiviteit in kaart te brengen, omdat deze: 1) ons niet vertelt waar verschillen vandaan komen, en 2) niet zonder meer rekening houdt met initiële niveauverschillen.

Binnen dit peilingsonderzoek zijn leerlingen genest binnen klassen en klassen op hun beurt weer genest binnen scholen. Soms hebben dan ook meerdere klassen vanuit een school deelgenomen aan dit onderzoek. Voor de multilevel analyses hebben we ervoor gekozen om klas als groepsvariabele te gebruiken. De belangrijkste reden voor deze keuze is dat we op deze manier het meest nauwkeurig de domeinspecifieke kenmerken in kaart kunnen brengen. In de klassen vindt immers het uiteindelijke onderwijs plaats. In Tabel 9.1 is het aantal scholen, klassen en gemiddeld aantal leerlingen per klas, uitgesplitst naar schooltype te vinden.

Tabel 9.1 – Aantallen scholen en klassen, Peil.Rekenen-wiskunde PO 2023

Schooltype	aantal scholen	aantal klassen	gem. leerlingen per klas	ICC klas	ICC school
bao	106	164	20.08	0.14	0.14
so4	28	60	8.40	0.31	0.13
sbo	49	96	10.30	0.23	0.13

Zoals in Tabel 9.1 is te zien, heeft het merendeel van de scholen met meerdere klassen aan het peilingsonderzoek deelgenomen.

In Tabel 9.2 zijn de variantiecomponenten op het leerlingniveau (σ_e^2) en het klasniveau (σ_u^2) en de ICC uitgesplitst per schooltype terug te vinden.

Tabel 9.2 – Nul model random effecten, Peil.Rekenen-wiskunde PO 2023

Random effect	sbo	so4	bao
Residual	0.71	0.89	0.82
school_klas	0.21	0.41	0.13
ICC	0.23	0.31	0.14

In Tabel 9.2 is te zien dat het aandeel van de totale variantie die aan het groepslidmaatschap toe te schrijven is veel hoger is in het sbo en met name in so4, in vergelijking met het bao. Waar in het bao de waarde van ICC 0.14 is, is deze in het sbo 0.23 en in het so4 zelfs 0.31. Dit wijst erop dat leerlingen op sbo- en so4-scholen, meer dan in het bao, op basis van vaardigheid of aan vaardigheid gerelateerde kenmerken in verschillende klassen worden ingedeeld.

In het vervolg van dit hoofdstuk wordt de invloed van de leerling, docent en schoolvariabelen op de uitkomsten beschreven. Daartoe wordt telkens eerst een kort overzicht van de verdeling van de achtergrondvariabelen gegeven, uitgesplitst naar schooltype. Ontbrekende waardes zijn hierin geïmputeerd met behulp van de overige antwoorden van de leerlingen. Leerlingen die uit de toetsdata zijn weggelaten vanwege te veel ontbrekende antwoorden zijn ook uit de beschrijvende tabellen weggelaten (zie hoofdstuk 6). Leerkrachten die niet gekoppeld konden worden aan een klas zijn ook uit de betreffende tabellen weggelaten. Dit betekent dat de uitkomsten van de beschrijvende tabellen in dit hoofdstuk enigszins (kunnen) afwijken van de tabellen die in hoofdstuk 8 zijn gepresenteerd. De tabellen in hoofdstuk 8 zijn gebaseerd op de oorspronkelijke, afzonderlijke datasets voor leerlingen, leerkrachten en scholen.

9.4 Leerling

Tabel 9.3 geeft de beschrijvende statistieken van de - gedeeltelijk geïmputeerde - leerlingkenmerken die worden gebruikt in de multilevel analyse. De statistieken in de beschrijvende tabellen zijn uitgesplitst naar schooltype.

In de tabel is te zien dat het gemiddeld aantal boeken thuis - de proxy voor het opleidingsniveau voor ouders - in het bao hoger is dan het sbo en so4. Heel opvallend is het veel grotere aandeel jongens in so4, maar ook in het sbo is het aandeel jongens groter dan in het bao. Het aandeel leerlingen dat vertraagd is, is in het bao een stuk kleiner dan in het sbo en so4.

De verschillen in de domeinspecifieke leerlingkenmerken tussen de schooltypes zijn klein. Voor de multilevel analyse worden de antwoordcategorieën *nooit* en *sommige lessen* uit de vraag over zelfstandig werken in de klas samengevoegd.

Tabel 9.3 – Leerlingkenmerken in de multilevel analyse, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde	sbo (938)	so4 (485)	bao (3187)
boeken	0-10	350 	155 	460 
	11-25	339 	157 	1010 
	26-100	165 	111 	1086 
	101-200	34 	26 	378 
	200+	50 	36 	253 
geslacht	meisje	331 	69 	1556 
	jongens	607 	416 	1631 
thuis taal	nooit	33 	13 	69 
	soms	239 	119 	470 
	bijna altijd	158 	111 	450 
	altijd	507 	241 	2197 
vertraagd	vertraagd	418 	169 	364 
	niet vertraagd	520 	316 	2823 
zelfstandig	nooit	14 	5 	17 
	sommige lessen	197 	75 	470 
	helft van de lessen	232 	119 	876 
	elke les	495 	286 	1824 
attributies	1-4	3.5 	3.3 	3.3 
helder	1-4	3.5 	3.4 	3.5 
klasklimaat	1-4	2.6 	2.3 	2.7 
mindset	1-4	3.0 	3.1 	3.2 
plezier	1-4	2.7 	2.5 	2.6 
zelfvertrouwen	1-4	2.8 	2.9 	2.9 

* Voor de categorische variabelen wordt het gemiddelde aantal leerlingen per categorie gegeven, na imputatie van ontbrekende antwoorden. De grijze balk geeft het relatieve aandeel. Voor de continue variabelen wordt het gemiddelde op de gebruikte schaal en kansdichtheid binnen het bereik van de schaal weergegeven.

9.4.1 Welke verschillen in leerlingprestaties zijn er naar leerlingkenmerken?

Voor het bao geldt dat het aantal boeken thuis positief samenhangt met de prestaties in rekenen-wiskunde (Tabel 9.4). Er is een redelijk sterk positief effect voor geslacht te zien in het so4 en sbo (jongens presteren beter dan meisjes), maar dit effect is veel minder sterk of afwezig in het bao. Het niet vertraagd zijn heeft daarentegen een positief effect in het bao maar dit effect lijkt minder sterk in het sbo en so4.

Tabel 9.4 – Multilevel analyse met leerlingkenmerken, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde		sbo		so4		bao
Categorisch							
boeken	11-25	-0.03		0.04		0.07	
	26-100	0.13		0.04		0.20	
	101-200	-0.18		-0.17		0.25	
	200+	-0.01		0.26		0.27	
geslacht	jongens	0.30		0.37		0.04	
thuis taal	soms	-0.12		0.19		-0.10	
	bijna altijd	0.08		0.44		0.04	
	altijd	-0.03		0.31		0.08	
vertraagd	niet vertraagd	0.06		0.20		0.30	
zelfstandig	helft van de lessen	0.17		0.33		0.12	
	elke les	0.25		0.41		0.16	
Continu							
attributies	1-4	-0.07		-0.15		-0.13	
helder	1-4	-0.11		-0.07		0.06	
klasklimaat	1-4	0.14		0.05		-0.02	
mindset	1-4	0.17		0.22		0.06	
plezier	1-4	-0.04		0.12		-0.05	
zelfvertrouwen	1-4	0.41		0.40		0.69	
Random effecten							
Residual		0.57		0.67		0.49	
school_klas		0.15		0.34		0.08	
R ² fixed		0.20		0.21		0.38	
R ² totaal		0.37		0.48		0.47	

* Het 95% betrouwbaarheidsinterval van het geschatte effect is grafisch weergegeven.

Van de domeinspecifieke leerlingkenmerken springt het hebben van zelfvertrouwen er uit. Dit heeft in alle drie de schooltypes een (middel)groot positief effect op de toetsresultaten. De richting van dit effect is overigens niet uit de analyse af te leiden. Het is dus evenwel mogelijk dat zelfvertrouwen een gevolg is van betere prestaties als dat het omgekeerde het geval is. Opvallend is ook dat het effect van plezier in rekenen zeer klein is.

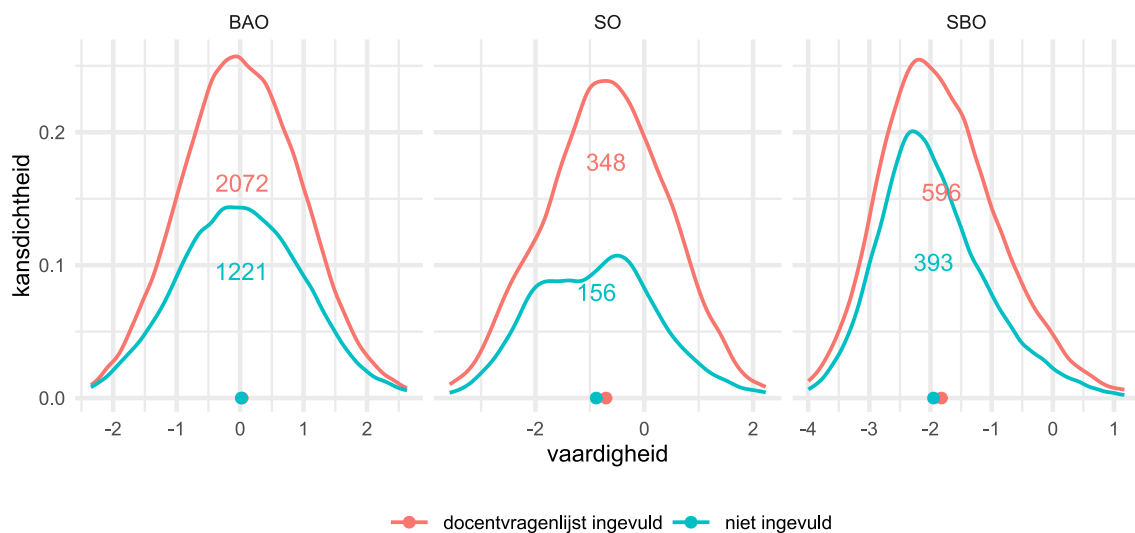
Meer zelfstandig werken lijkt daarnaast een positief effect te hebben. Ook hier zijn, net als bij zelfvertrouwen, twee richtingen voor het verband denkbaar. Leerlingen die beter zijn in rekenen mogen wellicht juist daarom vaker zelfstandig werken.

9.5 Leerkracht

De variabelen die worden gemeten door middel van de leerkrachtvragenlijst zijn in te delen in drie domeinen: algemene leerkrachtkenmerken, domeinspecifieke leerkrachtkenmerken en domeinspecifieke onderwijskenmerken. Binnen deze drie domeinen zijn er enkele variabelen binnen domeinspecifieke onderwijskenmerken die in zekere mate afhankelijk zijn van de school en niet of minder van de specifieke leerkracht, zoals de aanwezigheid van een rekencoördinator, methodetype en gebruik van digitale apparaten. Het merendeel van de variabelen heeft echter in grote mate betrekking op de leerkracht.

Het is belangrijk om te beseffen dat, ten tijde van de meting, verreweg de meeste van de deelnemende leerlingen minder dan één schooljaar rekenonderwijs hebben genoten van de bevroegde leerkracht. De vraag kan dan gesteld worden hoe groot de *a priori* te verwachten invloed van de huidige leerkracht op de relatieve prestaties van de leerlingen maximaal kan zijn. Aan de mogelijke invloed kan een bovengrens worden gesteld door te kijken naar de relatie tussen rekenprestaties aan het begin van groep 8 en het einde van groep 8. In de wetenschappelijke verantwoording van de Eindtoets Basisonderwijs 2010 (Van Boxtel, Engelen en Wijs, 2011) wordt een correlatie gegeven tussen de score op Rekenen-wiskunde in de Entreetoets 2009 en Rekenen-wiskunde in de Eindtoets 2010. Tussen de afname van deze toetsen ligt een periode van vijf à tien maanden die voor de meeste leerlingen de zomervakantie en de eerste helft van leerjaar 8 omvat. De steekproef bestond uit leerlingen die zowel de Entreetoets als Eindtoets hadden gemaakt en naar verwachting niet zouden doorstromen naar praktijkonderwijs of speciaal onderwijs. De correlatie tussen de toetsscores wordt gegeven als 0.84. De correlatie tussen de latente vaardigheden zal hoger liggen vanwege meetfout in beide toetsen. De betrouwbaarheid van het onderdeel Rekenen-wiskunde in de eindtoets wordt geschat op 0.92. De betrouwbaarheid van de Entreetoets wordt niet gerapporteerd. Als we aannemen dat deze gelijk is aan die van de Eindtoets dan is de geattenuëerde correlatie tussen de vaardigheden op de twee momenten gelijk aan $0.84/0.92 = 0.91$. Dit is een schatting van de latente correlatie tussen de vaardigheid op de twee momenten. De verklaarde variantie is hierdoor ongeveer 83% waardoor er nog maximaal 17% kan worden verklaard door genoten onderwijs van de betreffende leerkracht. Naar verwachting zullen effecten die kunnen worden toegeschreven aan de leerkracht dus klein zijn. Om een klein effect betrouwbaar vast te stellen is een grote steekproef nodig. De steekproef omvat weliswaar een vrij groot aantal leerlingen maar slechts een relatief gering aantal leerkrachten. Het is dan ook niet te verwachten dat deze effecten betrouwbaar gemeten kunnen worden. Dit betekent dat de betrouwbaarheidsintervallen breed zullen zijn in vergelijking tot de effectgroottes.

In totaal hebben 197 leerkrachten een vragenlijst ingevuld waarvoor een koppeling kon worden gemaakt met leerlingen in een klas. Voor 3016 van de 4786 leerlingen (63%) is een leerkrachtvragenlijst ingevuld. Het wel of niet invullen van de leerkrachtvragenlijst lijkt geen sterke relatie te hebben met de vaardigheid in het bao. Leerlingen in de steekproeven voor sbo en so4 voor wie de leerkracht een vragenlijst heeft ingevuld hebben gemiddeld een iets hogere vaardigheid dan de leerlingen waarvoor de leerkrachtvragenlijst niet is ingevuld, zie ook Figuur 9.1.



Figuur 9.1 – Vaardigheidsverdeling in de steekproef voor leerlingen waarvoor de leerkracht al dan niet de vragenlijst heeft ingevuld, Peil.Rekenen-wiskunde PO 2023. De steekproefgemiddelden worden weergegeven met een punt op de x-as. De getallen in de figuur geven het aantal leerlingen in elke categorie.

9.5.1 Welke verschillen in leerlingprestaties zijn er naar algemene leerkrachtkenmerken?

In Tabel 9.5 wordt een overzicht gegeven van de beschrijvende statistieken van de algemene kenmerken van de leerkrachten die in de multilevel analyses zijn meegenomen (gekoppeld zijn aan de leerlingen). De aantallen wijken daardoor iets af van de aantallen die in hoofdstuk 8 gerapporteerd worden.

Tabel 9.5 – Algemene kenmerken van de leerkrachten in de multilevel analyse, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde	sbo (53)	so4 (41)	ba0 (103)
geslacht	Vrouw	42	25	67
	Man	11	16	36
hoogste opleiding	Voortgezet onderwijs of middelbaar beroepsonderwijs	0	2	1
	Hoger beroepsonderwijs van 2-3 jaar	0		4
	Hoger beroepsonderwijs (HBO) - bachelor	21	24	71
	Hoger beroepsonderwijs (HBO) - master	28	11	22
	Universiteit (bachelor, master of gepromoveerd)	3	4	4
post HBO rekencoördinator	Ja	4	1	11
	Nee	49	40	92
leservaring	genormaliseerd	11.2	9.7	14.2

* Voor de categorische variabelen wordt het gemiddelde aantal leerkrachten per categorie gegeven, indien er sprake is van gedeeltelijke imputatie. De grijze balk geeft het relatieve aandeel. Voor de continue variabelen wordt het gemiddelde op de gebruikte schaal en kansdichtheid binnen het bereik van de schaal weergegeven.

