

PEILING SCHRIJFVAARDIGHEID EINDE (SPECIAAL) BASISONDERWIJS 2019 TECHNISCHE RAPPORTAGE

Lieneke Ritzema | Henk Pander Maat | Ninke Stukker | Bas Hemker |

Harm Naayer | Tecla Lampe | Amos van Gelderen | Hester de Boer

Colofon

Peil.Schrijfvaardigheid is uitgevoerd door een consortium bestaande uit GION Onderwijs/Onderzoek, Cito, Kohnstamm Instituut, Universiteit van Utrecht – afdeling Taal en communicatie, de faculteit der Letteren (afdeling Discourse and communication) en de vakgroep Orthopedagogiek van de Rijksuniversiteit Groningen en NHL-Stenden. Het project is uitgevoerd in opdracht van de Inspectie van het Onderwijs. Bij vermelding van dit rapport in een publicatie dient gerefereerd te worden aan:

Ritzema, E.S., Pander Maat, H., Stukker, N, Hemker, B.T., Naayer, H.M., Lampe, T., Van Gelderen, A., De Boer, H. (2019). *Peiling Schrijfvaardigheid einde (speciaal) Basisonderwijs 2019. Technische rapportage*. Groningen: GION Onderwijs/Onderzoek.

Projectmanagement	Lieneke Ritzema, Roel Bosker
Inhoudelijke input	Henk Pander Maat, Ninke Stukker, Amos van Gelderen, Anke Herder, Bé Poolman
Ontwerp schrijftaken	Amos van Gelderen
Verwerking schrijftaken	Harm Naayer, Demili Swierenga, Jan van Soeren
Beoordeling schrijftaken	Henk Pander Maat, Ninke Stukker, Bas Hemker
Ontwerp vragenlijsten	Amos van Gelderen, Lieneke Ritzema, Harm Naayer
Ontwerp interviews	Anke Herder, Amos van Gelderen, Lieneke Ritzema, Harm Naayer
Werving en begeleiding scholen	Yolande Emmelot, Dorothé Elshof, Harm Naayer, Denise Leijdelmeijer
Logistiek afnames	Yolande Emmelot, Dorothé Elshof, Harm Naayer, Barry de Groot, Denise Leijdelmeijer
Kwaliteitswaarborging afnames	Testleiders Kohnstamm Instituut, onderwijsinspecteurs
Psychometrie en methodologie	Bas Hemker, Tecla Lampe, Djurre Zijsling
Standaardbepaling	Tecla Lampe, Bas Hemker, Lieneke Ritzema
Analyse onderzoeksvragen	Bas Hemker, Henk Pander Maat, Ninke Stukker, Lieneke Ritzema, Hester de Boer
Didactische handreiking	Anke Herder
Redactie	Lieneke Ritzema, Harm Naayer
Eindredactie databestanden	Lieneke Ritzema, Harm Naayer, Djurre Zijsling

ISBN 90-367-xxxxx-3 (E-book)

© november 2020. GION Onderwijs/Onderzoek

Niets van deze uitgave mag worden veelevoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook zonder voorafgaande schriftelijke toestemming van de Directeur van het Instituut.

Inhoudsopgave

Voorwoord	1
Leeswijzer	2
Hoofdstuk 1. Inkadering peiling en instrumentarium.....	5
1.1 Kerndoelen, referentiekader en algemene kenmerken van schrijven.....	6
1.1.1 Algemene kenmerken van schrijven in het basisonderwijs	8
1.2 Opzet en resultaten eerder peilingsonderzoek.....	10
1.3 Domeinbeschrijving en veldraadpleging: richtlijnen voor het peilingsonderzoek.....	11
1.4 Opzet huidige peiling: leerlingvaardigheden en –kenmerken	12
1.4.1 De schrijftaken.....	12
1.4.2 Beoordeling schrijfproducten.....	15
1.4.3 Meting leerlingkenmerken	23
1.5 Opzet huidige peiling: onderwijsleerproces.....	23
1.5.1 Kenmerken van het schrijfonderwijs.....	23
1.5.2 Richtlijnen voor de meting van het onderwijsleerproces	24
1.6 Verdiepend onderzoek	25
1.6.1 Het observatie-instrument.....	26
1.6.2 De interviewleidraden	27
1.6.3 Training.....	27
1.7 Verschillen ten opzichte van de vorige peiling.....	28
Hoofdstuk 2. Steekproef, werving en onderzoeksgroep	29
2.1 Steekproeftrekking bo.....	29
2.1.1 Stratificatiecriteria.....	29
2.1.2. Exclusiecriteria.....	30
2.1.3 Gebruikte bronnen	31
2.1.4 Procedure	31
2.2 Steekproeftrekking in het speciaal basisonderwijs (sbo)	32
2.2.1 Stratificatiecriteria.....	32
2.2.2. Exclusiecriteria.....	32
2.2.3 Gebruikte bronnen	33
2.2.4 Procedure	33
2.3 Opzet en uitvoering van de werving	34
2.3.1 Informeren beoogde scholen	35

2.3.2	Inzet ambassadeurs.....	35
2.3.3	Verloop van de werving.....	36
2.3.4	Aanpassing planning: gevolgen voor de werving en voor de gerealiseerde steekproef.....	36
2.3.5	Representativiteit op schoolniveau.....	41
2.4	Gerealiseerde steekproef: leerlingen.....	44
2.4.1	Representativiteit op leerlingniveau.....	46
Hoofdstuk 3.	Design en afname.....	51
3.1	Uitgangspunten.....	51
3.1.1	Uitgangspunten design.....	53
3.2	Algemene opzet design.....	53
3.2.1	Toetsdesign voor de vragenlijst.....	54
3.2.2	Design voor de schrijftaken (bo).....	54
3.2.3	Design voor de schrijftaken (sbo).....	55
3.3	Toetsdesign over vaardigheden heen.....	55
3.4	Aantallen Leerlingen.....	56
Hoofdstuk 4.	Vorbereiding en uitvoering van de peiling.....	57
4.1	Werving en selectie van testleiders.....	57
4.2	Training van (hoofd)testleiders.....	57
4.3	Pilotonderzoek.....	58
4.4	Vorbereiding van de scholen.....	58
4.5	Organisatie van de afnames.....	59
4.6	Verloop van een gemiddelde afnamedag.....	60
4.7	Verdiepend onderzoek.....	61
Hoofdstuk 5:	Kwaliteit van de instrumenten.....	63
5.1	Stap 1: Data - van observatieformulier naar scores.....	63
5.1.1	Soorten vragen.....	64
5.1.2	Scores op de vragen - aandachtspunten.....	65
5.1.3	Ontbrekende waarden.....	66
5.2	Stap 2.1: Analyses per taak – een analyse van de selectievragen.....	67
5.2.1	Selectievragen en ontbrekende waarden – nieuwe taken.....	67
5.2.2	Selectievragen en ontbrekende waarden - ankertaken.....	68
5.2.3	Definitief verwijderde producten (over alle taken heen).....	69
5.2.4	Nieuwe taken: selectievraag 3.....	70
5.5	Stap 2.2: Analyses per taak: van OFI-itemscores naar taakscores.....	71