Het merendeel van de leerkrachten is vrouw. Dit geldt voor alle schooltypen. Er bestaan nauwelijks verschillen in de hoogst genoten opleiding van de leerkrachten tussen de verschillende schooltypen. De meeste leerkrachten hebben een afgeronde bachelor, of masteropleiding op HBO niveau. Het gemiddeld aantal jaren leservaring is iets hoger in het ba0 (14.2) in vergelijking met het sbo (11.2) en so4 (9.7).

Voor de multilevel analyse wordt de hoogst genoten opleiding ingedeeld in twee categorieën zodat we conclusies niet hoeven te baseren op slechts enkele leerkrachten. We voegen alle categorieën tot en met HBO bachelor samen, en doen hetzelfde voor HBO master en universiteit. In Tabel 9.6 zijn de effecten van de algemene leerkrachtkenmerken op de toetsresultaten te vinden.

Tabel 9.6 – Multilevel analyse met algemene leerkrachtkenmerken, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde		sbo		so4		bao
Categorisch							
geslacht	Man	-0.05		-0.27		0.17	
hoogste opleiding	HBO master of universitair	0.08		0.11		-0.10	
post HBO rekencoördinator	Nee	0.17		0.99		-0.04	
Continu							
leservaring	genormaliseerd	0.12		-0.03		-0.08	
Random effecten							
Residual		0.77		0.95		0.81	
school_klas		0.23		0.42		0.11	
R ² fixed		0.02		0.04		0.02	
R ² totaal		0.25		0.33		0.14	

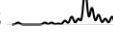
* Het 95% betrouwbaarheidsinterval van het geschatte effect is grafisch weergegeven.

De effecten van de algemene leerkrachtvariabelen zijn over het algemeen klein of hebben een breed betrouwbaarheidsinterval in vergelijking tot het geschatte effect. Mannelijke leerkrachten in het bao hebben een klein positief effect. Dit lijkt niet zo te zijn in het sbo en so4. Aangezien het hier slechts om 36 mannelijke leerkrachten gaat, is het de vraag hoe stabiel dit effect zal blijken te zijn in toekomstig onderzoek. In de voorlaatste Peiling Rekenen-wiskunde (Buisman e.a., 2021) was het effect van geslacht van de leerkracht nagenoeg 0, echter werd in die peiling geen apart model per schooltype geschat.

9.5.2 Welke verschillen in leerlingprestaties zijn er naar domeinspecifieke leerkrachtkenmerken?

In Tabel 9.7 is een overzicht te vinden van de beschrijvende statistieken van de domeinspecifieke kenmerken van de leerkrachten die in deze multilevel analyses zijn meegenomen.

Tabel 9.7 – Domeinspecifieke leerkrachtkenmerken in de multilevel analyse, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde	sbo (53)	so4 (41)	bao (103)
ervaren ondersteuning rekencoördinator	afwezig	16 	16 	21 
	erg laag	1	1 	3 
	laag	3 	1 	13 
	gemiddeld	21 	16 	34 
	hoog	9 	7 	31 
	erg hoog	3 		1
professionaliserings activiteiten	Geen	14 	16 	25 
	Minder dan 6 uur	14 	12 	25 
	6-15 uur	17 	9 	26 
	16-35 uur	1 	0	13 
	Meer dan 35 uur	7 	3 	14 
aansluiting rekenmethode	1-5	3.6 	3.7 	3.8 
mindset	1-4	2.7 	2.8 	3.0 
zelfvertrouwen	1-4	2.8 	2.7 	2.9 

* Voor de categorische variabelen wordt het gemiddelde aantal leerkrachten per categorie gegeven, indien er sprake is van gedeeltelijke imputatie. De grijze balk geeft het relatieve aandeel. Voor de continue variabelen wordt het gemiddelde op de gebruikte schaal en kansdichtheid binnen het bereik van de schaal weergegeven.

De verschillen tussen de schooltypen zijn klein met betrekking tot ervaren ondersteuning van de rekencoördinator⁸. Ook de andere verschillen zijn klein, waarbij de prestatieverwachtingen van leerlingen (fixed/growth mindset) en het zelfvertrouwen in vakdidactische vaardigheden (zeer licht) hoger zijn in het bao ten opzicht van het sbo en so4.

Voor de multilevel analyses wordt de ervaren ondersteuning van de rekencoördinator ingedeeld in vier categorieën, om zodoende zoveel mogelijk te voorkomen dat er conclusies worden getrokken op basis van één of enkele leerkrachten. Hiervoor worden respectievelijk de categorieën *erg laag* en *laag* en de categorieën *hoog* en *erg hoog* samengenomen. Voor de variabele *professionaliseringsactiviteiten* voegen we de hoogste twee categorieën samen. In Tabel 9.8 staan de effecten van de domeinspecifieke leerkrachtkenmerken op de toetsresultaten gepresenteerd.

8 Bij deze vraag is tijdens de analyse ook de categorie “afwezig” toegevoegd. Deze vraag kon namelijk worden overgeslagen als leerkrachten eerder hadden ingevuld dat er geen rekencoördinator aanwezig is.

Tabel 9.8 – Multilevel analyse met domeinspecifieke leerkrachtkenmerken, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde	sbo	so4	bao
Categorisch				
ervaren ondersteuning rekencoördinator	laag	0.19	-0.60	-0.08
	gemiddeld	0.08	-0.32	0.09
	hoog	-0.24	0.05	0.22
professionaliserings activiteiten	Minder dan 6 uur	-0.17	0.17	-0.10
	6-15 uur	0.00	-0.37	-0.14
	Meer dan 16 uur	-0.30	0.12	-0.15
Continu				
aansluiting rekenmethode	1-5	0.04	-0.41	0.05
mindset	1-4	-0.15	0.37	-0.06
zelfvertrouwen	1-4	0.29	-0.34	0.08
Random effecten				
Residual		0.77	0.95	0.81
school_klas		0.24	0.41	0.12
R ² fixed		0.05	0.09	0.02
R ² totaal		0.27	0.36	0.15

* Het 95% betrouwbaarheidsinterval van het geschatte effect is grafisch weergegeven.

De effecten van de domeinspecifieke leerkrachtvariabelen zijn over het algemeen klein of hebben een breed betrouwbaarheidsinterval in vergelijking tot het effect. De ervaren ondersteuning van de rekencoördinator lijkt een positief verband te hebben met prestaties in het so4 en het bao. Ook hier zijn echter de betrouwbaarheidsintervallen vrij breed in relatie tot het effect.

9.5.3 Welke verschillen in leerlingprestaties zijn er naar kenmerken van het reken-wiskundeonderwijs?

In Tabel 9.9 zijn de beschrijvende statistieken weergegeven van de domeinspecifieke onderwijskenmerken volgens de leerkrachten die in deze multilevel analyses zijn meegenomen.

Tabel 9.9 – Domeinspecifieke onderwijskenmerken, volgens leerkrachten in de multilevel analyse, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde	sbo (53)	so4 (41)	bao (103)
aanwezigheid rekencoördinator	Ja	37 	25 	82 
	Nee	16 	16 	21 
frequentie huiswerk	Ik geef nooit rekenhuiswerk op	27 	16 	24 
	Minder dan één keer per week	14 	12 	16 
	1 of 2 keer per week	11 	12 	63 
	3 of 4 keer per week		1 	
gebruik digitale apparaten	Ja	48 	35 	93 
	Nee	5 	6 	10 
gebruik rekenmachine	Ja, is altijd toegestaan		1 	
	Ja, is beperkt toegestaan	35 	33 	94 
	Nee	18 	7 	9 
meerdere strategieën	Nee, ik bied meestal één oplossingsstrategie aan	22 	12 	18 
	Ja, maar alleen voor de goed presterende leerlingen in rekenen	7 	7 	16 
	Ja, maar alleen voor de gemiddeld en goed presterende leerlingen in rekenen	9 	8 	37 
	Ja, maar alleen voor de zwak en gemiddeld presterende leerlingen in rekenen	3 		0
	Ja, ik bied alle leerlingen meerdere oplossingsstrategieën	12 	14 	31 

Variabele	Waarde	sbo (53)	so4 (41)	bao (103)
software/website	Alle of bijna alle lessen	22	13	63
	Ongeveer de helft van de lessen	12	7	7
	Sommige lessen	13	13	17
	Nooit of bijna nooit	6	8	16
type methode	traditioneel	15	19	14
	niet-traditioneel"	38	22	89
RW onderwijstijd	genormaliseerd	280.8	265.9	304.4
direct leren	1-4	3.7	3.7	3.7
feedback huiswerk	1-4	1.6	1.7	2.0
leeractiviteiten digitale apparaten	1-4	2.7	2.6	3.0
leerkracht gestuurd	1-4	2.9	3.0	3.0
ontdekkend leren	1-4	2.2	2.0	2.0
rekenmachine activiteiten	1-4	2.4	2.1	2.0
korte strategieën	1-4	2.9	3.4	3.7
lange strategieën	1-4	2.6	2.0	1.7

* Voor de categorische variabelen wordt het gemiddelde aantal leerkrachten per categorie gegeven, indien er sprake is van gedeeltelijke imputatie. De grijze balk geeft het relatieve aandeel. Voor de continue variabelen wordt het gemiddelde op de gebruikte schaal en kansdichtheid binnen het bereik van de schaal weergegeven.

In het bao is vaker een rekencoördinator aanwezig dan in het sbo en so4. Ook wordt er vaker huiswerk gegeven en wordt een rekenmachine en meerdere oplossingsstrategieën vaker (beperkt) toegestaan in het bao ten opzichte van het sbo en so4. Binnen het sbo en so4 wordt vaker niet-traditioneel onderwijs gegeven in vergelijking met het bao.

Voor de multilevel analyse hebben we de variabele gebruik rekenmachine gehercodeerd naar *nee* en *ja*. Bij frequentie huiswerk voor rekenen hebben we de categorie *3 of 4 keer per week* samengevoegd met *1 of 2 keer per week*. De variabele meerdere strategieën heeft ook categorieën met zeer weinig waarnemingen. Vanwege het nominale karakter van de antwoordcategorieën hebben we deze echter niet gehercodeerd. In Tabel 9.10 zijn de resultaten te vinden van de effecten van de domeinspecifieke onderwijskenmerken op de leeruitkomsten.

Tabel 9.10 - Multilevel analyse met domeinspecifieke onderwijskenmerken, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde	sbo	so4	ba0
Categorisch				
aanwezigheid rekencoördinator	Nee	0.04	0.21	-0.11
frequentie huiswerk	Minder dan één keer per week	-0.61	0.59	-0.15
	1 of 2 keer per week	-0.50	0.55	-0.04
gebruik digitale apparaten	Nee	0.23	0.24	0.09
gebruik rekenmachine	(beperkt) toegestaan	-0.66	1.35	0.33
meerdere strategieën	Ja, maar alleen voor de goed presterende leerlingen in rekenen	-0.23	0.13	0.00
	Ja, maar alleen voor de gemiddeld en goed presterende leerlingen in rekenen	0.00	0.39	0.17
	Ja, maar alleen voor de zwak en gemiddeld presterende leerlingen in rekenen	0.12		0.46
	Ja, ik bied alle leerlingen meerdere oplossingsstrategieën	0.31	0.47	0.19
software/website	Ongeveer de helft van de lessen	0.27	0.02	-0.15
	Sommige lessen	0.19	0.48	-0.00
	Nooit of bijna nooit	-0.33	0.46	-0.10
type methode	niet-traditioneel"	-0.05	-0.05	-0.12
Continu				
RW onderwijstijd	genormaliseerd	0.08	0.14	-0.07
direct leren	1-4	-0.56	-0.02	-0.03
feedback huiswerk	1-4	0.85	0.39	0.08
korte strategieën	1-4	0.06	-0.01	0.17
lange strategieën	1-4	0.05	-0.15	-0.01
leeractiviteiten digitale apparaten	1-4	-0.03	0.15	-0.03
leerkracht gestuurd	1-4	0.45	-0.16	-0.13
ontdekkend leren	1-4	0.01	-0.42	0.03
rekenmachine activiteiten	1-4	-0.38	0.33	0.17

Variabele	Waarde	sbo	so4	bao
Random effecten				
Residual		0.77	0.94	0.81
school_klas		0.15	0.27	0.13
R^2 fixed		0.16	0.20	0.03
R^2 totaal		0.29	0.38	0.17

* Het 95% betrouwbaarheidsinterval van het geschatte effect is grafisch weergegeven.

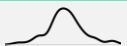
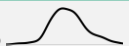
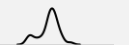
De effecten van de domeinspecifieke onderwijskenmerken op de leeruitkomsten zijn overwegend klein en/of hebben een breed betrouwbaarheidsinterval. In het sbo lijkt direct leren een vrij sterk negatief effect te hebben dat afwezig is in het sbo en bao. De gemiddelde score op dit construct is erg hoog in alle onderwijstypes; de meeste leerkrachten hebben een score van 4.

9.6 School

9.6.1 Welke verschillen in leerlingprestaties zijn er naar algemene schoolkenmerken?

In Tabel 9.11 is te zien dat (volgens opgave van de leerkracht) de meeste scholen een traditioneel onderwijsconcept hanteren. De prestatiegerichtheid van de scholen is het laagst in het sbo en het hoogst in het bao.

Tabel 9.11 – Algemene schoolkenmerken in de multilevel analyse, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde	sbo (53)	so4 (41)	bao (103)
onderwijsconcept	Dalton			5 
	Jenaplan	0		1
	Montessori	1		3 
	Ontwikkelingsgericht onderwijs	12 	6 	8 
	Traditioneel	25 	21 	76 
	Vrije school	1		0
	Nieuw leren/natuurlijk leren/competentiegericht leren	0	1 	2
	Ander onderwijsconcept	14 	13 	8 
werkdruk	1-4	2.6 	2.4 	2.6 
prestatie gerichtheid school	1-5	3.0 	2.7 	3.3 

* Voor de categorische variabelen wordt het gemiddelde aantal leerkrachten per categorie gegeven, indien er sprake is van gedeeltelijke imputatie. De grijze balk geeft het relatieve aandeel. Voor de continue variabelen wordt het gemiddelde op de gebruikte schaal en kansdichtheid binnen het bereik van de schaal weergegeven.

Voor de multilevel analyse hebben we het onderwijsconcept gehercodeerd naar *Traditioneel*, *Ontwikkelingsgericht* of *anders* om zodoende zeer kleine categorieën te vermijden. Tabel 9.12 laat zien dat het hanteren van een prestatiegericht schoolbeleid vermoedelijk maar een zeer klein effect heeft op de prestaties, wanneer gecontroleerd wordt voor schoolweging. Een opvallend resultaat is dat ervaren werkdruk in het sbo verband houdt met hogere prestaties, een effect dat we niet terug zien bij de andere schooltypes.

Tabel 9.12 – Meerniveau analyse met algemene schoolkenmerken, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde	sbo	so4	bao
Categorisch				
onderwijsconcept	Ontwikkelingsgericht onderwijs	-0.03 	-0.74 	-0.12 
	Traditioneel	0.19 	-0.05 	-0.05 
Continu				
aantal_leerlingen	genormaliseerd	0.07 	0.18 	0.03 
prestatie gerichtheid school	1-5	0.12 	0.65 	0.13 
schoolweging	genormaliseerd			-0.20 
werkdruk	1-4	0.44 	0.06 	0.04 
Random effecten				
Residual		0.77	0.95	0.80
school_klas		0.18	0.34	0.06
R ² fixed		0.07	0.08	0.07
R ² totaal		0.24	0.32	0.14

* Het 95% betrouwbaarheidsinterval van het geschatte effect is grafisch weergegeven.

9.7 Alle kenmerken in samenhang

We schatten tot slot een gecombineerd model waarbij we alle leerling-, leerkracht- en schoolvariabelen samen nemen. Het doel is om te komen tot een gecombineerd model waarin de belangrijkste verklarende variabelen zijn opgenomen. In eerste instantie wordt een multilevel model geschat met alle variabelen die hiervoor zijn besproken. Dit model is niet heel betrouwbaar te schatten vanwege het grote aantal verklarende variabelen, zeker ten opzichte van het aantal leerkrachten. Er zijn 25 variabelen gebaseerd op de leerkrachtvragenlijst opgenomen in de analyse, waar in het so4 bijvoorbeeld maar 41 leerkrachten in de analyse zijn opgenomen. Dit leidt ertoe dat de variantie van bepaalde combinaties van afhankelijke variabelen zeer laag is, waardoor de effecten van afzonderlijke variabelen niet exact vastgesteld kunnen worden. Omdat de uitkomsten van dit model misleidend kunnen zijn, geven we de uitkomsten van dit model hier niet.

We hebben vervolgens dit model gereduceerd door alleen de onafhankelijke variabelen te behouden waarvan de tweezijdige overschrijdingskans kleiner was dan 0.05. Het model is vervolgens opnieuw

geschat waarna wederom alleen de variabelen behouden zijn met een tweezijdige overschrijdingskans kleiner dan 0.05. Verdere stappen leidden niet tot een verdere reductie van variabelen.

Het is belangrijk om hierbij op te merken dat in deze procedure verklarende variabelen worden geselecteerd op basis van de overschrijdingskans (de p-waarde). Deze procedure leidt gemiddeld genomen tot een overschatting van de effecten, die in voorkomende gevallen zelfs fors kan zijn (Gelman en Carlin, 2014). Omgekeerd kunnen variabelen die uit dit model zijn weggelaten ook zeker een invloed hebben, mogelijk zelfs een relatief grote invloed. Deze is echter in het huidige onderzoek niet betrouwbaar vast te stellen. De effecten die in dit uiteindelijke model beschreven worden moeten daarom met enige voorzichtigheid worden beschouwd.

Zoals in paragraaf 9.4 is aangegeven, zijn de te verwachten effecten van leerkrachtvariabelen klein. Hierdoor is het niet verwonderlijk dat geen van deze effecten in een gecombineerd model een overschrijdingskans beneden 0.05 heeft. In het compacte model zijn dan ook geen leerkrachtkenmerken opgenomen. De verklarende variabelen bestaan uit leerlingkenmerken en, voor het bao, schoolgewicht. Het model is echter wel bepaald op de gecombineerde dataset waaruit leerlingen zijn verwijderd voor wie geen leerkrachtvragenlijst kon worden gekoppeld, wat betekent dat een aanzienlijk deel van de leerlingen is uitgesloten en de ingesloten leerlingen niet noodzakelijk representatief zijn voor de gehele steekproef (zie Figuur 9.1). We hebben daarom dit compacte model uiteindelijk nogmaals geschat op de gehele steekproef leerlingen waaruit we wel één bao-school met 47 leerlingen hebben moeten verwijderen omdat hiervoor geen informatie over de schoolweging bekend was en de variabele schoolweging wél in het compacte model is opgenomen.

Tabel 9.13 – Combinatie van kenmerken in het compacte model, Peil.Rekenen-wiskunde PO 2023*

Variabele	Waarde	sbo	so4	ba0
Categorisch				
boeken	11-25			0.05
	26-100			0.18
	101-200			0.22
	200+			0.23
geslacht	jongens	0.29	0.29	
vertraagd	niet vertraagd			0.31
zelfstandig	helft van de lessen	0.19		0.12
	elke les	0.28		0.16
Continu				
attributies	1-4		-0.12	-0.14
mindset	1-4			0.06
zelfvertrouwen	1-4	0.41	0.56	0.68
schoolweging				-0.16
Random effecten				
Residual		0.58	0.71	0.50
school_klas		0.19	0.39	0.07
R ² fixed		0.15	0.15	0.41
R ² totaal		0.36	0.45	0.48

* Het 95% betrouwbaarheidsinterval van het geschatte effect is grafisch weergegeven.

In vergelijking met het compacte model uit de peiling Rekenen-wiskunde 2019 (Buisman e.a., 2021, 182) zijn de variabelen thuistaal, rekenmethode en het docentoordeel over de passendheid van de rekenmethode in 2023 niet in het compacte model opgenomen. Zelfvertrouwen, attributies, mindset en schoolgewicht zijn zowel in 2019 als in 2023 voor één of meer van de schooltypes opgenomen in het compacte model. In 2023 zijn twee variabelen opgenomen in het compacte model die in 2019 niet zijn opgenomen. De eerste is het aantal *boeken* thuis, een proxymeting voor Sociaal-economische Status (SES). Daarnaast is in het compacte model in 2023 de variabele *vertraagd* opgenomen. In 2019 is in plaats daarvan de variabele leeftijd opgenomen voor het *ba0*, en het is mogelijk dat deze grotendeels hetzelfde meten.

9.8 Verklaarde variantie in de verschillende modellen

In Tabel 9.14 is voor het gemak nogmaals het nul model gegeven. Het nulmodel (of intercept-only model) verklaart geen variantie in de uitkomstvariabele (toetsscore). Het model splitst alleen de variantie in twee onafhankelijke componenten op: De variantie op leerlingniveau (residual) en de variantie op klasniveau (school_klas). Met name in het so4 is de spreiding op klasniveau hoog, zeker in vergelijking met het bao.

Tabel 9.14 – Nul model random effecten, Peil.Rekenen-wiskunde PO 2023*

Random effect	sbo	so4	bao
Residual	0.71	0.89	0.82
school_klas	0.21	0.41	0.13
ICC	0.23	0.31	0.14

Het leerlingmodel verklaart een gedeelte van de variantie in de toetsscores. Kijken we naar het sbo, dan zien we dat de variantie op het klasniveau gedaald is van respectievelijk 0.21 naar 0.15. Ook voor de andere schooltypes is de variantie op klasniveau gedaald.

Tabel 9.15 – Verklaarde variantie leerlingmodel, Peil.Rekenen-wiskunde PO 2023*

	sbo	so4	bao
Residual	0.57	0.67	0.49
school_klas	0.15	0.34	0.08
R2 fixed	0.20	0.21	0.38
R2 totaal	0.37	0.48	0.47

In het compacte model is voor het sbo en so4 alleen een selectie van leerlingkenmerken opgenomen. De verklaarde variantie van de fixed effecten is daardoor gedaald. In bao is ook een selectie van leerlingkenmerken opgenomen, maar daaraan is schoolgewicht toegevoegd. Daardoor neemt de verklaarde variantie van de fixed effecten licht toe in vergelijking met het leerlingmodel.

Tabel 9.16 – Verklaarde variantie compact model, Peil.Rekenen-wiskunde PO 2023*

Random effect	sbo	so4	bao
Residual	0.58	0.71	0.50
school_klas	0.19	0.39	0.07
R2 fixed	0.15	0.15	0.41
R2 totaal	0.36	0.45	0.48

10. Samenvatting

In deze samenvatting worden de belangrijkste uitkomsten beschreven aan de hand van de onderzoeksvragen die aan de peiling ten grondslag liggen.

1) Hoeveel procent van de leerlingen einde basisonderwijs en schoolverlaters speciaal basisonderwijs en speciaal onderwijs beheerst de referentieniveaus (1F en 1S) voor rekenen-wiskunde en wat is de vaardigheidsverdeling op de onderliggende schaal?

Onderstaande tabel geeft een overzicht van de percentages behaalde referentieniveaus (<1F, 1F, 1S) per schooltype.

Tabel 10.1 - Puntchattingen percentages beheersing referentieniveaus in de populatie, Peil.Rekenen-wiskunde PO 2023*

Referentieniveau	bao	sbo	so4
<1F	16,6	84,2	44,0
1F	50,6	14,1	41,4
1S	32,8	1,8	14,6

Op de onderliggende schaal is zoals verwacht de gemiddelde vaardigheid in het bao duidelijk hoger dan in het so4, en in het so4 duidelijk hoger dan in het sbo. De standaarddeviatie van de vaardigheid in het so4 is hoger dan in het sbo en bao.

2) Hoe kan de prestatie van leerlingen op de referentieniveaus worden getypeerd naar domeinen, subdomeinen, rekenonderwerpen en onderdelen?

Alle items in het peilingsonderzoek zijn op dezelfde schaal gezet met behulp van een IRT model. Voor leerlingen met referentieniveau 1S is het merendeel van de items relatief eenvoudig te beantwoorden. De prestatie op een domein of onderdeel is afhankelijk van de vaardigheid van een leerling en de relatieve moeilijkheid van de items binnen het domein of onderdeel. Van de domeinen is probleemoplossen de moeilijkste. Leerlingen onder 1F niveau kunnen gemiddeld circa 9% van de probleemoplossen items in de itembank goed beantwoorden. Leerlingen op 1F niveau kunnen gemiddeld circa 29% van deze items goed beantwoorden en leerlingen op 1S niveau circa 54%.

Voor de andere domeinen liggen deze percentages tussen 23% en 29% (kleiner dan 1F), tussen 51% en 58% (1F) en tussen 74% en 78% (2F) waarbij in alle gevallen verbanden het moeilijkst is voor de leerlingen. Een volledige beschrijving van de prestaties op alle domeinen en onderdelen wordt gegeven in hoofdstuk 7 en bijlage 2.

3) Hoe kan de prestatie van leerlingen voor rekenen met de rekenmachine worden getypeerd?

Leerlingen hebben meer moeite met de items met rekenmachine dan met de items zonder rekenmachine. Leerlingen met een referentieniveau onder 1F, 1F en 1S kunnen respectievelijk ongeveer 19%, 39% en 58% van de items met rekenmachine correct beantwoorden tegenover 27%, 55% en 76% van de items zonder rekenmachine. De schooltypes laten een vergelijkbaar patroon zien. Leerlingen op het sbo, so4 en bao kunnen respectievelijk ongeveer 19%, 32% en 43% van de items

met rekenmachine correct beantwoorden tegenover 26%, 45% en 59% van de items zonder rekenmachine.

4) Hoe kunnen per onderwijssoort de prestaties en het oplossingsproces van leerlingen voor probleemoplossen worden getypeerd?

De items binnen het domein probleemoplossen blijken erg moeilijk te zijn voor zowel leerlingen in het regulier basisonderwijs als voor leerlingen in het speciaal (basis)onderwijs. Leerlingen in basisonderwijs kunnen slechts percentage 35% van de opgaven over probleemopgaven correct beantwoorden, in so4 gaat het om 23% van de opgaven en in het sbo is dit 10%.⁹

5) Wat is de onderlinge samenhang in prestaties op domeinen, subdomeinen, rekenonderwerpen en onderdelen?

De onderlinge samenhang tussen alle domeinen en onderdelen is (zeer) hoog. De latente correlaties tussen de domeinen getallen, meten en meetkunde, verbanden, verhoudingen en probleemoplossen lagen tussen de 0,89 en 0,97. De latente correlatie tussen rekenen met rekenmachine en rekenen zonder rekenmachine was met 0,85 iets lager. Deze hoge onderliggende samenhangen tussen de domeinen en onderdelen leidde mede tot de beslissing om alle onderdelen op één schaal te kalibreren.

6) Welke verschillen zijn er tussen de leerlingprestaties einde basisonderwijs en einde speciaal basisonderwijs in 2023 en de leerlingprestaties in de peiling Rekenen-wiskunde in 2019?

De peiling uit 2019 is op papier afgenomen en de peiling in 2023 is digitaal afgenomen. Om te controleren voor een eventueel modus effect op de resultaten is een equivalentiestudie uitgevoerd als onderdeel van het proefonderzoek in 2022. Hieruit bleek dat dezelfde items digitaal minder goed worden beantwoord dan op papier. Hier is voor gecontroleerd door de papieren en digitale items in de IRT schaling als afzonderlijke items te beschouwen.

De vaardigheid voor het sbo en basisonderwijs op basis van de peiling in 2019 en het sbo, so4 en basisonderwijs in 2023 is geschat met behulp van zogenaamde *plausible values*. Bij een vergelijking van de gemiddeldes wordt de steekproeffout en de imputatiefout meegenomen in het betrouwbaarheidsinterval. Anders dan in voorgaande peilingen wordt hierbij ook rekening gehouden met de clustering van de steekproef wat leidt tot een realistische inschatting van de steekproeffout.

De verschillen tussen de gemiddelden van de peilingen in 2019 en 2023 in zowel het basisonderwijs en sbo zijn klein. De gevonden effecten zijn 0,03 (se=0,05) voor het basisonderwijs en 0,13 (se=0,09) voor het sbo. De precisie waarmee deze effecten te schatten is relatief klein ten opzichte van de gevonden effecten zodat er geen sterke conclusies aan de gevonden verschillen kunnen worden verbonden.

7) Hoe kunnen de deelnemende scholen worden getypeerd als het gaat om kenmerken van het onderwijs op het gebied van rekenen-wiskunde?

Kenmerken van het rekenonderwijs zijn vooral aan bod gekomen in de leerkrachtvragenlijst. Er zijn twee domeinspecifieke onderwijskenmerken aan schoolleiders voorgelegd: belemmeringen in het

9 Om inzicht te krijgen in de wijze waarop leerlingen tot hun antwoorden zijn gekomen (het oplossingsproces) zou procesdata nader geanalyseerd kunnen worden. Omdat in deze peiling de toets voor de eerste keer digitaal is afgenomen en vooraf nog niet bekend was hoe de afname (en daarmee ook de procesdata) in het hoofdonderzoek er uit zou zien, maakt deze analyse echter geen deel uit van dit peilingsonderzoek.

rekenonderwijs en de aanwezigheid en taken van een rekencoördinator. Schoolleiders blijken nauwelijks belemmeringen te ervaren in het onderwijs vanwege gebrek aan (geschikte) voorzieningen voor het rekenonderwijs. In het bao is op 81% van de scholen een rekencoördinator aanwezig, in het sbo 58% en so4 is dit 39%. Het aantal ingevulde schoolvragenlijsten is echter beperkt en deze uitkomsten moeten daarom met de nodige voorzichtigheid geïnterpreteerd worden.

Als we kijken naar de informatie die met de leerkrachtvragenlijst verzameld is over de rekenlessen van de getoetste klassen, dan kan het rekenonderwijs als volgt getypeerd worden:

- Volgens opgave van de leerkrachten wordt in groep 8 gemiddeld ongeveer 5 uur per week aan rekenen-wiskunde besteed. In het laatste jaar van het sbo is dit 4,7 uur en in het laatste jaar van het so4 is dit 4,4 uur per week.
- Gemiddeld genomen krijgen leerlingen in zowel groep 8 als in het laatste leerjaar s(b)o in meer dan de helft van hun lessen *leerkrachtgestuurde* klassikale instructie. Het *directe instructiemodel* wordt (gemiddeld genomen) in ruim meer dan de helft van de rekenlessen toegepast.
- Van de leerkrachten in groep 8 geeft 30% aan dat zij aan al hun leerlingen meerdere oplossingsstrategieën aanbieden en 52% dat zij dit alleen aanbieden voor goed en/of gemiddeld presterende leerlingen. Slechts 18% van de bao-leerkrachten biedt meestal maar één oplossingsstrategie voor rekenen aan. Van de sbo-leerkrachten biedt 43% hun leerlingen meestal één oplossingsstrategie aan, onder de so4-leerkrachten is dit 29%.
- De ruime meerderheid van de leerkrachten in bao en s(b)o geeft aan dat in hun rekenlessen door de leerlingen digitale apparaten worden gebruikt. Deze worden in alle drie de onderwijstypen het meest gebruikt voor het oefenen van sommen en procedures.
- In groep 8 wordt door 85% van de bao-leerkrachten aangegeven dat zij rekensoftware en/of rekenwebsites in hun lessen gebruiken. In sbo en so4 geeft 89%, respectievelijk 79% aan rekensoftware en/of rekenwebsites te gebruiken. Als er rekensoftware wordt gebruikt, dan betreft dit meestal software die bij de rekenmethode hoort.
- Verreweg de meeste leerkrachten geven aan dat het gebruik van een rekenmachine tijdens de rekenlessen beperkt is toegestaan. Van de ondervraagde sbo-leerkrachten geeft bijna een derde aan dat zij het gebruik van een rekenmachine in de les helemaal niet toestaan, in de overige twee onderwijstypen ligt dit percentage een stuk lager (bao: 10%, so4: 17%). De rekenmachine wordt in alle drie de onderwijstypes relatief het meest gebruikt voor het oplossen van complexe vraagstukken.
- In groep 8 wordt aanmerkelijk meer rekenhuiswerk gegeven dan in sbo en so4. Zo geeft 61% van de groep 8 leerkrachten één of twee keer per week rekenhuiswerk. Bijna een kwart (24%) geeft in groep 8 nooit huiswerk voor rekenen. Dit percentage ligt in het laatste leerjaar van het s(b)o aanmerkelijk hoger; in het sbo en so4 geven respectievelijk 53% en 41% van de leerkrachten hun leerlingen nooit huiswerk op voor rekenen.