5.4	Stap 2.3: Analyses per taak: Beoordelaarsbetrouwbaarheid.....	73
5.5	Stap 3: Van taak naar schaal - OPLM-analyses.....	77
5.5.1	Vorbereiding data: taken combineren.....	77
5.5.2	Vorbereiden design.....	79
5.6	Stap 4: OPLM analyses.....	81
5.6.1	Vooranalyses.....	81
5.6.2	OPLM schaal analyses.....	81
5.6.3	Schattingen met de OPLM schalen.....	84
5.7	Vragenlijstonderzoek.....	85
Hoofdstuk 6.	Scholen, leerkrachten, leerlingen en onderwijs-leerproces.....	95
6.1	Gebruikte informatiebronnen: dataverzameling en respons.....	95
6.2	Kenmerken van de deelnemende scholen.....	99
6.2.1	Organisatie van taal/schrijven op school.....	99
6.2.2	Aandachtspunten in het taalbeleid.....	100
6.2.3	Visie op het schrijfonderwijs.....	101
6.3	Leerkrachtkenmerken.....	102
6.3.1	Achtergrondkenmerken van de leerkrachten.....	102
6.3.2	Professionalisering en self efficacy.....	104
6.3.3	Eigen inschatting aanwezige kennis en vaardigheden.....	104
6.3.4	Sterke punten en lastige aspecten in het schrijfonderwijs.....	105
6.4	Visie en afstemming.....	109
6.4.1	Visie op schrijfonderwijs.....	109
6.4.2	Leerplan (uitvoering visie/doelen in planning en voorbereiding).....	114
6.4.3	Aandacht voor product en procesaspecten.....	116
6.4.4	Afstemming met collega's en doorgaande leerlijn.....	120
6.5	De onderwijspraktijk.....	122
6.5.1	Het onderwijsaanbod.....	122
6.5.2	De lespraktijk.....	133
6.5.3	Feedback, differentiatie, beoordeling en systematisch volgen.....	157
6.6	Resultaten vragenlijst leerlingen.....	170
6.6.1	Achtergrondinformatie.....	170
6.6.2	Ervaring met schrijven.....	172
6.6.3	Schrijfplezier.....	176
6.6.4	Self efficacy.....	178

6.6.5 Resultaten interviews: het schrijfproces beschreven	180
Hoofdstuk 7. Standaardbepaling.....	187
7.1 Inleiding.....	187
7.2 Vorbereidende activiteiten	189
7.2.1 Het expertpanel.....	189
7.2.2 De te gebruiken schalen om standaarden op te bepalen	190
7.3 Standaardbepaling: methode en selectie schrijfproducten	193
7.3.1 Methode van de standaardbepaling	193
7.3.2 Dubbele controle vooraf van oplopende volgorde schrijfproducten.....	197
7.4 Procedure en resultaten standaardbepaling.....	198
7.4.1 Eerste ronde ‘rangefinding’	199
7.4.2 Tweede en derde ronde ‘pinpointing’	204
7.5 Advies referentieniveaus 1F en 1S/2F.....	209
7.6 Bijlagen bij dit hoofdstuk omtrent de standaardsetting	211
Hoofdstuk 8. Resultaten.....	213
8.1 Behaalde percentages referentieniveaus 1F en 1S/2F.....	214
8.1.1 Referentieniveaus en achtergrondkenmerken van de leerlingen.....	215
8.1.2 Relatie tussen referentieniveaus, schrijfniveaus en OFI-aspecten	216
8.2 Verdeling van de vaardigheid.....	217
8.3 Typering van de prestaties van leerlingen	217
8.3.1 Beheersing opgaven	220
8.3.2 Beheersing opgaven door verschillende typen leerlingen	222
8.4 Kwalitatieve typering van leerlingteksten.....	232
8.4.1 Inleiding	232
8.4.2 Een analyse van de itemscores.....	234
8.4.3 Kwalitatieve beschrijving van de vaardigheidsniveaus aan de hand van T04.....	237
8.4.4 Kwalitatieve beschrijving van de vaardigheidsniveaus aan de hand van T10.....	245
8.4.5 Tot slot.....	252
8.5 Algemene indruk, correctheid en aantal woorden	253
8.5.1 Resultaten algemene indruk, correctheid en aantal woorden	253
8.5.2 Relatie algemene indruk en correctheidskenmerken met vaardigheid	258
8.6 Trendvergelijking.....	261
8.6.1 Vergelijking data van de oude en de nieuwe peiling.....	263
8.6.1 Vergelijking van scores per taak (bo en sbo).....	271

8.6.2 Vergelijking scores additionele aspecten – bo	275
Hoofdstuk 9. Meerniveau-analyses	277
9.1 Gehanteerde strategie voor de meerniveau-analyses	277
9.1.1 Multipele imputatie van ontbrekende gegevens	277
9.1.2 Meerniveau regressie-analyse	279
9.1.3 Berekening van effecten.....	280
9.2 Resultaten meerniveau-analyses	281
Hoofdstuk 10. Overzicht van databestanden en variabelen.....	293
10.1 Opbouw van databestanden	293
10.2 Overzicht van variabelen in het hoofdbestand	294
Geraadpleegde literatuur	297

Voorwoord

Voor u ligt het technisch rapport van de derde peiling schrijfvaardigheid einde (speciaal) basisonderwijs, Peil.Schrijfvaardigheid. Dit peilingsonderzoek heeft plaatsgevonden in de periode april – juli 2019 en is uitgevoerd door een consortium bestaande uit de Rijksuniversiteit Groningen (GION Onderwijs/Onderzoek, Faculteit der Letteren - afdeling Discourse and communication, en de afdeling Orthopedagogiek), de Universiteit Utrecht (Faculteit der Letteren - afdeling Taal en communicatie), het Kohnstamm Instituut en NHL Stenden Hogeschool¹. Het onderzoeksdesign en de psychometrische analyses zijn uitgevoerd door Cito, dat tevens heeft ondersteund bij de logistieke operatie. Cito heeft daarnaast, in nauwe afstemming met het consortium, de standaardbepaling uitgevoerd ter bepaling van de referentieniveaus 1F en 1S/2F.

De peiling is uitgezet door NRO en is uitgevoerd onder regie van de Inspectie van het Onderwijs. Doel van het peilingsonderzoek is om de schrijfvaardigheid, de vaardigheid in het schrijven van teksten, van de groep 8-leerlingen in het regulier basisonderwijs en de schoolverlaters (leerlingen van het laatste leerjaar die doorstromen naar het voortgezet (speciaal) onderwijs) in het speciaal basisonderwijs in kaart te brengen. Daarnaast dient het onderzoek om de ontwikkelingen in die vaardigheid over de tijd vast te stellen en een beeld te krijgen van het onderwijsaanbod, in relatie tot de resultaten van de leerlingen. Data zijn verzameld via twee deelonderzoeken: 1) een hoofdonderzoek, waarin de schrijfvaardigheid is getoetst aan de hand van schrijftaken en waarin vragenlijsten zijn afgenomen (periode april – half juni 2019), gevolgd door 2) een verdiepend onderzoek, bestaande uit lesobservaties en interviews met leerkrachten en leerlingen (afname april – juni 2019). Deze deelonderzoeken volgden elkaar op. Afnames voor het verdiepend onderzoek vonden plaats ná de hoofdafname op de school. Doel van het verdiepend onderzoek was om additionele informatie over het schrijfonderwijs en de reguliere lespraktijk te verzamelen en om meer zicht te krijgen op de schrijfaanpak van leerlingen.

Hoogeveen (2017) heeft voor de huidige peiling een domeinbeschrijving opgesteld met daarin richtlijnen voor de te peilen vaardigheden. Daarnaast vormde de review van Janssen en Van Weijen (2017) de basis voor de te onderzoeken kenmerken van het onderwijsleerproces. Deze richtlijnen zijn gebaseerd op het wettelijke en referentiekader en de resultaten van een veldraadpleging. De huidige peiling sluit grotendeels bij deze gegeven richtlijnen aan. Binnen het hoofdonderzoek zijn 12 schrijftaken gebruikt, waarvan er negen voor deze peiling zijn ontwikkeld en drie reeds in de vorige peiling zijn ingezet (de ankertaken). Om de kwaliteit van de door de leerlingen geschreven teksten vast te stellen is gebruik gemaakt van analytische beoordelingsmodellen, vergelijkbaar aan de vorige peiling (Kuhlemeijer et al., 2013). Uitgangspunt voor de vaststelling van de kwaliteit van de leerlingteksten was hun communicatieve effectiviteit, die in de huidige peiling is bepaald aan de hand van de drie aspecten Opbouw, Formulering en Inhoud. Naast de schrijfvaardigheid is ook informatie verzameld over de mate waarin leerlingen schrijven, en hun attitude en hun zelfwaardering ten opzichte van schrijven.