8) Hoeveel variabiliteit in de leerlingprestaties is toe te schrijven aan scholen, klassen en leerlingen?

Het aandeel van de totale variantie in leerlingprestaties die aan het groepslidmaatschap is toe te schrijven is veel hoger in het s(b)o, met name in het so4, in vergelijking met het bao. Waar in het bao de waarde van de ICC relatief laag is (0,14), is deze in het sbo 0,23 en in het so4 zelfs 0,31. Dit wijst erop dat leerlingen op sbo- en so4-scholen, meer dan in het bao, op basis van vaardigheid en capaciteiten in verschillende klassen worden ingedeeld.

9) Welke verschillen in leerlingprestaties zijn er tussen leerlingen naar algemene leerlingkenmerken?

Voor het bao geldt dat het aantal boeken thuis positief samenhangt met de prestaties in rekenen-wiskunde. Er is een redelijk sterk positief effect voor geslacht te zien in het so4 en sbo (jongens presteren beter dan meisjes), maar dit effect is veel minder sterk (of afwezig) in het bao. Het niet vertraagd zijn heeft daarentegen een positief effect in het bao, maar dit effect¹⁰ lijkt minder sterk in het sbo en so4. Thuistaal heeft geen duidelijk effect.

10) Welke verschillen in leerlingprestaties zijn er tussen leerlingen naar domeinspecifieke leerlingkenmerken en domeinspecifieke onderwijskenmerken?

Van de domeinspecifieke leerlingkenmerken springt het hebben van zelfvertrouwen er uit. Dit heeft in alle drie de schooltypes een positief verband met de prestaties. De richting van dit effect is niet uit de analyses af te leiden. Mogelijk is deze relatie wederkerig, waarbij prestaties het zelfvertrouwen beïnvloeden en zelfvertrouwen de prestaties (zie bijvoorbeeld Mejía-Rodríguez et al., 2020).

Attributies van positieve prestaties aan eigen inzet heeft een klein effect dat in alle schooltypes negatief lijkt, hoewel in het so4 en sbo het 95% betrouwbaarheidsinterval ook een zeer klein positief effect omvat. Het hebben van plezier in rekenen hangt niet of nauwelijks samen met leerprestaties.

Zelfstandig werken heeft een positief verband met leerprestaties. Ook hier geldt echter dat de richting van het verband niet uit de analyse is af te leiden. Het is dus evenwel mogelijk dat zelfstandig werken leidt tot betere leerprestaties als dat de leerkracht leerlingen met een hoge vaardigheid vaker zelfstandig laat werken.

11) Welke verschillen in leerlingprestaties zijn er naar algemene leerkrachtkenmerken?

De samenhang tussen de algemene leerkrachtvariabelen en de leerlingprestaties is over het algemeen zeer beperkt, of het effect heeft een breed betrouwbaarheidsinterval in vergelijking tot het geschatte effect. Mannelijke leerkrachten in het bao lijken een klein positief effect te hebben. Dit effect is echter gebaseerd op slechts 36 mannelijke leerkrachten en hetzelfde effect komt niet naar voren in het so4 en sbo. In hoofdstuk 9 werd ook besproken dat de te verwachten effecten van leerkrachten in het laatste leerjaar noodzakelijk klein zijn, aangezien de latente correlatie tussen Rekenen-wiskunde eind groep 7 en Rekenen-wiskunde eind groep 8 op basis van ander onderzoek ongeveer 0,92 is (Van Boxtel, et al., 2011)

12) Welke verschillen in leerlingprestaties zijn er naar domeinspecifieke leerkrachtkenmerken?

De effecten van de domeinspecifieke leerkrachtvariabelen zijn over het algemeen klein of hebben een breed betrouwbaarheidsinterval in vergelijking tot het effect. De ervaren ondersteuning van de rekencoördinator lijkt een positief verband te hebben met prestaties in het so4 en het bao. Ook hier zijn echter de betrouwbaarheidsintervallen vrij breed in relatie tot het effect.

10 De uitkomsten van de multilevel-analyses worden in het algemeen beschreven in termen van 'heeft een effect op'. Hiermee wordt echter geen (direct) causaal verband verondersteld, maar geeft alleen weer dat beide variabelen samenhangen. De richting van deze samenhang, en/of mogelijke wederkerigheid kan met de cross-sectionele data niet worden onderzocht.

13) Welke verschillen in leerlingprestaties zijn er naar kenmerken van het onderwijs (domeinspecifieke onderwijskenmerken) op het gebied van rekenen-wiskunde?

De effecten van de domeinspecifieke onderwijskenmerken op de leeruitkomsten zijn overwegend klein en/of hebben een breed betrouwbaarheidsinterval. In het sbo lijkt directe instructie negatief samen te hangen met leerprestaties. Deze relatie lijkt afwezig in het so4 en bao.

14) Welke verschillen in leerlingprestaties zijn er naar algemene schoolkenmerken?

Schoolweging heeft een duidelijke negatief verband met leerprestaties. Een prestatiegericht schoolbeleid heeft (vermoedelijk) maar een zeer klein effect op de prestaties, wanneer gecontroleerd wordt voor schoolweging. Een opmerkelijke uitkomst is dat een hogere ervaren werkdruk in het sbo samen lijkt te hangen met hogere leerlingprestaties. Deze relatie zien we niet terug bij de andere schooltypes.

Referenties

- Bechger, T. M., & Maris, G. (2014). A Statistical Test for Differential Item Pair Functioning. *Psychometrika*, 80, 317–340.
- Bechger, T. M., Maris, G., Verstralen, H. F. M., & Béguin, A. (2003). Using Classical Test Theory in Combination with Item Response Theory. *Applied Psychological Measurement*, 27(5), 319–334. doi:[10.1177/0146621603257518](https://doi.org/10.1177/0146621603257518)
- Bethlehem, J. (2009). *Applied survey methods: A statistical perspective*. John Wiley & Sons.
- Birnbaum, A., Lord, F. M., & Novick, M. R. (1968). Statistical theories of mental test scores. *Some latent trait models and their use in inferring an examinee's ability*. Addison-Wesley, Reading, MA.
- Bock, R. D., & Zimowski, M. F. (1997). Multiple Group IRT. In W. J. van der Linden & R. K. Hambleton (Red.), *Handbook of Modern Item Response Theory* (pp. 433–448). New York, NY: Springer New York. doi:[10.1007/978-1-4757-2691-6_25](https://doi.org/10.1007/978-1-4757-2691-6_25)
- Buisman, M., Kuijper, S., Hickendorff, M., Kuijpers, R. E., Keuning, J., Walet, L., ... Kronenburg, F. van. (2021). *Rekenvaardigheden in het (S)BO. Technisch rapport Peil.onderwijs Rekenen-wiskunde einde (speciaal) basisonderwijs 2019*. Amsterdam: Kohnstamm Instituut.
- Eggen, T. J. H. M., & Verhelst, N. D. (2011). Item calibration in incomplete testing designs. *Psicologica*, 32(1), 107–132.
- Gelman, A., & Carlin, J. (2014). Beyond Power Calculations: Assessing Type S (Sign) and Type M (Magnitude) Errors. *Perspectives on Psychological Science*, 9(6), 641–651. doi:[10.1177/1745691614551642](https://doi.org/10.1177/1745691614551642)
- Gelman, A., & Carlin, J. (2017). Some Natural Solutions to the p-Value Communication Problem—and Why They Won't Work. *Journal of the American Statistical Association*, 112(519), 899–901.
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis, third edition*. <https://stat.columbia.edu/~gelman/book/BDA3.pdf>
- Gelman, A., Hill, J., & Vehtari, A. (2020). *Regression and Other Stories*. Cambridge University Press.
- Haelermans, C., Van der Velden, R., Aarts, B., Bijlsma, I., Jacobs, M., Smeets, C., ... Van Wetten, S. (2021). Balans na anderhalf jaar: vertraging leergroei het grootst voor rekenen-wiskunde. *NCO Factsheet*, 8.
- Hickendorff, M. (2020). *Addendum Rapport Rekenen op de Basisschool: Review van de Samenhang Tussen Beïnvloedbare Factoren in het Onderwijsleerproces en de Rekenwiskundeprestaties van Basisschoolleerlingen*. Universiteit Leiden, Instituut Pedagogische Wetenschappen, afdeling Onderwijswetenschappen.
- Hickendorff, M., Mostert, T., Van Dijk, C., Jansen, L., Van der Zee, L., & Auer, M. F. (2017). *Rekenen op de basisschool. Review van de samenhang tussen beïnvloedbare factoren in het onderwijsleerproces en de rekenwiskundeprestaties van basisschoolleerlingen*. Leiden: Universiteit Leiden.
- Hox, J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications*. Routledge.
- Koch, G. G. (2004). Intraclass correlation coefficient. *Encyclopedia of statistical sciences*.
- Köhler, C., & Hartig, J. (2017). Practical Significance of Item Misfit in Educational Assessments. *Applied Psychological Measurement*, 41(5), 388–400. doi:[10.1177/0146621617692978](https://doi.org/10.1177/0146621617692978)
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling and linking*.
- Lumley, T. (2010). *Complex Surveys: A Guide to Analysis Using R: A Guide to Analysis Using R*. John Wiley; Sons.

- Marsman, M., Maris, G., Bechger, T., & Glas, C. (2016). What can we learn from Plausible Values? *Psychometrika*, 81, 274–289.
- Meelissen, M. R. M., Maassen, N. A. M., Gubbels, J., Langen, A. M. L. van, Valk, J., Dood, D., C., Wolbers, M. (2023). *Resultaten PISA-2022 in vogelvlucht*. Universiteit Twente.
doi:[10.3990/1.9789036559461](https://doi.org/10.3990/1.9789036559461)
- Mejía-Rodríguez, A. M., Luyten, H., & Meelissen, M. R. M. (2020). Gender Differences in mathematics self-concept across the world: An exploration of student and parent data of TIMSS 2015. *International Journal of Science and Mathematics Education*. doi:[10.1007/s10763-020-10100-x](https://doi.org/10.1007/s10763-020-10100-x)
- Mullis, M., I. V.S., & Davier, M. von (Red.). (2021). *TIMSS 2023 Assessment Frameworks*. Boston College, TIMSS & PIRLS International Study Center website.
<https://timssandpirls.bc.edu/timss2023>.
- Rubin, D. B. (2018). Multiple imputation. In *Flexible Imputation of Missing Data, Second Edition* (pp. 29–62). Chapman; Hall/CRC.
- Van Boxtel, H., Engelen, R., & Wijs, A. de. (2011). *Wetenschappelijke verantwoording van de Eindtoets Basisonderwijs 2010*. Cito, Arnhem.
- Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in R. *Journal of statistical software*, 45, 1–67.
- Van Zanten, M., & Oldengram, S. (2020). *Rekenen-wiskunde in het primair onderwijs 2020. Domeinbeschrijving ten behoeve van peilingsonderzoek*. Amersfoort: SLO.
- Verhelst, N. D., Glas, C. A. W., & Verstralen, H. H. F. M. (1995). *One-parameter logistic model, OPLM*.

Bijlagen

Bijlage 1 Representativiteitsanalyse

B1.1 Methode

De toets van Peil.Rekenen-wiskunde PO 2023 is afgenomen bij 4954 leerlingen op 183 scholen. Dit is exclusief leerlingen die niets hebben ingevuld. Het aantal en percentage leerlingen per onderwijstype is in Tabel B1.1 te vinden.

Tabel B1.1 – Scholen en leerlingen in de steekproef, Peil.Rekenen-wiskunde PO 2023

Schooltype	Scholen		Leerlingen	
	aantal	percentage	aantal	percentage
bao	106	58	3.382	68
so4	28	15	525	11
sbo	49	27	1.047	21
Totaal	184	100	4.954	100

De representativiteit van de responsgegevens wordt beoordeeld door de verdeling van de relevante kenmerken in de populatie te vergelijken met de verdeling van dit kenmerk in de responsgroep (Bethlehem, 2009). Als deze twee verdelingen overlappen, kan gesteld worden dat de responsgegevens representatief zijn *met betrekking tot* dat kenmerk. De kenmerken die we zullen beoordelen op representativiteit zijn achtereenvolgens de prestatie-indicator, achterstandsscore, stedelijkheid, provincie, schoolgrootte, denominatie en geografische spreiding. Dit doen we door zowel visueel de verdeling in responsgroep en populatie met elkaar te vergelijken, als een statistische toets op het verschil tussen de twee verdelingen uit te voeren. Voor de discrete variabele is met behulp van een chi-kwadraattoets en psi-Cramer onderzocht of de verschillen tussen de onderzoeksgroep en de gehele populatie met betrekking tot diverse kenmerken significant en relevant zijn. Voor de continue variabele is de Kolmogorov-Smirnovtoets gebruikt. De maten prestatie-indicator, achterstandsscore, stedelijkheid en schoolgrootte zullen we eerst afzonderlijk toelichten.

Prestatie-indicator

De prestatie-indicator van de school is gebaseerd op de percentages behaalde referentieniveaus op de verschillende eindtoetsen sinds 2015. Hierbij hebben we het gemiddelde percentage leerlingen op de school berekend dat niveau 2F (rekenen en taal) *niet* behaalt.

Achterstandsscore

De achterstandsscore is een geaggregeerde schoolmaat voor de verwachte schoolprestatie van de leerlingen op de school gezien de omgevingskenmerken en ongeacht de intelligentie van de leerlingen. Een hogere achterstandsscore betekent dat er meer leerlingen op de school zitten die gegeven hun achterstand in aanmerking komen voor additionele financiering. In deze analyses vergelijken we de achterstandsscores met drempel van alle scholen in de populatie en de scholen in de steekproef. Voor so4- en sbo-scholen zijn geen achterstandsscores beschikbaar.

Stedelijkheid

Stedelijkheid is een derde schoolkenmerk dat we beoordelen op representativiteit. Op grond van de omgevingsadressendichtheid (het gemiddeld aantal adressen per vierkante kilometer binnen een cirkel met een straal van een kilometer) is aan iedere buurt, wijk of gemeente een stedelijkheidsklasse toegekend¹¹. De volgende klassenindeling is gehanteerd:

1. zeer sterk stedelijk ≥ 2500 adressen per km
2. sterk stedelijk 1500 - 2500 adressen per km
3. matig stedelijk 1000 - 1500 adressen per km
4. weinig stedelijk 500 - 1000 adressen per km
5. niet stedelijk < 500 adressen per km

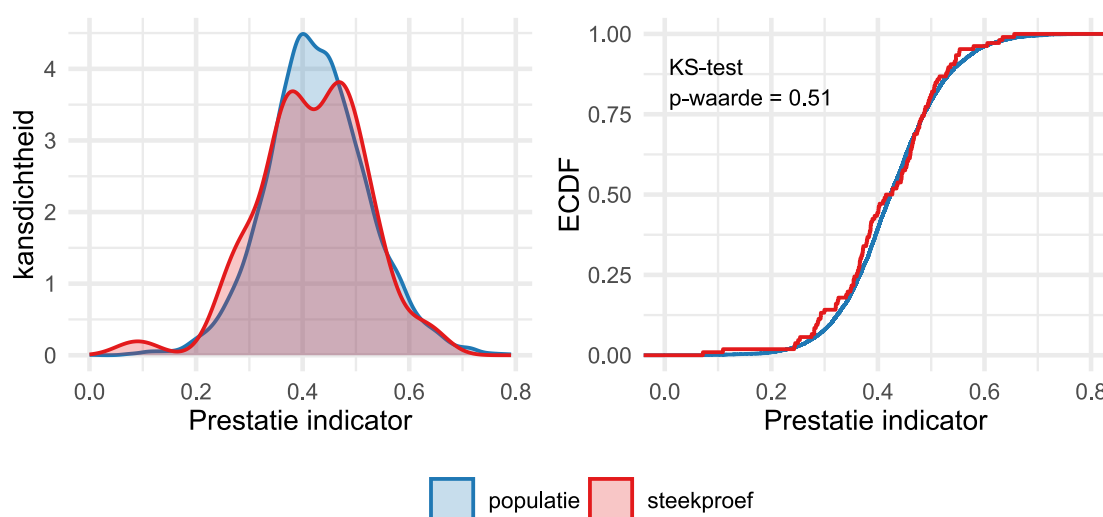
De gegevens zijn afkomstig van CBS Statline.