¹ De uitvoering van de peiling was grotendeels gebaseerd op de oorspronkelijke onderzoeksopzet in het voorstel voor de peiling Schrijfvaardigheid (De Glopper, Minnaert, Van Gelderen, Pander Maat & Kuijpers, 2017).

Omdat schrijfvaardigheid een onderdeel vormt van de referentieniveaus Nederlandse taal vormde een standaardbepaling onderdeel van deze peiling. Dit gebeurde onder regievoering van het College voor Toetsen en Examens (CvTE). Voor twee referentieniveaus, 1F en 1S/2F, hebben experts in januari 2020 een referentiecesuur bepaald door zich te buigen over de vraag bij welk schrijfproduct een leerling het betreffende referentieniveau toonde. Het is voor het eerst dat er standaarden voor schrijfvaardigheid zijn bepaald aan de hand van de referentiekader.

Naast het in kaart brengen van de huidige stand van zaken omtrent de schrijfvaardigheid van de leerlingen in groep 8 en van de schoolverlaters was een tweede doelstelling van deze peiling een trendvergelijking met de schrijfprestaties in de vorige peiling. Aan de hand van de drie ankertaken is nagegaan in hoeverre de schrijfvaardigheid van de leerlingen anno 2019 verschilt van die van de leerlingen in 2009. Hoewel met behulp van de resultaten van de huidige peiling een vergelijking met de resultaten van de vorige peiling kan worden gemaakt, moet worden opgemerkt dat het voorbij de scope van de huidige peiling en dit rapport gaat om verschillen tussen 2009 en 2019 te verklaren. Tenslotte zijn analyses uitgevoerd ter verklaring van verschillen in schrijfvaardigheid tussen leerlingen. Hiertoe is onderzocht of bepaalde kenmerken van de leerlingen, hun leerkrachten en de scholen, en het onderwijsaanbod op het gebied van schrijven voorspellend zijn voor de schrijfvaardigheid van de leerlingen.

In het kader van het peilingsonderzoek is tevens een didactische handreiking opgesteld, die gericht is op leerkrachten van de groepen 6, 7 en 8 van het (speciaal) basisonderwijs en tot doel heeft een impuls te geven aan het schrijfonderwijs.

Leeswijzer

In het huidige rapport wordt een verantwoording gegeven van de theoretische achtergrond, de opzet en het design van de peiling schrijfvaardigheid. Daarnaast wordt de gegevensverzameling beschreven, evenals de kwaliteit van het instrumentarium en schaling, de analyse van de resultaten (zowel kwantitatief als kwalitatief) en de analyses voor het vaststellen van verbanden tussen resultaten en voorspellers. Dit rapport is bedoeld als achtergrond bij de publicatie van de Inspectie van het Onderwijs over de peiling schrijfvaardigheid in 2020. De rapportage is tevens bedoeld voor de gebruikers van de data die in deze peiling zijn verzameld en die voor secundaire analyse via DANS aan onderzoekers ter beschikking worden gesteld. Zij zouden in deze rapportage voldoende achtergrond moeten kunnen vinden om met de data te kunnen werken.

In het eerste hoofdstuk wordt een korte schets gegeven van de theoretische achtergronden die van invloed waren op de constructie van het instrumentarium en de beoordeling van de leerlingteksten. In dit hoofdstuk worden het wettelijke en referentiekader beschreven, alsmede het centrale uitgangspunt binnen deze peiling om de communicatieve effectiviteit van de leerlingteksten te beschouwen als kenmerkend voor hun kwaliteit, inclusief een onderbouwing voor het vaststellen van die kwaliteit aan de hand van de aspecten Opbouw, Formulering en Inhoud. Verder worden kenmerken van het schrijfonderwijs in Nederland beschreven met daaraan gekoppeld de manier waarop zij binnen deze peiling met behulp van vragenlijsten, interviews en lesobservaties in kaart zijn gebracht. Hoofdstuk 1 wordt afgesloten met een bespreking van de verschillen tussen de vorige en huidige peiling. In het tweede hoofdstuk wordt het steekproefkader, de wijze van steekproeftrekking en de werving van scholen toegelicht. Daarbij wordt weergegeven op welke manier is geprobeerd om een representatieve steekproef te realiseren en wordt op zowel school- als leerlingniveau beschreven in hoeverre dit is gelukt. Het derde hoofdstuk bevat een toelichting op het gekozen onderzoeksdesign

en de uitgangspunten die daaraan ten grondslag lagen. Gekozen is voor een incompleet design als gevolg van de beperkte tijd die beschikbaar was voor de afname van de peiling. De voorbereiding en de uitvoering van de peiling staan centraal in hoofdstuk 4. Beschreven wordt hoe de testleiders zijn voorbereid op de afnames, hoe de instrumenten tijdens pilots zijn getest (en welke aanpassingen als gevolg van deze pilots hebben plaatsgevonden) en hoe de scholen zijn voorbereid op de daadwerkelijke peilingsafname. Het vijfde hoofdstuk is gewijd aan de psychometrische kwaliteit van de beoordelingsmodellen en de schaling. Kenmerken van de school, leerkracht en leerlingen alsmede van het onderwijsleerproces staan centraal in hoofdstuk 6. Hierbij is vragenlijstinformatie uit het hoofdonderzoek aangevuld met de resultaten uit de interviews en de lesobservaties. Hoofdstuk 7 gaat in op de manier waarop er, na beoordeling van de leerlingteksten, door experts standaarden zijn bepaald voor de referentieniveaus 1F en 1S/2F. De schrijffprestaties van de leerlingen staan centraal in hoofdstuk 8. De resultaten van de bo- en sbo-leerlingen worden besproken, zowel in termen van vaardigheidsscores als van de referentieniveaus. De resultaten van de leerlingen zijn kwantitatief getypeerd en daarnaast zijn voor twee taken de leerlingteksten kwalitatief getypeerd. Het hoofdstuk wordt afgesloten met een trendvergelijking, waarin de schrijffprestaties van leerlingen in 2019 worden vergeleken met die van de leerlingen in 2009. In hoofdstuk 9 worden de meerniveau-modellen besproken, waarmee is onderzocht of kenmerken van leerlingen, hun leerkrachten, de scholen en het onderwijsleerproces voorspellend zijn voor de schrijffprestaties van leerlingen. Het laatste hoofdstuk, hoofdstuk 10, bevat een beschrijving van de opgeleverde bestanden en een overzicht van variabelen. Deze informatie is met name voor gebruikers van de data voor secundaire analyse van toepassing. Afsluitend treft u een referentielijst. De bijlagen, waarnaar in dit rapport wordt verwezen, zijn opgenomen in een supplement.

Hoofdstuk 1. Inkadering peiling en instrumentarium

Met behulp van peilingsonderzoeken wordt van tijd tot tijd in kaart gebracht wat de opbrengsten van het onderwijs in een bepaald leergebied zijn. Sinds 2014 voert de Inspectie van het Onderwijs de regie over de peilingsonderzoeken onder de naam Peil.onderwijs. Deze onderzoeken worden sinds 2016 via het NRO uitgezet. De peilingen richten zich op de leergebieden die onderscheiden worden in de kerndoelen en vinden plaats in groep 8 van het basisonderwijs en/of bij schoolverlaters in het speciaal basisonderwijs. Peilingen in het primair onderwijs bieden in de volle breedte zicht op leerlingresultaten, ontwikkelingen in resultaten en resultaten in relatie tot het onderwijsleerproces (aanbod, tijd en kwaliteit). De uitkomsten hiervan vormen de empirische basis voor (maatschappelijke) dialogen over de inhoud, kwaliteit en het niveau van het onderwijs en bieden bruikbare inzichten aan scholen, schoolbesturen, beleidsmakers en methode- en toetsontwikkelaars.