Schoolgrootte

De schoolgrootte van een school is bepaald door het aantal leerlingen op een school te sommeren¹². Deze gegevens zijn afkomstig van DUO.

B1.2 Regulier basisonderwijs

Prestatie-indicator bao-scholen



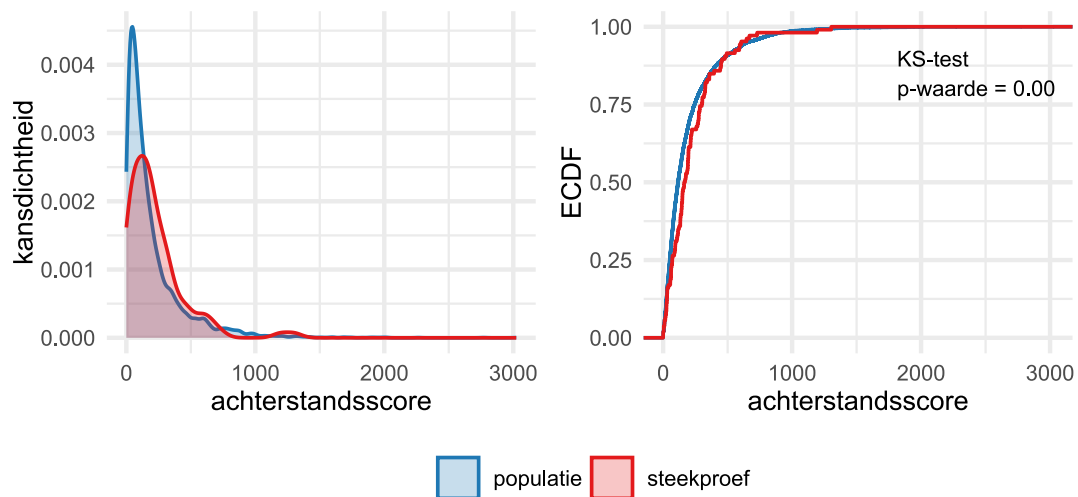
Figuur B1.1 - Prestatie-indicator bao-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

¹¹ https://www.cbs.nl/nl-nl/cijfers/detail/84583NED?q=stedelijkheid#MateVanStedelijkheid_115

¹² Binnen de analyses hebben we gebruik gemaakt van de totale schoolomvang en niet van het aantal groep 8-leerlingen op een school. We hebben hiervoor gekozen, omdat het aantal groep 8-leerlingen in een klas niet op populatieniveau bekend is. Zolang de samenhang tussen het totaal aantal leerlingen op een school en het aantal groep 8-leerlingen voldoende hoog is, zullen eventuele wegingsresultaten niet van elkaar verschillen. De gedachte dat er een grote samenhang bestaat tussen het totaal aantal leerlingen op een school en het aantal groep 8-leerlingen op diezelfde school klinkt intuïtief aannemelijk. We kunnen dit echter ook grotendeels empirisch valideren met behulp van openbaar beschikbare populatiegegevens. De toevoeging 'grotendeels' heeft betrekking op het gegeven dat het aantal leerlingen in groep 8 niet op populatieniveau bekend is (en in die zin dus ook niet als alternatief gebruikt kan worden in de analyse). Wat wel bekend is, is het aantal leerlingen per leeftijd (in jaren) op een school. De correlatie tussen het totaal aantal leerlingen op een school en het aantal 11-jarigen leerlingen op een school is 0.91. Dit is ruim voldoende evidentie om aan te nemen dat de uitkomsten van de weging niet verandert op het moment dat totale schoolgrootte wordt gebruikt. Zeker als men bedenkt dat ook het aantal 11-jarigen - een andere schatting van het onbekend aantal groep 8 leerlingen - al een proxy is.

In Figuur B1.1 is te zien dat de verdeling van de prestatie-indicator in de steekproef vrijwel overlapt met de verdeling van deze waarden van alle scholen in de populatie. Dit wordt verder onderstreept door het resultaat van de Kolmogorov-Smirnovtoets die te vinden is aan de rechterkant van de figuur. Deze statistische toets beoordeelt het verschil tussen twee verdelingen en laat hier een niet-significant verschil zien.

Achterstandsscores bao-scholen

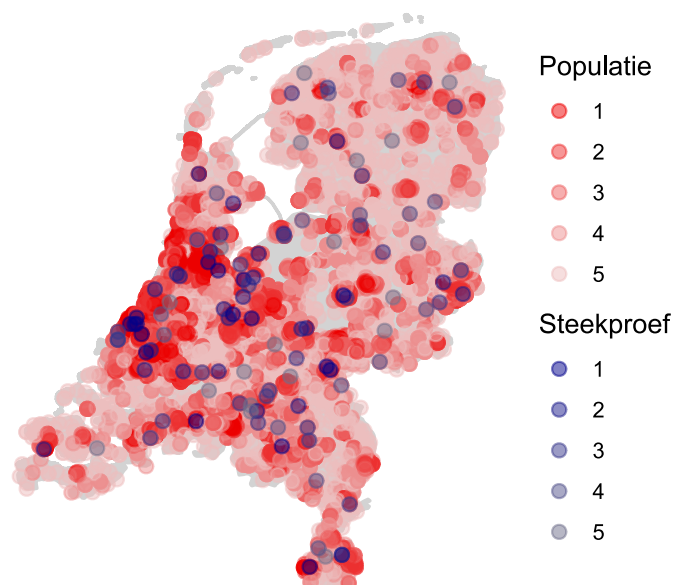


Figuur B1.2 – Achterstandsscores bao-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

Zoals in Figuur B1.2 te zien is, hebben de basisscholen in de steekproef licht hogere achterstandsscores dan in de populatie. De verschillen zijn echter klein.

Stedelijkheid bao-scholen

Stedelijkheid



Figuur B1.3 – Stedelijkheid bao-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

Tabel B1.2 – Verdeling naar stedelijkheid populatie en steekproef bao-scholen en resultaat chi-kwadraat toets, Peil.Rekenen-wiskunde PO 2023

Stedelijkheid	Groep	Aantal	Percentage	
1	steekproef	17	16.0	
	populatie	974	15.7	
2	steekproef	27	25.5	
	populatie	1,491	24.0	
3	steekproef	28	26.4	
	populatie	1,098	17.7	
4	steekproef	16	15.1	
	populatie	1,144	18.4	
5	steekproef	18	17.0	
	populatie	1,501	24.2	
Chi-kwadraat toets				
aantal:	$\chi^2 = 7.43$, df = 4, p = 0.11			
percentage:	$\chi^2 = 3.36$, df = 4, p = 0.50			

De basisscholen in de responsgroep zijn representatief met betrekking tot stedelijkheid, zoals in bovenstaande figuur en tabel af te lezen valt.

Provincie bao-scholen

Alle provincies zijn vertegenwoordigd in de responsgroep en de complete verdeling over alle provincies heen is ook vergelijkbaar met de verdeling van de scholen in de populatie (Tabel B1.3).

Tabel B1.3 – Verdeling naar provincie populatie en steekproef bao-scholen en resultaat chi-kwadraat toets, Peil.Rekenen-wiskunde PO 2023

	Groep	Aantal	Percentage	
Zuid-Holland	steekproef	18	17.0	
	populatie	1,111	17.9	
Noord-Brabant	steekproef	17	16.0	
	populatie	821	13.2	
Noord-Holland	steekproef	15	14.2	
	populatie	863	13.9	
Gelderland	steekproef	15	14.2	
	populatie	856	13.8	
Overijssel	steekproef	10	9.4	
	populatie	499	8.0	
Friesland	steekproef	8	7.5	
	populatie	367	5.9	
Limburg	steekproef	7	6.6	

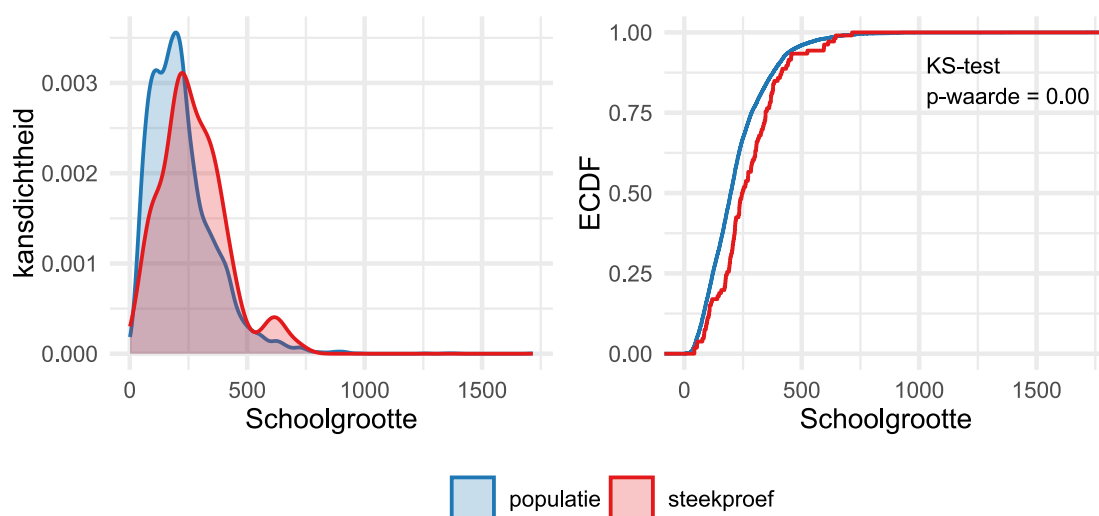
	Groep	Aantal	Percentage	
Utrecht	populatie	328	5.3	
	steekproef	4	3.8	
	populatie	484	7.8	
Groningen	steekproef	4	3.8	
	populatie	252	4.1	
Flevoland	steekproef	5	4.7	
	populatie	181	2.9	
Drenthe	steekproef	1	0.9	
	populatie	253	4.1	
Zeeland	steekproef	2	1.9	
	populatie	193	3.1	

Chi-kwadraat toets

aantal: $\chi^2 = 8.14$, df = 11, p = 0.70

percentage: $\chi^2 = 4.85$, df = 11, p = 0.94

Schoolgrootte bao-scholen



Figuur B1.4-- Schoolgrootte bao-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

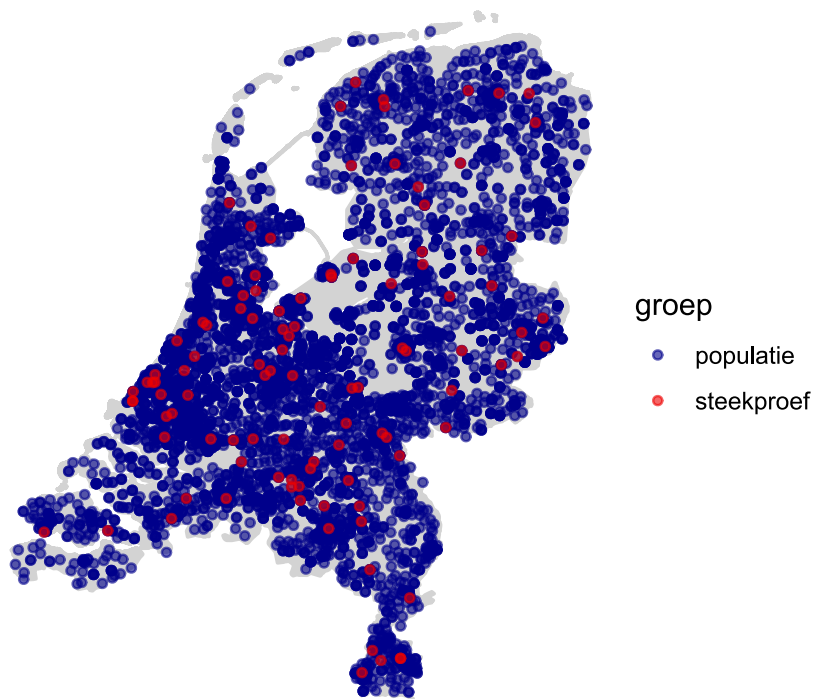
Wat betreft schoolgrootte zijn de verdelingen van de populatie en de onderzoeksgroep minder vergelijkbaar. Figuur B1.4 geeft aan dat in verhouding de steekproef meer grote en minder kleine scholen heeft geselecteerd. De Kolmogorov-Smirnovtoets geeft een zeer kleine p-waarde wat aangeeft dat het verschil in verdelingen significant is. De impact en mogelijke significantie van deze representatie onderzoeken we nader door de schoolgrootte te categoriseren en als discreet kenmerk te benaderen.

Denominatie bao-scholen

Tabel B1.4 – Verdeling naar denominatie populatie en steekproef bao-scholen en resultaat chi-kwadraat toets, Peil.Rekenen-wiskunde PO 2023

	Groep	Aantal	Percentage	
Rooms-Katholiek	steekproef	44	41.5	
	populatie	1,856	29.9	
Openbaar	steekproef	31	29.2	
	populatie	1,943	31.3	
Protestants-Christelijk	steekproef	24	22.6	
	populatie	1,546	24.9	
Algemeen bijzonder	steekproef	3	2.8	
	populatie	320	5.2	
Reformatorisch	steekproef	2	1.9	
	populatie	161	2.6	
Chi-kwadraat toets				
aantal:	$\chi^2 = 5.87$, df = 4, p = 0.21			
percentage:	$\chi^2 = 2.76$, df = 4, p = 0.60			

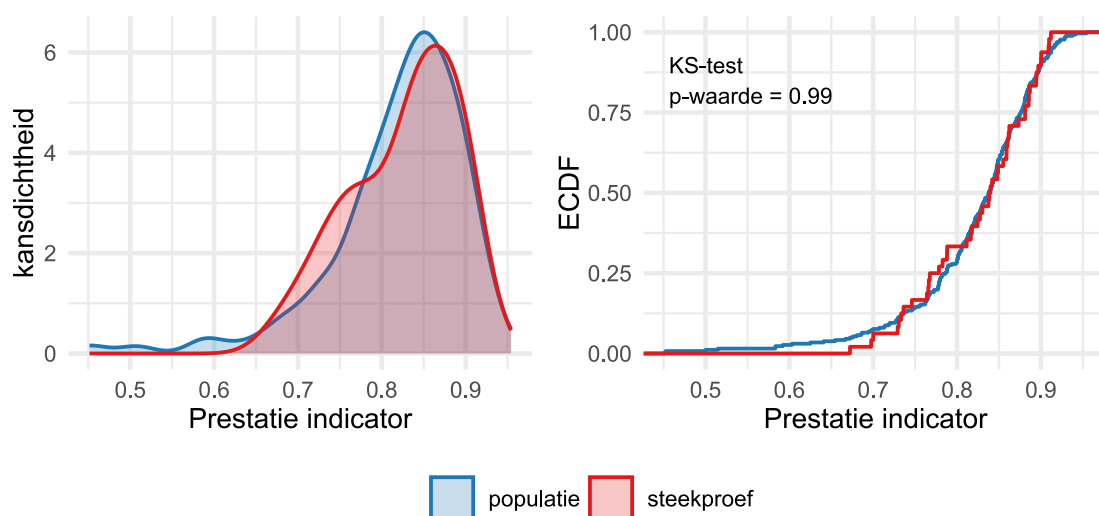
Geografische spreiding bao-scholen



Figuur B1.5 – Geografische spreiding bao-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

B1.3 Speciaal basisonderwijs

Prestatie-indicator sbo-scholen

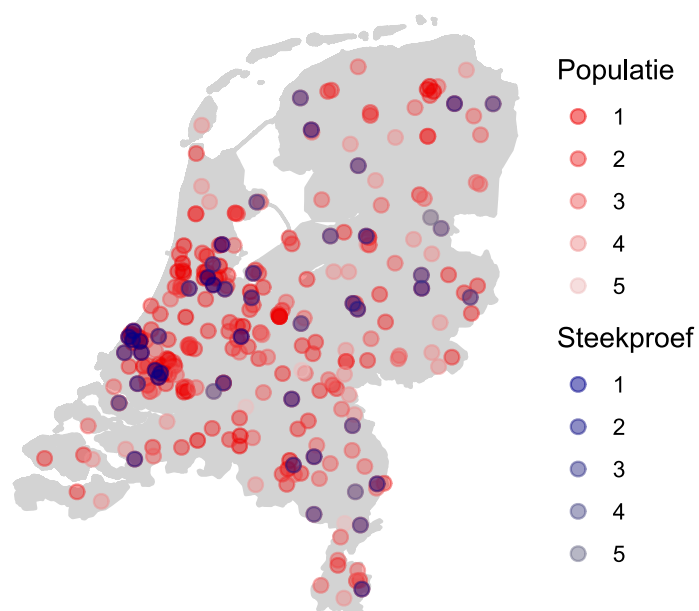


Figuur B1.6 – Prestatie-indicator sbo-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

In Figuur B1.6 is te zien dat de verdeling van de prestatie-indicator in de steekproef en de populatie vrijwel hetzelfde is. Dit blijkt ook uit het resultaat van de Kolmogorov-Smirnovtoets die te vinden is in het rechterdeel van de figuur.