Het huidige peilingsonderzoek betreft het leergebied schrijfvaardigheid. Schrijfvaardigheid is onmisbaar in de communicatie met anderen, zeker in onze moderne samenleving waarin de invloed van ICT en mobiele telefoons steeds groter wordt. Door te schrijven kunnen mensen een boodschap overbrengen en hun gedachten en gevoelens uitdrukken. Dit is relevant voor volwassenen, maar ook voor leerlingen in het basisonderwijs. Met behulp van de peiling schrijfvaardigheid einde (speciaal) basisonderwijs is vastgesteld in welke mate leerlingen in staat zijn een diversiteit aan teksten te schrijven, ook in vergelijking tot de vorige peiling. Naast de schrijfvaardigheid is ook de houding van leerlingen ten opzichte van schrijven van teksten en schrijfonderwijs onderzocht. Tevens is het onderwijsleerproces zowel in het reguliere basisonderwijs als in het speciaal basisonderwijs in kaart gebracht. Basis voor dit peilingsonderzoek was de domeinbeschrijving die door SLO is opgesteld en waarin wordt beschreven 'wat de wettelijke eisen zijn voor de inhoud van het specifieke domein en hoe deze inhoudelijke eisen in de praktijk van het onderwijs vorm krijgen' (Hoogeveen, 2017, p. 5), aangevuld met de uitkomsten van een veldraadpleging over de te peilen onderdelen.

In dit hoofdstuk wordt allereerst het wettelijk kader (kerndoelen en referentieniveaus) beschreven, dat richtinggevend is voor het schrijfonderwijs in Nederland (paragraaf 1.1), waarna beknopt wordt ingegaan op (de resultaten van) eerdere peilingsonderzoeken naar schrijfvaardigheid (paragraaf 1.2). Naast het wettelijk kader leverde een ten behoeve van de peiling opgestelde domeinbeschrijving (aangevuld met een veldraadpleging) minimale richtlijnen op voor te peilen leerlingwerken en -kenmerken. Deze staan centraal in paragraaf 1.3. De voor de huidige peiling ontworpen taken voor het meten van de schrijfvaardigheid en de gemaakte keuzes voor de beoordeling van de schrijftaken die hierop zijn gebaseerd, komen aan bod in paragraaf 1.4. In paragraaf 1.5 staan de kenmerken van het huidige Nederlandse schrijfonderwijs centraal, alsmede de richtlijnen voor het meten van dit onderwijsleerproces. Vervolgens wordt beschreven op welke wijze het onderwijsleerproces in deze peiling is gemeten. Dit is deels gedaan met behulp van een verdiepend onderzoek dat op de deelnemende scholen plaatsvond. In paragraaf 1.6 wordt een beschrijving gegeven van dit verdiepende onderzoek en de daarbij gebruikte instrumenten. In paragraaf 1.7 wordt tenslotte weergegeven hoe de huidige peiling zich verhoudt tot de vorige peiling, die in 2009 is uitgevoerd.

1.1 Kerndoelen, referentiekader en algemene kenmerken van schrijven

Binnen het peilingsonderzoek wordt nagegaan welk schrijfvaardigheidsniveau bij de leerlingen aan het einde van het (speciaal) basisonderwijs bereikt wordt en hoe het onderwijsleerproces op het gebied van schrijfonderwijs vorm krijgt op de scholen. Het wettelijk kader, weergegeven door de kerndoelen, en het referentiekader fungeren daarbij als kader voor de te toetsen vaardigheid. Doel van het schrijfonderwijs in het primair onderwijs is het leggen van een basis in schriftelijke geletterdheid, zodat leerlingen in staat zijn in schoolse en buitenschoolse situaties schriftelijk te communiceren. Binnen het wettelijke kader, de kerndoelen voor het primair onderwijs (2006), zijn de doelstellingen voor het schrijfonderwijs weergegeven onder het kopje 'schriftelijk onderwijs'. Drie kerndoelen (5, 8 en 9) zijn vastgesteld:

- Kerndoel 5 - De leerlingen leren naar inhoud en vorm teksten te schrijven met verschillende functies: zoals informeren, instrueren, overtuigen of plezier verschaffen.
- Kerndoel 8 – De leerlingen leren informatie en meningen te ordenen bij het schrijven van een brief, een verslag, een formulier of een werkstuk. Zij besteden daarbij aandacht aan zinsbouw, correcte spelling, een leesbaar handschrift, bladspiegel, eventueel beeldende elementen en kleur.
- Kerndoel 9 – De leerlingen krijgen plezier in het lezen en schrijven van voor hen bestemde verhalen, gedichten en informatieve teksten.

Daarnaast raken ook de kerndoelen 1, 10 en 12 aan aspecten van schrijfvaardigheid, ook al zijn zij oorspronkelijk voor andere domeinen vastgesteld:

- Kerndoel 1: De leerlingen leren informatie te verwerven uit gesproken taal. Ze leren tevens die informatie, mondeling of schriftelijk, gestructureerd weer te geven (mondeling onderwijs).
- Kerndoel 10: De leerlingen leren bij de doelen onder 'mondeling taalonderwijs' en 'schriftelijk taalonderwijs' strategieën te herkennen, te verwoorden, te gebruiken en te beoordelen (Taalbeschouwing, waaronder strategieën).
- Kerndoel 12: De leerlingen verwerven een adequate woordenschat en strategieën voor het begrijpen van voor hen onbekende woorden. Onder 'woordenschat' vallen ook begrippen die het leerlingen mogelijk maken over taal te denken en te spreken.

De kerndoelen schrijven dus voor dat het voor het primair onderwijs van belang is dat leerlingen leren om teksten met verschillende functies en verschillende vormen te schrijven, en dat zij leren om informatie geordend en gestructureerd in de tekst weer te geven. Daarnaast wordt beoogd dat het schrijfonderwijs leerlingen motiveert om te schrijven.

De kerndoelen zijn omschreven in algemene termen en geven nauwelijks indicaties voor de wijze waarop het onderwijs moet worden ingericht en de beoogde opbrengsten (oftewel de beheersingsniveaus van leerlingen). De referentieniveaus voor schrijven beogen specifiekere omschrijvingen te geven van de vaardigheden en kennis die leerlingen moeten beheersen op een bepaald moment in hun schoolloopbaan, en geven specifiekere omschrijvingen van de onderwijsinhoud dan de kerndoelen (Staatsblad 2010 265). Daarbij komen de kerndoelen 1, 5, 8, 9 en 10 terug in de referentieniveaus voor schrijfvaardigheid.

Het referentiekader is voor alle taaldomeinen, en dus ook voor schrijfvaardigheid, opgebouwd uit drie onderdelen: een algemene beschrijving, een omschrijving van de taken en een omschrijving van de taakuitvoering (Expertgroep Doorlopende Leerlijnen Taal en Rekenen (2008). Binnen dit referentiekader zijn voor het einde van het basisonderwijs (groep 8) twee referentieniveaus vastgelegd: het fundamenteel niveau 1F en het streefniveau 1S/2F; de inschatting was dat deze niveaus door respectievelijk 75% van de leerlingen en 50% van de leerlingen zouden worden behaald. De ambitie van de Expertgroep was zelfs hoger: respectievelijk minimaal 85% en 65% zou idealiter respectievelijk niveau 1F en 1S/2F halen. Tabel 1.1 laat algemene beschrijvingen zien van de vaardigheden die op 1F en 2F dienen te worden beheerst.

Tabel 1.1 Algemene niveaubeschrijving schrijven

Taak	1F	2F
Algemene omschrijving	Kan korte, eenvoudige teksten schrijven over alledaagse onderwerpen of over onderwerpen uit de leefwereld.	Kan samenhangende teksten schrijven met een eenvoudige lineaire opbouw over uiteenlopende vertrouwde onderwerpen binnen school, werk en maatschappij.

Deze algemene beschrijvingen van de twee niveaus zijn nader uitgewerkt voor de verschillende soorten taken die leerlingen zouden moeten beheersen (Tabel 1.2).

Tabel 1.2: Niveaubeschrijvingen per type tekst

Taak	1F	2F
Correspondentie	Kan een briefje, kaart of e-mail schrijven om informatie te vragen, iemand te bedanken, te feliciteren, uit te nodigen e.d.	Kan e-mails of informele brieven schrijven en daarbij meningen of gevoelens uitdrukken. Kan met behulp van standaardformuleringen eenvoudige zakelijke brieven produceren en schriftelijke verzoeken opstellen.
Formulieren invullen, berichten, advertenties en aantekeningen	Kan een kort bericht, een boodschap met eenvoudige informatie schrijven. Kan eenvoudige standaardformulieren invullen. Kan aantekeningen maken en overzichtelijk weergeven.	Kan notities, berichten en instructies schrijven waarin eenvoudige informatie van onmiddellijke relevantie voor vrienden, docenten en anderen wordt overgebracht. Kan een advertentie opstellen om bijvoorbeeld spullen te verkopen. Kan aantekeningen maken tijdens een uitleg of les.
Verslagen, werkstukken, samenvattingen en artikelen	Kan een verslag of een werkstuk schrijven en daarbij stukjes informatie uit verscheidene bronnen samenvatten	Kan verslagen en werkstukken schrijven met behulp van een stramien en daarbij informatie uit verscheidene bronnen samenvoegen. Kan onderhoudende teksten schrijven en overtuigen met argumenten. Kan een collage, een krant of muurkrant maken.
Vrij schrijven	Kan eigen ideeën, ervaringen, gebeurtenissen en fantasieën opschrijven in een verhaal, in een informatieve tekst of in een gedicht.	Kan eigen ideeën, ervaringen, gebeurtenissen en fantasieën opschrijven in een verhaal, in een informatieve tekst of in een gedicht.

Het derde onderdeel van het referentiekader bevat omschrijvingen van kenmerken van de taakuitvoering. Zoals kan worden afgeleid uit Tabel 1.3 op de volgende bladzijde worden deze kenmerken gespecificeerd voor zes aspecten: 1) samenhang, 2) afstemming op doel, 3) afstemming op publiek, 4) woordgebruik en woordenschat, 5) spelling, interpunctie en grammatica en 6) leesbaarheid.

Zoals aangegeven in de domeinbeschrijving dienen deze kenmerken in dienst te staan van een succesvolle tekst, die zich kenmerkt door communicatieve effectiviteit en de eisen aan adequaat en correct taalgebruik (Hoogeveen, 2017, Expertgroep Doorlopende Leerlijnen taal en rekenen, 2008). Opmerkelijk is overigens dat er geen kenmerken zijn gespecificeerd voor afstemming op doel op het niveau 1F en dat er bij het aspect spelling, interpunctie en grammatica wordt verwezen naar niveaubeschrijvingen onder Taalbeschouwing/Taalverzorging.

1.1.1 Algemene kenmerken van schrijven in het basisonderwijs

De in de vorige paragraaf beschreven kerndoelen en referentieniveaus zijn door verschillende organisaties uitgewerkt in doorlopende leerlijnen. Zo heeft Van Gelderen een praktische vertaling van de referentieniveaus geschreven (Van Gelderen, 2010) en heeft SLO een uitwerking naar leerlijnen per kerndoel opgesteld, onderverdeeld in de jaargroepen 1-2, 3-4, 5-6 en 7-8 (<http://tule.slo.nl/Nederlands>). Ook andere organisaties, te weten het Expertisecentrum Nederlands (<https://www.leerlijnentaal.nl/schrijven.html>) en CED (basisleerlijnen Nederlandse taal, <https://www.leerlijnen.cedgroep.nl/po-sbo>) hebben de kerndoelen en het referentiekader uitgewerkt in leerlijnen. Deze uitgewerkte leerlijnen bieden leerkrachten houvast om het curriculum voor schrijven praktisch in te vullen.

Tabel 1.3. Omschrijvingen van de kenmerken van de taakuitvoering in de referentieniveaus

Kenmerken van de taakuitvoering	Niveau 1F	Niveau 2F
Samenhang	De informatie is zodanig geordend dat de lezer de gedachtegang gemakkelijk kan volgen en het schrijfdoel bereikt wordt. De meest bekende voegwoorden (en, maar, want, omdat) zijn correct gebruikt, met andere voegwoorden komen nog fouten voor. Fouten met verwijswoorden komen voor. Samenhang in de tekst en binnen samengestelde zinnen is niet altijd duidelijk.	Gebruikt veel voorkomende verbindingswoorden (als, hoewel) correct. De tekst bevat een volgorde: inleiding, kern en slot. Kan alinea's maken en inhoudelijke verbanden expliciet aangeven. Maakt soms nog onduidelijke verwijzingen en fouten in de structuur van de tekst
Afstemming op doel		Kan in teksten met een eenvoudige lineaire structuur trouw blijven aan het doel van het schrijfproduct.
Afstemming op publiek	Kan eigen ideeën, ervaringen, gebeurtenissen en fantasieën opschrijven in een verhaal, in een informatieve tekst of in een gedicht	Past het woordgebruik en toon aan het publiek aan

Tabel 1.3. Omschrijvingen van de kenmerken van de taakuitvoering in de referentieniveaus (vervolg)

Kenmerken van de taakuitvoering	Niveau 1F	Niveau 2F
Woordgebruik en woordenschat	Gebruikt voornamelijk frequent voorkomende woorden.	Fouten met voorzetsels en idiomatische uitdrukkingen komen nog voor. Varieert het woordgebruik.
Spelling, interpunctie en grammatica	Zie niveaubeschrijving taal-beschouwing/taalverzorging	Zie niveaubeschrijving taal-beschouwing/taalverzorging
Leesbaarheid	Hanteert titel. Voorziet een brief op de gebruikelijke plaats van datering, adressering, aanhef en ondertekening. Besteedt aandacht aan de opmaak van de tekst (handschrift, bladspiegel, eventueel beeldende elementen en kleur).	Gebruikt titel en tekstkopjes. Heeft bij langere teksten (meer dan twee A4) ondersteuning nodig bij aanbrengen van de lay-out.

In zijn beschrijving van het schrijfonderwijs in het primair onderwijs noemt Van Gelderen (2010) twee primaire tekstfuncties die van belang zijn om de doelen van de leerlijnen (en daarmee de beheersingsdoelen binnen het referentiekader 1F en 1S/2F) te behalen: de communicatieve functie en de expressieve functie. De communicatieve functie verwijst daarbij naar het kunnen beschrijven, instrueren en argumenteren, terwijl de expressieve functie gericht is op ‘het authentiek uitdrukken van eigen ideeën (persoonlijk, origineel of fantasierijk)’ (Van Gelderen, 2010, p. 18). De derde functie van schrijfonderwijs, de conceptualisatiefunctie of *schrijvend leren*, is erop gericht dat leerlingen (en volwassenen) ook kunnen leren door te schrijven. Door informatie al schrijvend te verwerken en te structureren kan er begrip over een bepaald onderwerp of verband ontstaan. Deze conceptualisatiefunctie van schrijven is niet expliciet in het referentiekader voor de niveaus 1F en 1S/2F opgenomen (zie Van Gelderen, 2010).² Gezien de aansluiting van de peiling bij dit referentiekader is ervoor gekozen deze functie nu niet te peilen bij de leerlingen. Wel is aan leerkrachten gevraagd in hoeverre zij aandacht aan deze conceptualisatiefunctie besteden met als doel bewust kennis en inzicht van leerlingen in de bovenbouw te vergroten door hen teksten te laten schrijven.