Stedelijkheid sbo-scholen

Stedelijkheid



Figuur B1.7 – Stedelijkheid sbo-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

Tabel B1.5 – Verdeling naar stedelijkheid populatie en steekproef sbo-scholen en resultaat chi-kwadraat toets, Peil.Rekenen-wiskunde PO 2023

Stedelijkheid	Groep	Aantal	Percentage	
1	steekproef	12	24.5	
	populatie	59	21.9	
2	steekproef	21	42.9	
	populatie	120	44.4	
3	steekproef	12	24.5	
	populatie	59	21.9	
4	steekproef	3	6.1	
	populatie	28	10.4	
5	steekproef	1	2.0	
	populatie	4	1.5	
Chi-kwadraat toets				
aantal:	$\chi^2 = 1.14$, df = 4, p = 0.89			
percentage:	$\chi^2 = 1.51$, df = 4, p = 0.82			

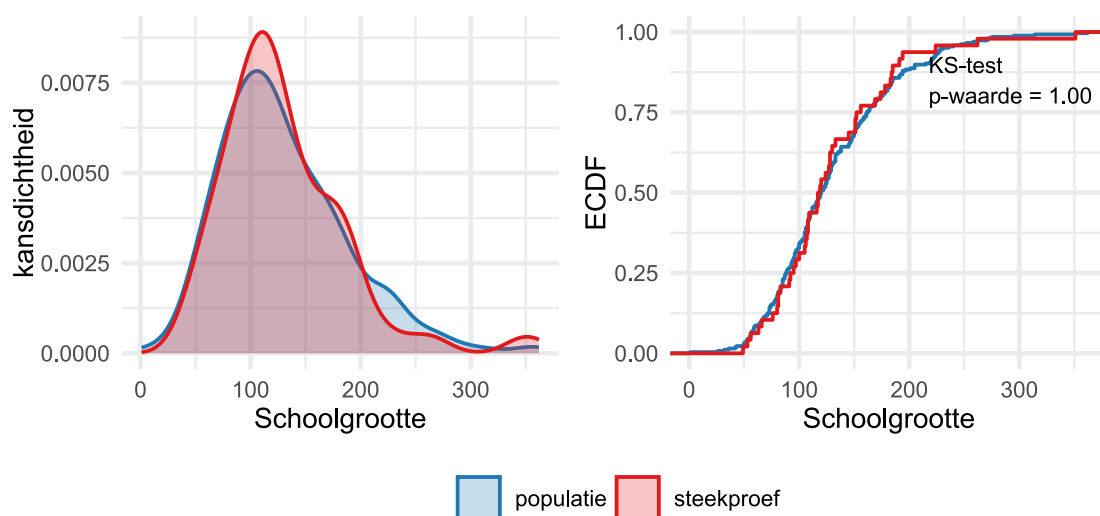
Tabel B1.6 – Verdeling naar provincie populatie en steekproef sbo-scholen en resultaat chi-kwadraat toets, Peil.Rekenen-wiskunde PO 2023

	Groep	Aantal	Percentage	
Zuid-Holland	steekproef	13	26.5	
	populatie	60	22.2	
Noord-Holland	steekproef	8	16.3	
	populatie	44	16.3	
Noord-Brabant	steekproef	6	12.2	
	populatie	31	11.5	
Overijssel	steekproef	6	12.2	
	populatie	20	7.4	
Gelderland	steekproef	3	6.1	
	populatie	34	12.6	
Limburg	steekproef	5	10.2	
	populatie	19	7.0	
Friesland	steekproef	3	6.1	
	populatie	11	4.1	
Utrecht	steekproef	1	2.0	
	populatie	19	7.0	
Groningen	steekproef	2	4.1	
	populatie	10	3.7	
Flevoland	steekproef	2	4.1	
	populatie	7	2.6	
Drenthe	populatie	8	3.0	
Zeeland	populatie	7	2.6	
Chi-kwadraat toets				
aantal:	$\chi^2 = 8.71$, df = 11, p = 0.65			
percentage:	$\chi^2 = 13.48$, df = 11, p = 0.26			

Alle provincies zijn vertegenwoordigd in de responsgroep en de complete verdeling over alle provincies heen is ook vergelijkbaar met de verdeling van de scholen in de populatie.

Schoolgrootte sbo-scholen

De verdeling naar schoolgrootte van de sbo-scholen in de populatie en de steekproef is zeer vergelijkbaar (Figuur B1.8).



Figuur B1.8 – Schoolgrootte sbo-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

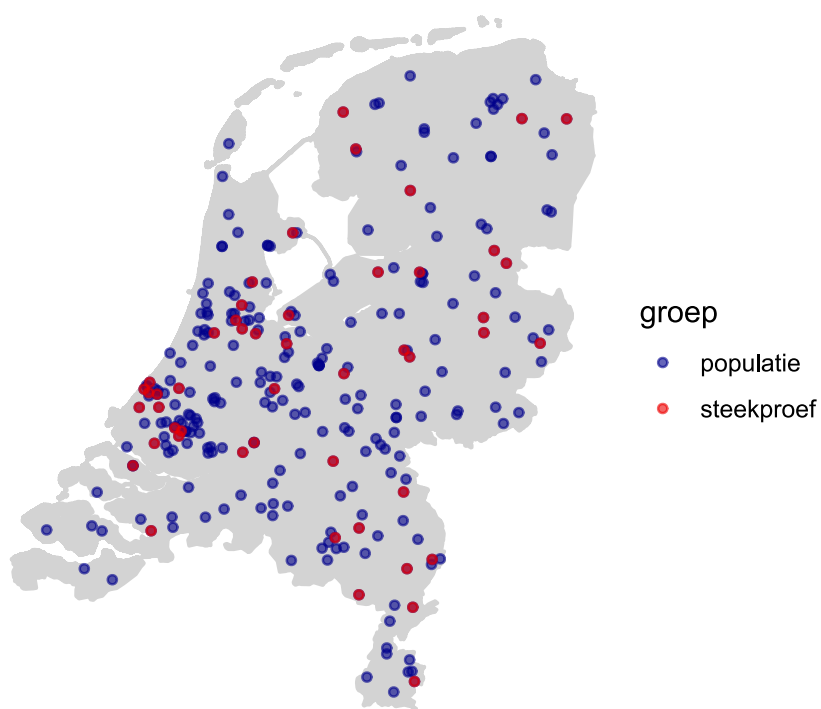
Denominatie sbo-scholen

Ook de denominaties van de sbo-scholen in de steekproef zijn zeer vergelijkbaar met die in de populatie (Tabel B1.7).

Tabel B1.7 – Verdeling naar denominatie populatie en steekproef sbo-scholen en resultaat chi-kwadraat toets, Peil.Rekenen-Wiskunde PO 2023

	Groep	Aantal	Percentage	
Rooms-Katholiek	steekproef	15	30.6	
	populatie	75	27.8	
Openbaar	steekproef	13	26.5	
	populatie	69	25.6	
Protestants-Christelijk	steekproef	11	22.4	
	populatie	56	20.7	
Algemeen bijzonder	steekproef	8	16.3	
	populatie	54	20.0	
Reformatorisch	steekproef	2	4.1	
	populatie	8	3.0	
Chi-kwadraat toets				
aantal:	$\chi^2 = 0.59$, df = 4, p = 0.96			
percentage:	$\chi^2 = 0.73$, df = 4, p = 0.95			

Geografische spreiding sbo-scholen

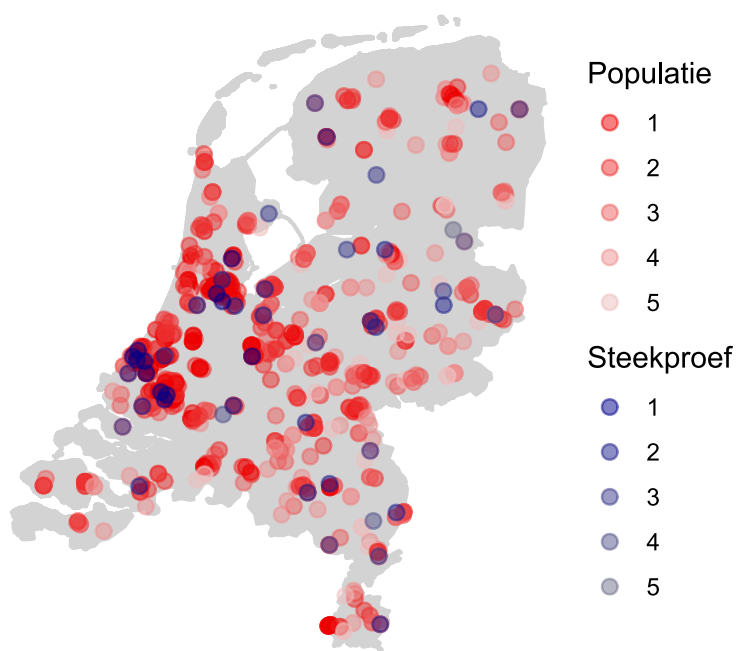


Figuur B1.9 – Geografische spreiding sbo-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

B1.4 Speciaal onderwijs

De steekproef van de so4-scholen is kleiner dan die van bao en sbo. Hierna blijkt ook dat deze op de verschillende variabelen wat minder representatief is. In de populatie is bijvoorbeeld meer variatie in schoolgrootte dan in de steekproef. De overeenkomst in stedelijkheid is wel vrij goed.

Stedelijkheid



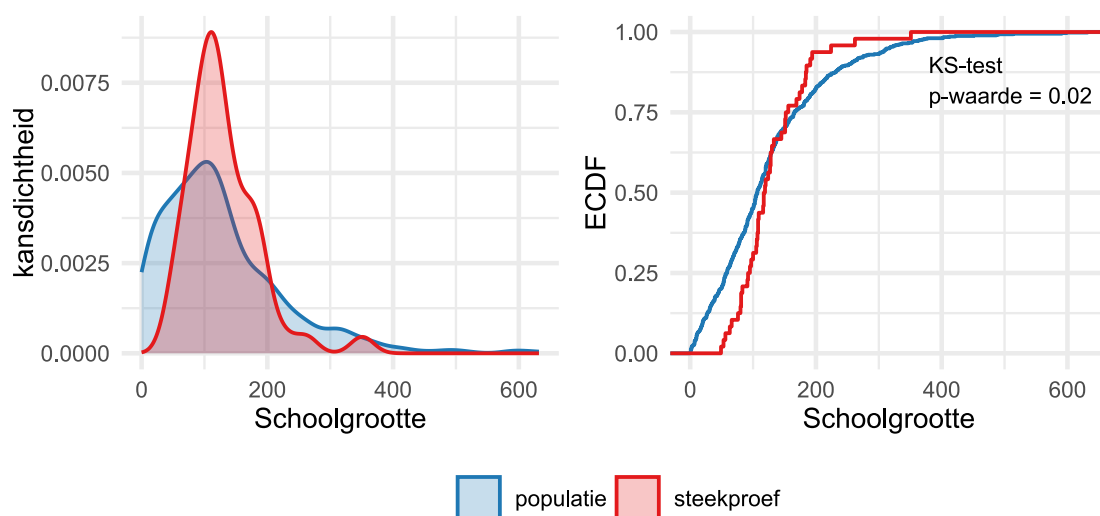
Figuur B1.10 – Stedelijkheid so4-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

Tabel B1.8 – Verdeling naar stedelijkheid populatie en steekproef so4-scholen en resultaat chi-kwadraat toets, Peil.Rekenen-wiskunde PO 2023

Stedelijkheid	Groep	Aantal	Percentage	
1	steekproef	12	24.5	
	populatie	159	25.9	
2	steekproef	21	42.9	
	populatie	221	36.1	
3	steekproef	12	24.5	
	populatie	112	18.3	
4	steekproef	3	6.1	
	populatie	82	13.4	
5	steekproef	1	2.0	
	populatie	39	6.4	
Chi-kwadraat toets				
aantal:	$\chi^2 = 4.81$, df = 4, p = 0.31			
percentage:	$\chi^2 = 6.45$, df = 4, p = 0.17			

Schoolgrootte so4-scholen

In de populatie van so4-scholen is meer variatie in schoolgrootte dan in de steekproef (Figuur B1.11).



Figuur B1.11 - Schoolgrootte so4-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

Denominatie so4-scholen

Ook wat betreft denominatie wijken de so4-scholen in de steekproef af van de populatie

Tabel B1.9 – Verdeling naar denominatie populatie en steekproef so4-scholen en resultaat chi-kwadraat toets, Peil.Rekenen-wiskunde PO 2023

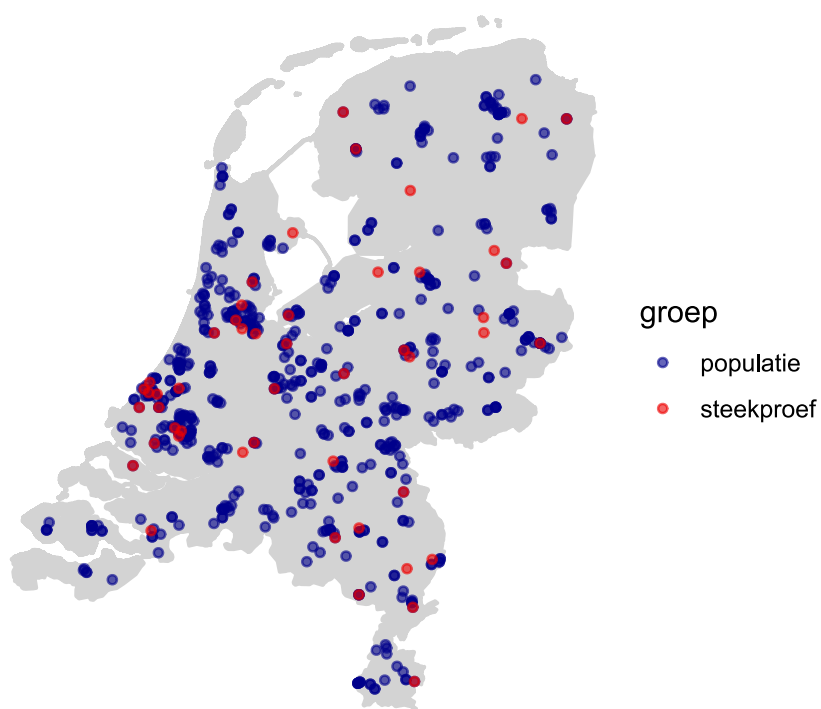
	Groep	Aantal	Percentage	
Algemeen bijzonder	steekproef	8	16.3	
	populatie	270	44.0	
Rooms-Katholiek	steekproef	15	30.6	
	populatie	130	21.2	
Openbaar	steekproef	13	26.5	
	populatie	106	17.3	
Protestants-Christelijk	steekproef	11	22.4	
	populatie	75	12.2	
Reformatorisch	steekproef	2	4.1	
	populatie	11	1.8	

Chi-kwadraat toets

aantal: $\chi^2 = 16.59$, $df = 4$, $p = 0.00$

percentage: $\chi^2 = 20.23$, $df = 4$, $p = 0.00$

Geografische spreiding so4-scholen



Figuur B1.12 – Geografische spreiding so4-scholen populatie en steekproef, Peil.Rekenen-wiskunde PO 2023

B1.5 Representativiteit op leerlingniveau

In Tabel B1.10 en B1.11 worden de populatie en steekproef van leerlingen op bao- en sbo-scholen op enkele kenmerken vergeleken. Voor so4 zijn deze percentages niet weergegeven, maar is de samenvatting van de toetsing te vinden in Tabel B1.12.

Tabel B1.10 – Vergelijking populatie en steekproef baso-leerlingen, in percentages, Peil.Rekenen-wiskunde PO 2023

			Populatie	Steekproef
Schoolkenmerken	Regio	Noord	14.2	14.1
		Oost	24.7	27.4
		West	42.5	35.9
		Zuid	18.6	22.6
	Schoolgrootte	≤100	17.8	11.3
		101-150	15.4	7.5
		151-200	17.5	12.3
		201-250	16.4	21.7
		251-300	9.8	10.4
		301-350	7.1	14.1
		>350	15.9	22.6
Leerlingenkenmerken	Geslacht	Jongen	50.0	49.1
		Meisje	50.0	46.2
		Anders	0.0	0.8
		Onbekend	0.0	3.9

Tabel B1.11 – Vergelijking populatie en steekproef sbo-leerlingen, in percentages, Peil.Rekenen-wiskunde PO 2023

			Populatie	Steekproef
Schoolkenmerken	Regio	Noord	10.9	10.4
		Oost	23.3	22.9
		West	46.5	45.8
		Zuid	19.3	20.8
	Schoolgrootte	≤100	42.0	33.3
		101-150	28.1	35.4
		151-200	14.5	25.0
		201-250	7.2	2.1
		251-300	3.4	2.1
		301-350	2.4	0.0
		>350	2.5	2.1
Leerlingkenmerken	Geslacht	Jongen	64.8	61.4
		Meisje	35.2	32.7
		Anders	0.0	2.6
		Onbekend	0.0	3.3

In de twee voorafgaande tabellen valt op dat voor het kenmerk schoolgrootte de percentuele verdeling van de onderzoeksgroep van zowel bao- als sbo-leerlingen afwijkt van de populatie. Verder is te zien dat voor het kenmerk regio de percentages ook enigszins afwijken voor bao, met name regio West. Regio West is ondervertegenwoordigd in de onderzoeksgroep voor bao. Voor sbo zijn de percentages voor het kenmerk regio vergelijkbaar. Voor geslacht zijn de percentages van de onderzoeksgroep goed vergelijkbaar met die in de populatie voor zowel bao als sbo. Om te kijken of de verschillen tussen de onderzoeksgroep en de populatie significant zijn, zijn in een volgende stap chi-kwadraattoetsen uitgevoerd. De resultaten zijn zichtbaar in Tabel B1.12.

Tabel B1.12 – Toetsing van de verschillen tussen populatie en steekproef van leerlingen naar schooltype, Peil.Rekenen-wiskunde PO 2023

Variabele	Schooltype	N	Chi	df	p	ΦCramer	Kwalificatie
Regio	bao	106	2.38	3	0.50	0.13	Klein tot gemiddeld
	sbo	49	0.07	3	0.99	0.01	Klein
	so4						
Schoolgrootte	bao	106	42.00	36	0.23	0.68	Groot
	sbo	49	28.00	24	0.26	0.82	Groot
	so4	27	23.33	20	0.27	1.00	Groot
Geslacht	bao	3.289	6.00	4	0.19	0.04	Klein
	sbo	1.018	6.00	4	0.20	0.09	Klein
	so4						
Prestatie-indicator	bao	106			0.77		Klein
	sbo	49			0.98		Klein
	so4						
Achterstandsscore	bao	106			0.67		Klein
	sbo						
	so4						

De chi-kwadraatafstand, oftewel de waarde van de chi-kwadraattoets, staat in functie van de steekproefgrootte en de grootte van de kruistabel (of het aantal vrijheidsgraden), wat betekent dat zelfs een zeer klein verschil statistisch significant kan zijn bij een grote steekproef ($n > 500$), ook al zijn de effectgroottes van de verschillen klein of zelfs onbeduidend. Daarom is er voor de kenmerken met significante chi-kwadraattoetsen daarnaast ook gekeken naar de grootte van het effect, met behulp van $\Phi\text{Cramer} = \sqrt{\chi^2 / (n \times df)}$. Een waarde van $0.1/\sqrt{df}$ is gekwalificeerd als klein effect, $0.3/\sqrt{df}$ als een gemiddeld effect en $0.5/\sqrt{df}$ als een groot effect. Daarnaast is ook de significantie van de Kolmogorv-Smirnov toets voor de continue kenmerken aangegeven. De effecten van het kenmerk regio zijn voor bao klein tot gemiddeld en voor sbo klein. De effecten van het kenmerk geslacht en prestatie-indicator zijn voor zowel bao als sbo klein. Het effect van achterstandsscore voor bao is klein. Voor sbo zijn geen achterstandsscores bekend. Voor so4 zijn niet genoeg gegevens bekend om de kenmerken regio, geslacht, prestatie-indicator en achterstandsscore te kunnen benaderen. Er is geen reden om aan te nemen dat de kenmerken regio en geslacht, prestatie-indicator en achterstandsscore niet representatief zijn. We kunnen ze dus als voldoende representatief beschouwen en er hoeft geen weging te worden toegepast.

Voor alle schooltypen is de chi-kwadraattoets van het kenmerk schoolgrootte significant. Dat betekent dat de schoolgrootte dus niet als voldoende representatief kan worden beschouwd.

B1.5 Wel of niet wegen?

Om het effect van niet-representativiteit van schoolgrootte in kaart te brengen, hebben we de ongewogen met de gewogen (concept-)hoofresultaten met elkaar vergeleken. Dit hebben we gedaan door het toepassen van een wegingsfactor naar schoolgrootte en gekeken of dit effect heeft op de scoreverdeling van leerlingen en de verdeling van leerlingen over de referentieniveaus. De wegingsfactor is berekend door het aandeel scholen in een schoolgrootte-categorie in de populatie te delen door het aandeel in de onderzoeksgroep. (poststratificatie). De gemiddelde wegingsfactor per schoolgrootte-categorie is weergegeven in Tabel B1.13.

Tabel B1.13 - Gemiddelde wegingsfactor per schooltype en categorie van schoolgrootte, Peil.Rekenen-wiskunde PO 2023

		Populatie		Steekproef		
Schooltype	School-grootte	Aantal	%	Aantal	%	Gewicht
bao	≤100	1.085	17.76	12	11.32	1.57
	101-150	944	15.45	8	7.55	2.05
	151-200	1.068	17.48	13	12.26	1.43
	201-250	1.002	16.40	23	21.70	0.76
	251-300	601	9.84	11	10.38	0.95
	301-350	437	7.15	15	14.15	0.51
	>350	973	15.92	24	22.64	0.70
sbo	≤100	351	41.99	16	33.33	1.26
	101-150	235	28.11	17	35.42	0.79
	151-200	121	14.47	12	25.00	0.58
	201-250	60	7.18	1	2.08	3.45
	251-300	28	3.35	1	2.08	1.61
	301-350	20	2.39	0	0.00	
	>350	21	2.51	1	2.08	1.21
so4	≤100	37	39.36	8	29.63	1.33
	101-150	30	31.91	14	51.85	0.62
	151-200	8	8.51	2	7.41	1.15
	201-250	6	6.38	1	3.70	1.72
	251-300	4	4.26	1	3.70	1.15
	301-350	4	4.26	0	0.00	
	>350	5	5.32	1	3.70	1.44

De verschillen in scoreverdeling van leerlingen wanneer uitgegaan wordt van gewogen versus ongewogen scores zijn bekeken en bleken geen zichtbare verschillen te geven. De verdeling over de referentieniveaugroepen <1F, 1F en 1S op basis van gewogen versus ongewogen scores is

weergegeven in Tabel B1.14. Uit een chi-kwadraattoets blijkt dat de verschillen tussen de verdeling over referentieniveaus afwezig tot maximaal klein zijn voor alle schooltypen. Voor bao is Φ Cramer 0.06, voor sbo 0.1 en voor so4 0.15. Op basis van deze uitkomsten is besloten om geen weging toe te passen naar schoolgrootte.

Tabel B1.14 – Gewogen versus ongewogen aantallen en percentages leerlingen op de referentieniveaus, Peil.Rekenen-wiskunde PO 2023

		Gewogen		Ongewogen	
Schooltype	Referentie-niveau	Aantal	%	Aantal	%
bao	<1F	522	17.9	712	20.2
	1F	1.849	63.6	1.862	52.9
	1S	538	18.5	944	26.8
	Totaal	2.909		3.518	
so4	<1F	694	83.6	887	84.2
	1F	115	13.9	144	13.7
	1S	21	2.5	22	2.1
	Totaal	830		1.053	
sbo	<1F	179	42.7	267	46.0
	1F	190	45.4	245	42.2
	1S	50	11.9	69	11.9
	Totaal	419		581	
Totaal	<1F	1.469	35.3	1.866	36.2
	1F	1.789	43.0	2.251	43.7
	1S	900	21.6	1.035	20.1
	Totaal	4.158		5.152	

B1.6 Conclusie

De representativiteit van de responsgegevens hebben we beoordeeld door de verdeling van de relevante kenmerken in de populatie te vergelijken met de verdeling van dit kenmerk in de responsgroep. Binnen de hoofdafname van Peil.Rekenen-wiskunde PO 2023 zien we dat de verdeling van deze kenmerken voor zowel bao als sbo mooi vergelijkbaar zijn met de verdeling van deze kenmerken in de populatie. Er is dus geen aanleiding om te veronderstellen dat de non-respons tot een selectieve vertekening gaat leiden in de uiteindelijke (toets)resultaten doordat een relevante groep scholen is onder- of oververtegenwoordigd in de responsgroep.

Bijlage 2 Prestaties op onderdelen

B2.1 Inleiding

In Tabel B2.1 staan de prestaties op onderdelen van de schaal, uitgedrukt in de percentages van items in de bank voor het gegeven onderdeel die de betreffende groep leerlingen naar verwachting correct beantwoorden. Vragen die een leerling niet heeft gemaakt zijn geïmputeerd.

Tabel B2.1 - Prestaties op onderdelen, Peil.Rekenen-wiskunde PO 2023. De groep laagvaardige leerlingen betreft de 10% leerlingen met de laagste vaardigheid in een schooltype. De hoogvaardige leerlingen zijn de 10% meest vaardige leerlingen binnen het schooltype.

Onderdeel	Categorie	bao			sbo			so4		
		laag vaar dig	alle	hoog vaar dig	laag vaar dig	alle	hoog vaar dig	laag vaar dig	alle	hoog vaar dig
context	context	26.9	61.7	85.6	7.5	26.3	58.7	12.7	46.7	79.3
	kaal	31.3	62.1	88.2	9.5	28.8	62.6	15.6	48.2	82.4
domein	getallen	30.3	61.6	87.9	9.0	28.1	62.0	14.8	47.4	82.1
	meten en meetkunde	30.2	60.9	87.7	9.4	28.2	61.7	15.4	47.1	81.8
	probleem-oplossen	9.4	34.6	69.8	1.7	10.0	33.4	3.2	23.1	58.8
	verbanden	23.9	55.4	84.5	5.8	22.4	56.0	10.4	41.5	78.3
	verhoudingen	26.3	58.7	84.6	6.2	24.5	60.1	11.1	44.6	79.0
rekenmachine	ja	19.8	43.3	69.2	6.2	18.7	43.5	9.7	32.4	61.5
	nee	27.9	59.2	86.6	8.0	26.1	59.8	13.4	45.2	80.5
rekenonderwerp	combinaties van bewerkingen	18.3	52.0	84.9	4.2	18.3	51.5	7.5	36.6	76.8
	gewicht	36.4	70.4	93.3	9.1	33.6	73.1	16.3	55.7	90.0
	inhoud	23.3	53.0	82.7	6.5	22.0	53.1	11.4	39.4	75.2
	lengte en omtrek	27.7	57.7	85.9	9.7	26.7	57.9	15.1	44.6	79.3
	oppervlakte	25.7	58.3	87.0	6.8	24.2	58.8	11.6	43.8	80.6
	optellen en aftrekken	34.0	65.3	88.7	9.0	30.6	66.2	15.6	50.9	83.7
	tijd	38.2	65.7	88.7	15.2	35.4	66.1	23.0	53.6	83.6
	tijd en snelheid	6.3	33.6	74.0	1.0	7.8	33.3	2.1	21.3	61.9
	vermenigvuldigen en delen	26.6	57.5	86.3	8.7	25.5	57.7	13.3	43.8	79.5
subdomein	bewerkingen	26.6	58.5	86.7	7.4	25.0	58.7	12.3	44.0	80.1
	getalbegrip en getalrelaties	35.6	66.0	90.2	11.4	32.4	66.6	18.6	52.4	85.1
	meetkunde	42.4	73.8	94.6	13.8	38.1	75.1	22.4	59.1	91.0
	meten	28.3	58.8	86.4	8.7	26.7	59.5	14.4	45.3	80.1
	probleem-oplossen	9.4	34.6	69.8	1.7	10.0	33.4	3.2	23.1	58.8
	procenten	34.9	71.9	95.0	7.7	31.1	74.8	13.7	56.5	91.8
	procenten (incl. breuken)	22.3	51.0	77.1	5.8	21.5	51.7	11.3	38.4	70.4
	verhoudingen	35.1	58.7	90.9	9.1	24.5	71.8	14.9	44.6	87.2
	verhoudingen (incl. breuken)	17.6	50.2	83.2	3.5	17.3	50.6	6.8	35.7	75.1

In Tabel B2.2 staan de prestaties op onderdelen van de schaal, uitgedrukt in de percentages van items in de bank voor het gegeven onderdeel die leerlingen met een bepaald referentieniveau naar verwachting correct beantwoorden. Vragen die een leerling niet heeft gemaakt, zijn geïmputeerd.

Tabel B2.2 – Prestaties op onderdelen naar referentieniveau, Peil.Rekenen-wiskunde PO 2023.

Onderdeel	Categorie	Referentieniveau		
		<1F	1F	1S
context	context	24.6	54.1	77.2
	kaal	28.4	58.4	80.4
domein	getallen	28.9	57.8	78.4
	meten en meetkunde	28.8	57.0	77.9
	probleemoplossen	9.5	28.5	53.6
	verbanden	22.9	51.0	73.5
	verhoudingen	25.2	55.2	75.5
rekenmachine	ja	18.1	39.1	59.6
	nee	25.4	55.2	78.3
rekenonderwerp	combinaties van bewerkingen	16.8	46.5	74.6
	gewicht	33.1	68.1	87.8
	inhoud	21.2	48.1	72.8
	lengte en omtrek	25.8	53.4	77.0
	oppervlakte	23.3	54.0	78.6
	optellen en aftrekken	30.5	62.5	82.1
	tijd	35.3	62.6	81.8
	tijd en snelheid	5.9	25.4	58.4
	vermenigvuldigen en delen	24.6	53.1	77.3
subdomein	bewerkingen	24.2	54.3	78.2
	getalbegrip en getalrelaties	32.4	62.8	83.2
	meetkunde	38.3	71.6	89.7
	meten	25.9	54.7	77.9
	probleemoplossen	9.5	28.5	53.6
	procenten	30.9	70.0	90.1
	procenten (incl. breuken)	20.6	47.2	68.8
	verhoudingen	25.2	55.2	75.5
	verhoudingen (incl. breuken)	16.0	44.7	72.5

B2.2 Gestandaardiseerde effecten

In de meerniveau analyse zijn omwille van de inzichtelijkheid effecten gerapporteerd die direct zijn te relateren aan de onderliggende schalen, zoals uitgelegd in hoofdstuk 7. Voor de volledigheid worden in deze bijlage ook de gestandaardiseerde effecten gegeven.