Van Gelderen (2010) benadrukt tevens het belang van het schrijven met een goed omschreven inhoudelijk schrijfdoel, gekoppeld aan het lesdoel. Als het lesdoel bijvoorbeeld is dat leerlingen een spannend verhaal moeten schrijven, is het belangrijk dat zij zich ervan bewust zijn dat zij moeten proberen een bepaald effect bij de lezer teweeg te brengen (namelijk spanning en nieuwsgierigheid) en dat zij hierop ook feedback krijgen. Zonder een dergelijk schrijfdoel (het beoogde effect) is het voor leerlingen namelijk moeilijk te focussen op datgene wat geleerd moet worden (en succes te ervaren

² Overigens zou kunnen worden beargumenteerd dat hetgeen in het referentiekader staat onder ‘Verslagen, werkstukken, samenvattingen en artikelen’ raakt aan het schrijvend leren. Binnen deze peiling is echter de algemene karakteristiek van Van Gelderen (2010) aangehouden.

als ze dat doel hebben bereikt). Leerkrachten zouden leerlingen veelvuldig moeten laten oefenen met het schrijven van verschillende soorten teksten mét communicatief schrijfdoel en gespecificeerd publiek, zodat leerlingen zich bewust zijn van de doelgroep waarvoor geschreven moet worden. Gezien het belang van schrijfdoel en publiek zijn beide in alle te peilen taken gespecificeerd.

Centraal uitgangspunt binnen de huidige peiling zijn de hierboven genoemde kerndoelen en de beschrijvingen voor de referentieniveaus 1F en 1S/2F binnen het referentiekader. Tezamen vormen zij het landelijke leerplankader. De wettelijke eisen die zijn weergegeven middels de kerndoelen waren ook leidend bij de vorige peilingen. Het referentiekader was in 2009, het afnamemoment van de vorige peiling, nog niet wettelijk ingevoerd. In de volgende paragraaf wordt beknopt weergegeven hoe de vorige peilingen zijn opgezet en welke resultaten daarin zichtbaar waren.

1.2 Opzet en resultaten eerder peilingsonderzoek

Ten behoeve van de huidige peiling heeft Hoogeveen (2017) een domeinbeschrijving opgesteld, waarin een overzicht wordt gegeven van de gepeilde vaardigheidsaspecten en uitkomsten voor schrijven binnen diverse peilingen vanaf 1984. Drie onderdelen zijn door de jaren heen gepeild, te weten: handschrift, taalverzorging en schrijfvaardigheid. In 1999 is er voor het eerst een peilingsonderzoek uitgevoerd dat specifiek gericht was op schrijfvaardigheid (Krom, et al., 2004) in plaats van het bredere domein Taal. De tweede schrijfpeiling is in 2009 uitgevoerd (Kuhlemeijer et al., 2013) en de huidige peiling in het voorjaar van 2019. Omdat ook de huidige peiling primair tot doel had de schrijfvaardigheid van leerlingen te meten wordt in deze paragraaf een beknopte beschrijving gegeven van de opzet en uitkomsten van de eerdere peilingen schrijfvaardigheid en dan met name van de peiling uit 2009. Eén van de doelen van peilingsonderzoeken is namelijk een trendvergelijking met de vorige peiling. Voor nadere informatie over de eerdere peilingen wordt verwezen naar Krom et al. (2004), Kuhlemeijer et al. (2013), Hoogeveen (2017) en Bonset en Hoogeveen (2007, 2015).

Binnen voorgaande peilingen zijn de gebruikte schrijftaken onder te verdelen in de teksttypen informatief/rapporterend, directief/instructief en argumenterend. De eerdere peilingen omvatten tevens beschouwende teksten, die in de tweede peiling (2009) niet meer voorkwamen. Wel werden in deze tweede peiling verhalende teksttypen ingezet. De taakaanpak van leerlingen en (de aandacht voor) de conceptuele functie van schrijven werden niet gepeild.

In de beide schrijfpeilingen is voor de beoordeling van de schrijfproducten gebruik gemaakt van analytische modellen. Kuhlemeijer ea. (2013, p. 40) beschrijft: *'In een analytische beoordeling beperken de beoordelaars zich tot het registreren van eenvoudig waarneembare kenmerken of eigenschappen van het schrijfproduct (Rijlaarsdam, Van den Bergh & Couzijn, 1996; Feenstra, 2012a, b). Een voorbeeld is het beoordelen van de inhoudelijke kwaliteit aan de hand van het tellen van het aantal inhoudselementen dat in de tekst voorkomt.'* In de eerste schrijfpeiling (Krom et al., 2004) is de kwaliteit van de leerlingwerken op deze manier beoordeeld aan de hand van de drie criteria: (1) *communicatie* (doel- en publieksgerichtheid), (2) *inhoud en organisatie* en (3) *structuur*, terwijl bij de tweede peiling deze drie criteria werden aangevuld met *stijl* (Kuhlemeijer et al., 2013). De beoordelingscategorie *stijl* richtte zich op 'stijl algemeen' (geluidseffecten of het denken of sprekend opvoeren van personages), 'stijl lay-out' (gebruik van tussenkopjes, witregels) en 'stijl verrassend taalgebruik' (bijzondere zinssneden, bijzonder taalgebruik in vergelijking met andere leerlingen). Beoordeling vond in de tweede peiling plaats door getrainde leerkrachten basisonderwijs.

Voor de peiling uit 2009 (Kuhlemeijer, et al., 2013) werden 12 schrijfopdrachten ontwikkeld, waarvan zes ook in 1999 waren gebruikt. Met behulp van deze 12 taken werd niet alleen de schrijfvaardigheid van leerlingen in groep 8 in het regulier basisonderwijs gemeten, maar ook van leerlingen van groep 5 in het reguliere basisonderwijs en schoolverlaters in het sbo. De resultaten in 2009 lieten zien dat de verschillen in schrijfvaardigheid tussen leerlingen groot waren, ook binnen de drie onderscheiden groepen. Over het geheel bleken leerlingen moeite te hebben communicatief effectieve teksten te produceren, hun teksten met stijlmiddelen te verlevendigen en hun teksten correct te structureren (respectievelijk 44%, 89% en 34% van de leerlingen beheerste deze aspecten onvoldoende³, terwijl 29%, 0% en 32% van hen deze aspecten goed beheerste). Ook bleek dat leerlingen van groep 8 weliswaar significant hoger presteerden dan leerlingen in groep 5, maar het verschil was klein. Sbo-leerlingen presteerden ongeveer op het niveau van de groep 5-leerlingen. Een vergelijking met de resultaten uit de peiling van 1999 liet zien dat er weinig verschil in schrijfvaardigheid was tussen beide momenten⁴.

1.3 Domeinbeschrijving en veldraadpleging: richtlijnen voor het peilingsonderzoek

De door Hoogeveen opgestelde domeinbeschrijving (2017), die voor de huidige peiling is opgesteld, beschrijft de doelen van het schrijfonderwijs (het wettelijke en referentiekader), resultaten van voorgaande peilingsonderzoeken en de uitvoering van het schrijfonderwijs in Nederland. Daarnaast geeft zij de resultaten van een veldraadpleging weer, zowel voor de te peilen vaardigheden als de te onderzoeken kenmerken van het onderwijsleerproces. Samengevat zijn de door de experts voorgestelde minimale richtlijnen voor subdomeinen, taken en leerlingkenmerken de volgende:

- “Dat de vaardigheden van deze leerlingen [leerlingen einde sbo] met behulp van dezelfde instrumenten gepeild kunnen worden als in het primair onderwijs” (p. 25).
- “De peiling meet het beheersingsniveau van leerlingen ten aanzien van het referentiekader taal voor het domein schrijfvaardigheid. De subdomeinen taalverzorging en grammatica (ondersteunende vaardigheden) en handschrift dienen niet apart gepeild te worden” (p. 26).
- “Voor schrijfvaardigheid dient te kunnen worden vastgesteld of en in welke mate leerlingen presteren op niveau 1F en 2F/1S” (p. 27).
- “De peiling meet minimaal het beheersingsniveau van leerlingen ten aanzien van de taken die in het referentiekader opgenomen zijn: 1) correspondentie, 2) formulieren invullen, berichten, advertenties en aantekeningen, 3) verslagen, werkstukken, samenvattingen, artikelen. Uit elke categorie dient minimaal 1 tekstsoort gepeild te worden”. Het verdient voorkeur om de schrijftaken te benoemen in termen van de verschillende functies van schrijven, zoals in eerdere schrijfpeilingen gebeurde (informerende, overtuigende etc.) (p. 27).
- “Bij de peiling dienen in principe alle kenmerken van de taakuitvoering meegenomen te worden: samenhang (op verschillende tekstniveaus en opgesplitst in inhoud en organisatie), afstemming op doel, afstemming op publiek (informeel en formeel taalgebruik), woordgebruik en woordenschat, spelling, interpunctie en grammatica (zie Taalverzorging: accuraat gebruik

³ Per item is de p-waarde berekend. Een p-waarde < .50 is geïnterpreteerd als onvoldoende, terwijl een p-waarde ≥ .80 een goede beheersing representeerde.