Deze standaardisatie is een transformatie die wordt gegeven door:

$$\text{gestandaardiseerd effect} = \frac{\text{effect} \times (\text{stand. deviatie onafh. variabele})}{\text{stand. deviatie uitkomst variabele}}$$

We zullen de relatie tussen de reguliere en gestandaardiseerde effecten illustreren met een rekenvoorbeeld. Stel dat er een effect van geslacht is van 0.3 voor de categorie jongens in het sbo en het so4. De vaardigheid in het so4 en sbo wordt gemeten op dezelfde vaardigheidsschaal, dus deze effecten zijn vergelijkbaar en hetzelfde, onder gelijke waardes voor de andere variabelen is de gemiddelde score van jongens 0.3 hoger dan die van meisjes¹³. De eerste stap in de normalisatie is om het effect te delen door de standaarddeviatie van de uitkomstvariabele. Die is ongeveer 0.97 voor sbo en 1.15 voor so4. Hierdoor wordt het effect respectievelijk 0.31 voor sbo en 0.26 voor so4. De tweede stap is het effect vermenigvuldigen met de standaarddeviatie van de variabele geslacht. In Tabel 9.3 lezen we dat 607 van de 938 sbo-leerlingen jongens zijn en 416 van de 485 so4-leerlingen jongens zijn. Dit leidt tot een standaarddeviatie van 0.48 voor geslacht in het sbo en 0.35 in so4. Het resulterende gestandaardiseerde effect is daardoor 0.15 in het sbo en 0.09 in het so4. Dit is het effect op de afzonderlijk gestandaardiseerde, en daardoor niet vergelijkbare, uitkomstvariabelen van 1 standaarddeviatie van de variabele geslacht¹⁴.

Hierna geven we voor elk van de meerniveau modellen in hoofdstuk 7 de gestandaardiseerde effecten.

13 We negeren in deze uiteenzetting eventuele verschillen in betrouwbaarheidsintervallen. Het effect van geslacht is omwille van het voorbeeld gelijkgesteld in het sbo en so4, dit is géén resultaat van de analyse.

14 Voor de volledigheid: op de schaal van de genormaliseerde onafhankelijke variabele *geslacht* is de afstand tussen jongens en meisjes 2.09 in het sbo en 2.86 in het SO.

Tabel B2.3 – Gestandaardiseerde effecten voor de meerniveau analyse met leerlingkenmerken, Peil.Rekenen-wiskunde PO 2023

Variabele	Waarde	sbo			so4			bao		
		beta	betr. interval		beta	betr. interval		beta	betr. interval	
boeken	101-200	-0.03	-0.09	0.03	-0.03	-0.12	0.05	0.08	0.04	0.12
	11-25	-0.01	-0.08	0.05	0.02	-0.07	0.11	0.03	-0.01	0.07
	200+	-0.00	-0.06	0.06	0.06	-0.02	0.14	0.07	0.04	0.11
	26-100	0.05	-0.01	0.11	0.02	-0.07	0.10	0.10	0.05	0.14
geslacht	jongens	0.15	0.09	0.21	0.11	0.04	0.19	0.02	-0.01	0.05
thuis taal	altijd	-0.02	-0.18	0.14	0.14	-0.10	0.38	0.04	-0.06	0.13
	bijna altijd	0.03	-0.10	0.16	0.16	-0.04	0.37	0.02	-0.06	0.09
	soms	-0.05	-0.20	0.09	0.07	-0.14	0.28	-0.04	-0.11	0.04
vertraagd	niet vertraagd	0.03	-0.03	0.09	0.08	0.01	0.16	0.10	0.07	0.13
zelfstandig	elke les	0.13	0.05	0.20	0.18	0.07	0.29	0.08	0.04	0.12
	helft van de lessen	0.07	0.00	0.15	0.13	0.02	0.23	0.05	0.01	0.09
attributies	1-4	-0.04	-0.11	0.03	-0.09	-0.17	0.00	-0.08	-0.11	-0.05
helder	1-4	-0.06	-0.13	0.01	-0.03	-0.12	0.06	0.03	-0.01	0.06
klasklimaat	1-4	0.10	0.04	0.17	0.03	-0.05	0.12	-0.01	-0.04	0.02
mindset	1-4	0.10	0.03	0.16	0.12	0.04	0.20	0.04	0.01	0.07
plezier	1-4	-0.04	-0.11	0.04	0.09	-0.01	0.18	-0.04	-0.07	0.00
zelfvertrouwen	1-4	0.30	0.23	0.38	0.26	0.17	0.35	0.55	0.51	0.58

Tabel B2.4 – Gestandaardiseerde effecten voor de meerniveau analyse met algemene leerkrachtkenmerken, Peil.Rekenen-wiskunde PO 2023.

Variabele	Waarde	sbo			so4			bao		
		beta	betr. interval		beta	betr. interval		beta	betr. interval	
geslacht	Man	-0.02	-0.18	0.14	-0.11	-0.32	0.09	0.08	0.00	0.16
hoogste opleiding	HBO master of universitair	0.04	-0.12	0.20	0.05	-0.16	0.26	-0.05	-0.13	0.03
post HBO rekencoördinator	Nee	0.05	-0.11	0.21	0.16	-0.07	0.39	-0.01	-0.09	0.07
leservaring		0.12	-0.04	0.28	-0.03	-0.25	0.20	-0.09	-0.17	-0.01

Tabel B2.5 – Gestandaardiseerde effecten voor de meerniveau analyse met domeinspecifieke leerkrachtkenmerken, Peil.Rekenen-wiskunde PO 2023

Variable	Waarde	sbo			so4			bao		
		beta	betr. interval		beta	betr. interval		beta	betr. interval	
ervaren ondersteuning rekencoördinator	gemiddeld	0.04	-0.16	0.24	-0.14	-0.38	0.11	0.05	-0.07	0.16
	hoog	-0.10	-0.30	0.10	0.02	-0.20	0.23	0.10	-0.01	0.21
	laag	0.06	-0.13	0.25	-0.11	-0.33	0.10	-0.03	-0.13	0.07
professionaliseringsactiviteiten	6-15 uur	0.00	-0.22	0.22	-0.12	-0.35	0.10	-0.06	-0.17	0.04
	Meer dan 16 uur	-0.10	-0.29	0.09	0.02	-0.23	0.28	-0.06	-0.16	0.04
	Minder dan 6 uur	-0.08	-0.28	0.13	0.07	-0.17	0.32	-0.05	-0.15	0.06
aansluiting rekenmethode	1-5	0.02	-0.15	0.19	-0.14	-0.39	0.11	0.04	-0.05	0.13
mindset	1-4	-0.06	-0.23	0.11	0.14	-0.13	0.41	-0.03	-0.12	0.06
zelfvertrouwen	1-4	0.16	-0.00	0.32	-0.17	-0.42	0.07	0.04	-0.05	0.13

Tabel B2.6 – Gestandaardiseerde effecten voor de meerniveau analyse met domeinspecifieke onderwijskenmerken, Peil.Rekenen-wiskunde PO 2023

Variabele	Waarde	sbo			so4			bao		
		beta	betr. interval		beta	betr. interval		beta	betr. interval	
aanwezigheid rekencoördinator	Nee	0.02	-0.15	0.18	0.09	-0.14	0.32	-0.04	-0.14	0.06
frequentie huiswerk	1 of 2 keer per week	-0.21	-0.65	0.22	0.23	-0.22	0.67	-0.02	-0.30	0.25
	Minder dan één keer per week	-0.27	-0.72	0.18	0.24	-0.18	0.65	-0.06	-0.25	0.14
gebruik digitale apparaten	Nee	0.07	-0.33	0.46	0.07	-0.37	0.51	0.03	-0.12	0.17
gebruik rekenmachine	(beperkt) toegestaan	-0.31	-0.79	0.17	0.37	-0.01	0.75	0.09	-0.06	0.25
meerdere strategieën	Ja, ik bied alle leerlingen meerdere oplossings-strategieën	0.13	-0.06	0.32	0.19	-0.10	0.48	0.09	-0.05	0.23
	Ja, maar alleen voor de gemiddeld en goed presterende leerlingen in rekenen	0.00	-0.16	0.16	0.14	-0.13	0.41	0.08	-0.05	0.22
	Ja, maar alleen voor de goed presterende leerlingen in rekenen	-0.08	-0.32	0.16	0.04	-0.22	0.30	0.00	-0.11	0.12
	Ja, maar alleen voor de zwak en gemiddeld presterende leerlingen in rekenen	0.03	-0.16	0.21				0.06	-0.05	0.16
software/website	Nooit of bijna nooit	-0.11	-0.54	0.32	0.15	-0.29	0.59	-0.04	-0.19	0.12
	Ongeveer de helft van de lessen	0.11	-0.06	0.29	0.01	-0.29	0.30	-0.04	-0.15	0.06
	Sommige lessen	0.08	-0.14	0.30	0.19	-0.15	0.54	-0.00	-0.12	0.11
type methode	niet-traditioneel"	-0.02	-0.23	0.19	-0.02	-0.32	0.27	-0.04	-0.14	0.06
RW onderwijstijd		0.08	-0.12	0.28	0.12	-0.17	0.41	-0.07	-0.17	0.03
direct leren	1-4	-0.28	-0.47	-0.08	-0.01	-0.27	0.25	-0.02	-0.12	0.08
feedback huiswerk	1-4	0.55	-0.02	1.11	0.22	-0.23	0.67	0.05	-0.18	0.28
korte strategieën	1-4	0.06	-0.17	0.29	-0.00	-0.31	0.30	0.09	-0.01	0.19
lange strategieën	1-4	0.06	-0.18	0.29	-0.13	-0.39	0.13	-0.01	-0.12	0.10
leeractiviteiten digitale apparaten	1-4	-0.02	-0.20	0.17	0.08	-0.20	0.36	-0.02	-0.13	0.10
leerkracht gestuurd	1-4	0.18	-0.00	0.36	-0.05	-0.28	0.19	-0.05	-0.15	0.05
ontdekkend leren	1-4	0.00	-0.18	0.18	-0.26	-0.46	-0.07	0.02	-0.09	0.13
rekenmachine activiteiten	1-4	-0.38	-0.90	0.14	0.20	-0.31	0.70	0.12	-0.03	0.28

Tabel B2.7 – Gestandaardiseerde effecten voor de meerniveau analyse met algemene schoolkenmerken, Peil.Rekenen-wiskunde PO 2023

Variabele	Waarde	sbo			so4			bao		
		beta	betr. interval		beta	betr. interval		beta	betr. interval	
onderwijs-concept	Ontwikkelings-gericht onderwijs									
	Traditioneel									
aantal leerlingen		0.07	-0.08	0.21	0.16	-0.04	0.35	0.03	-0.04	0.11
prestatie gerichtheid school	1-5	0.04	-0.11	0.19	0.20	0.00	0.39	0.06	-0.04	0.16
schoolweging								-0.21	-0.31	-0.12
werkdruk	1-4	0.23	0.09	0.37	0.02	-0.17	0.21	0.02	-0.06	0.09

Tabel B2.8-- Gestandaardiseerde effecten voor de combinatie van kenmerken in het compacte model, Peil.Rekenen-wiskunde PO 2023

Variabele	Waarde	sbo			so4			bao		
		beta	betr. interval		beta	betr. interval		beta	betr. interval	
boeken	101-200							0.07	0.03	0.11
	11-25							0.02	-0.02	0.07
	200+							0.06	0.03	0.10
	26-100							0.09	0.04	0.13
geslacht	jongens	0.15	0.09	0.21	0.09	0.01	0.17			
vertraagd	niet vertraagd							0.10	0.07	0.13
zelfstandig	elke les	0.14	0.07	0.22				0.08	0.04	0.12
	helft van de lessen	0.09	0.01	0.16				0.05	0.01	0.09
attributies	1-4				-0.06	-0.14	0.02	-0.09	-0.11	-0.06
mindset	1-4							0.04	0.01	0.07
zelfvertrouwen	1-4	0.30	0.24	0.36	0.36	0.28	0.44	0.53	0.50	0.56
								-0.17	-0.22	-0.11

Bijlage 3 Overzicht variabelen multilevel analyses

LEERLINGVRAGENLIJST*		
Categorie	Kenmerk	Variabelnaam
Algemene leerlingkenmerken	Geslacht	gender
	Vertraagd	vertraagd
	Thuis taal	v1.3
	Aantal boeken thuis	v1.4
	Voorlopig schooladvies	Uitstroomadvies
Domeinspecifieke leerlingkenmerken	Plezier in rekenen	plezier
	Zelfvertrouwen in rekenen	zelfvertrouwen
	Attributies positieve rekenprestaties aan eigen inzet	attributies
	Growth/fixed mindset m.b.t. rekenen	mindset
Domeinspecifieke onderwijskenmerken	Helderheid rekeninstructie	helder
	Ordelijkheid rekenles	klasklimaat
	Zelfstandig werken	v2.3

* vaardigheid begrijpend lezen en voorlopig schooladvies verzameld met leerlingadministratielijst

LEERKRACHTVRAGENLIJST		
Categorie	Kenmerk	Variabelnaam
Algemene leerkrachtkenmerken	Geslacht	A1.2
	Leservaring in betreffende onderwijssoort	A1.1 (bao)
		A1.1_2_3 (sbo)
	Opleiding:	A1.1_3_3 (so4)
	- hoogste opleiding met diploma afgerond - post HBO rekencoördinator gevolgd	A1.4 A1.7
Domeinspecifieke leerkrachtkenmerken	Zelfvertrouwen in vakdidactische vaardigheden	Zelfvertrouwen
	Fixed en growth mindset m.b.t. rekenen	mindset
	Aansluiting rekenmethode	oordeel geschiktheid reken-wiskunde methode
	Ervaren ondersteuning rekencoördinator	steunrc
	Professionaliseringsactiviteiten op het gebied van rekenen-wiskunde: tijdsbesteding	A7.2

LEERKRACHTVRAGENLIJST (vervolg)		
Categorie	Kenmerk	Variabelnaam
Domeinspecifieke onderwijskenmerken	Aanwezigheid rekencoördinator	A2.3
	Onderwijstijd rekenen-wiskunde	minreken (bao) A4.1 (sbo en so4)
	Methode wel/niet traditioneel	methodetype
	Werkvormen:	
	- leerkracht-gestuurde klassikale instructie	lkr gestuurd
	- directe instructie/ontdekkend leren	direct/ontdekkend leren
	Aanleren oplossingsstrategieën	
	- meer dan één oplossingsstrategie	A4.6
	- behandelde oplossingsstrategieën	strat_lang strat_lkort
	Technologische hulpmiddelen:	
	- gebruik digitale apparaten les	A5.1
	- frequentie software/websites	A5.3A
	- leeractiviteiten met digitale apparaten	ictleeract
	- gebruik rekenmachine	A5.5
	- leeractiviteiten met rekenmachine	rekenact
	Huiswerk:	
	- frequentie huiswerk	A6.1
	- feedback huiswerk	huiswerk
	Vormen van evalueren	evaluatie
	Differentiatie	differentiatie
Algemene schoolkenmerken	Onderwijsconcept	A2.1
	Prestatiegerichtheid school	prestatie_sch
	Werkdruk	werkdruk

VARIABLEN SCHOOLNIVEAU (DUO)		
Categorie	Kenmerk	Variabelnaam
Algemene schoolkenmerken	Schoolgewicht (steekproefinformatie)	Schoolgewicht
	Schoolgrootte (steekproefinformatie)	Grade4Students