⁴ Voor meer specifieke informatie over de resultaten op de onderscheiden domeinen Inhoud, Opbouw en Structuur, Stijl en Afstemming op doel wordt verwezen naar Kuhlemeijer et al., 2013).

van eenvoudige zinsconstructies), leesbaarheid (titel, opmaak, handschrift, bladspiegel, beeldende elementen en kleur)” (p. 28). Het wordt daarbij van belang geacht “dat spelling, interpunctie, grammatica, en handschriftkwaliteit gepeild worden ten dienste van de communicatieve effectiviteit van teksten” (p.29).

- “Aan schrijffuncties die belangrijk geacht worden voor het vervolgonderwijs [wordt] een hoge prioriteit toegekend: argumenteren/overtuigen, activeren, informeren, uitleggen, instrueren, informeren, verzoeken, amuseren” (p. 28).

Ten aanzien van de beoordeling zijn geen strikte richtlijnen geformuleerd. Hiervoor kunnen de beoordelingsaspecten uit vorige peilingen (zoals inhoud, organisatie en structuur, doel- en publieksgerichtheid en stijl) of de kenmerken van de taakuitvoering binnen het referentiekader worden gehanteerd. Aan deze kenmerken zou volgens de experts het kenmerk ‘inhoud’ moeten worden toegevoegd. Inhoudselementen zijn noodzakelijk (of wenselijk) om te komen tot begrip van een tekst. Zij laten zien in hoeverre een leerling zich aan de gegeven opdracht houdt.

In de huidige peiling wordt grotendeels aangesloten bij de in de domeinbeschrijving genoemde richtlijnen. In deze domeinbeschrijving worden daarnaast nog “nadere specificaties ten aanzien van de setting” (p. 29-30) gegeven. Hierbij wordt onder andere ingegaan op het schrijfproces, meer specifiek op de revisievaardigheden. In de huidige peiling zijn per klas twee leerlingen in interviews bevestigd over het schrijfproces. Leerlingen is gevraagd hoe zij zich oriënteren op een taak, hoe zij hun schrijfproces reguleren en evalueren en *of en hoe* zij een voorbeeldtekst zouden reviseren. De feitelijke *revisievaardigheid* van alle leerlingen is echter niet gepeild, zoals in de nadere specificaties wel wenselijk werd geacht. Reden hiervoor was dat de afnametijd op een school beperkt is en wij primair per leerling meerdere schrijftaken met verschillende teksthandelingen wilden afnemen, zodat we voldoende producten zouden hebben om betrouwbare uitspraken te doen. Bovendien meenden wij aan de hand van een concreet voorbeeld in de interviews een behoorlijk beeld te kunnen geven van de aspecten die leerlingen zouden willen verbeteren.

1.4 Opzet huidige peiling: leerlingvaardigheden en –kenmerken

1.4.1 De schrijftaken

Wat betreft de te peilen vaardigheden sluit de huidige peiling op relevante aspecten bij de domeinbeschrijving aan. Allereerst zijn taken ontworpen die zowel in het basisonderwijs als in het speciaal basisonderwijs konden worden afgenomen. Geschiktheid voor beide doelgroepen is in een pilot nagegaan. Gepoogd is om taken te construeren die motiverend en zo authentiek mogelijk zijn, dat wil zeggen dat ze qua thematiek, herkenbaarheid en publiek aansluiten bij de belevingswereld van leerlingen. Ook wordt per taakomschrijving in het referentiekader minimaal één tekstsoort gepeild. Het gaat daarbij om de categorieën 1) correspondentie, 2) formulieren invullen, berichten, advertenties en aantekeningen, 3) verslagen, werkstukken, samenvattingen, artikelen, 4) vrij schrijven (verhalen, informatieve teksten, gedichten).

Voor de te peilen taken is gebruik gemaakt van een model met onderscheidingen in tekstfuncties en bijbehorende teksthandelingen, en tekstsoorten (genres) (De Gloppe, Minnaert, Van Gelderen, Pander Maat & Kuijpers, 2017). Tekstfuncties verwijzen naar effecten die bij de lezer nagestreefd worden; teksthandelingen betreffen tekstuele middelen die gebruikt kunnen worden voor het nastreven van die effecten (Pander Maat 2002). Gekozen is voor teksthandelingen die het meest

relevant lijken voor het taalonderwijs op het niveau einde (speciaal) basisonderwijs. Tekstfuncties en teksthandelingen zijn gekoppeld in voor leerlingen herkenbare, alledaagse genres. Iedere tekstfunctie wordt in de peiling door twee taken vertegenwoordigd, waarbij ervoor is gezorgd dat de vier onderscheiden schrijftaken in het referentiekader door minimaal één taak werden gespeild. Schrijftaak 1, correspondentie, wordt gespeild aan de hand van een brief; schrijftaak 2 aan de hand van een bericht; schrijftaak 3 aan de hand van een samenvatting en artikel; en schrijftaak 4 aan de hand van een verhaal en informatieve tekst.

De uiteindelijk ontworpen taken met bijbehorende genres, tekstfuncties, -handelingen en beoogd publiek staan beschreven in Tabel 1.5. Er is gebruik gemaakt van twaalf schrijftaken, waarvan er negen nieuw ontworpen zijn. Met behulp van de twaalf taken werden zes teksthandelingen getoetst; per teksthandeling zijn twee taken ontwikkeld. Om een trendvergelijking mogelijk te maken zijn daarnaast drie zogenaamde ankertaken aan de leerlingen voorgelegd. Dit zijn taken die ook in de vorige peiling zijn gebruikt. De taakinhoud en de beoordelingswijze was in deze peiling ongewijzigd ten opzichte van de peiling in 2009.

Tabel 1.5: Overzicht gebruikte taken in de peiling

Taak	Titel	Genre	Publiek	Tekstfunctie	Teksthandeling
T01 (anker)	Wim	Verhaal	Leeftijdgenoten	Amuserend	Vertellen
T02	Verboden boek	Verhaal	Ongespecificeerd	Amuserend	Vertellen
T03 (anker)	Unicef	Samenvatting/stukje op basis van tekst	Leeftijdgenoten	Informatief	Beschrijven
T04	Brief aan Annika	Brief	Leeftijdgenoot uit ander land	Informatief	Beschrijven
T05	Nepnieuws	Tijdschriftartikel	Leeftijdgenoten	Informatief	Uitleggen
T06	Vlinder	'Biologietekst'	Iemand zonder plaatjes erbij	Informatief	Uitleggen
T07	Pannenkoeken	Recept	Leeftijdgenoot uit ander land	Instructief	Instrueren
T08 (anker)	Poekie	Flyer over verdwenen huisdier	Buurtgenoten	Instructief	Instrueren
T09	Interviewhulp	Brief	Leeftijdgenoot	Directief	Verzoeken
T10	Speelveld	Brief	Burgemeester	Directief	Verzoeken (en argumenteren)
T11	Kortere vakantie	Brief	Schooldirecteur	Persuasief	Argumenteren
T12	Frisdrank en snoep	Stuk in de krant	Krantenlezers	Persuasief	Argumenteren

De taken kunnen als volgt worden beschreven:

T1: Wim (ankertaak)
De leerling moet op avontuurlijke wijze beschrijven hoe het kwam dat Wim met natte voeten en een scheur in zijn T-shirt maar zonder fiets thuiskwam
T2: Verboden boek
De leerling moet een spannend verhaal schrijven voor een schrijfwedstrijd. Daarbij is de start van het verhaal gegeven, dat door de leerling moet worden aangevuld. In het gegeven begin van het verhaal logeren twee kinderen bij oma, waar zij op zolder een boek vinden waarop staat dat dit niet geopend mag worden. De kinderen openen het boek toch en de leerling moet het verhaal afmaken.
T3: Unicef (ankertaak)
De leerling schrijft een eigen stukje voor de muurkrant over het werk van Unicef in arme landen. De opdracht is duidelijk te maken wat Unicef doet en wat de leerling daarvan vindt.
T4: Nederland
De leerling schrijft een brief aan een nieuw, Zweeds meisje dat in de klas komt. In deze brief moet de leerling antwoord geven op vijf vragen die het meisje over Nederlandse gebruiken heeft gesteld. De vragen gaan over hoe Nederland eruitziet, hoe dicht mensen op elkaar wonen, wat Nederlanders lekker vinden, van welk voedsel Nederlanders houden en wat Nederlanders in hun vrije tijd doen. Bij vier thema's wordt informatie over de situatie in Zweden gegeven.
T5: Nephieuws
De leerling schrijft een artikel voor een fictief informatief jeugdtijdschrift (geschikt voor 9-12-jarigen). Leerlingen lezen hiervoor twee korte teksten over nephieuws en moeten in hun artikel uitleg geven over wat nephieuws is, waarom het verspreid wordt en wat eraan gedaan kan worden.
T6: Vlinder
De leerling legt aan de hand van zes plaatjes de fases van de metamorfose van een vlinder uit. Deze uitleg wordt geschreven voor een medeleerling die plaatjes niet ziet. In de opdracht staat beschreven bij welk plaatje de leerling moet beginnen (de vlinder) en waar de leerling moet eindigen (weer dezelfde vlinder). Bij ieder plaatje wordt een kernwoord gegeven.
T7: Pannenkoeken
Leerlingen schrijven aan de hand van een ingrediëntenlijst en plaatjes, waarop de handelingen chronologisch staan weergegeven, een recept voor pannenkoeken. Dit recept is bedoeld voor een kinderkookboek, waaraan leeftijdgenoten uit andere landen ook een bijdrage leveren.

T8: Poekie (ankertaak)
Naar aanleiding van het verdwijnen van de poes Poekie schrijft de leerling een briefje aan buurtgenoten waarin Poekie beschreven wordt. De opdracht bevat daartoe enige relevante gegevens, evenals twee foto's van het dier. In verband met het terugbezorgen van Poekie dient de leerling in het briefje eveneens te vermelden hoe hij of zij te bereiken is.
T9: Interviewhulp
De leerling moet een interview houden over vroeger, toen mensen nog geen internet hadden. De leerling weet niet hoe hij/zij een interview moet houden en schrijft een briefje aan een voormalig klasgenoot, die ervaring heeft met interviews. De leerling dient deze voormalige klasgenoot beleefd om hulp te vragen en duidelijk te maken op welke vragen hij/zij antwoord nodig heeft.
T10: Speelveld
Leerlingen schrijven namens de klas een brief aan de burgemeester met het verzoek een speelveld te behouden en geplande huizen op een andere plaats te bouwen. De leerling dient uit te leggen hoe belangrijk dit speelveld voor ze is.
T11: Kortere vakantie
Leerlingen schrijven namens de klas een brief aan de directeur van de school, waarin zij beredeneren waarom zij een kortere zomervakantie geen goed idee vinden. Zij proberen de directeur aan de hand van zoveel mogelijk argumenten te overtuigen van dit idee af te zien. Enkele argumenten staan in de begeleidende tekst die vooraf door de leerlingen moet worden gelezen.
T12: Frisdrank en snoep
Leerlingen schrijven een stuk voor de krant waarin zij hun mening over het verbieden van de verkoop van frisdrank en snoep aan kinderen moeten geven en onderbouwen met argumenten. Een aantal mogelijke argumenten staan in de begeleidende tekst die de leerlingen vooraf moeten lezen. Leerlingen wordt expliciet gevraagd een titel toe te voegen.

1.4.2 Beoordeling schrijfproducten

Het uitgangspunt bij de modelontwikkeling was de keuze voor descriptieve schalen, zoals in de vorige schrijfpeiling (Kuhlemeier, Van Til, Hemker, De Klijn & Feenstra 2013). We citeren Kuhlemeier et al., pagina 41:

Bij het ontwerpen van een beoordelingschaal is het niet alleen van belang welke aspecten worden beoordeeld. Ook de manier waarop dat gebeurt, is bepalend voor de objectiviteit en betrouwbaarheid van de beoordelingen. In het algemeen worden drie typen beoordelingschalen onderscheiden: normatieve, descriptieve en productschalen (...). Bij normatieve schalen geven beoordelaars waarderingen aan teksten van leerlingen zoals: 'goed', 'matig' en 'zwak'. Het nadeel ervan is dat deze oordelen over het algemeen zeer subjectief zijn en de oordelen van de beoordelaars sterk uiteenlopen. Bij descriptieve schalen, ook wel rubrieken genoemd, worden de schaalpunten van inhoudelijke

beschrijvingen voorzien. Die geven de beoordelaar steun bij het beoordelen. Bij een productschaal, ook wel ankerschaal genoemd, zijn de scoreniveaus (d.w.z. de antwoordmogelijkheden) verduidelijkt aan de hand van complete schrijfproducten van leerlingen, ook wel ankerteksten genoemd. De ankerteksten zijn zorgvuldig geselecteerde teksten die de verschillende schaalpunten illustreren. De taak van de beoordelaars is dan te vergelijken met een sorteertaak. De beoordelaar legt het te beoordelen werk naast de ankers en bepaalt vervolgens met welke ankertekst het te beoordelen werk het meest overeenkomt. In de peiling schrijfvaardigheid is gekozen voor een descriptieve schaal. Elke schaal bestaat uit een eenvoudige, concrete vraag met een beperkt aantal antwoordmogelijkheden. De schaalpunten zijn voorzien van verhelderende toelichtingen waarin fragmenten uit 'echte' teksten van leerlingen verwerkt zijn. De keuze van deze fragmenten is gebaseerd op een vooronderzoek waarbij de schalen op een aantal werken van leerlingen zijn uitgetoetst. De voorbeelden fungeren als een soort anker dat de beoordelaars duidelijkheid geeft als ze er met behulp van de vragen alleen niet uitkomen.

Kuhlemeier et al. (2013) benaderen de schaalkeuze vanuit overwegingen rondom betrouwbaarheid. Maar er zijn veel meer goede redenen om met descriptieve schalen te werken. Allereerst houdt descriptief scoren de normen achter de beoordeling transparant. Doordat er wordt gewerkt met zo concreet mogelijke vragen over de aanwezigheid van tekstkenmerken in het schrijfproduct kan precies worden nagegaan hoe een oordeel tot stand is gekomen. Ten tweede geeft descriptief scoren meer informatie over wat leerlingen kunnen dan normatief scoren of scoren via productschalen. Ten derde maakt descriptief scoren het mogelijk om tekst- en stijlwetenschappelijke kennis in te zetten in de beoordeling van schrijfproducten. Tenslotte maakt descriptief scoren het makkelijker om de resultaten van verschillende peilingen door de jaren heen met elkaar te vergelijken.

Vanwege de genoemde redenen zijn wij doorgegaan op de weg ingeslagen in eerdere peilingen. Gepoogd is de schalen zo descriptief mogelijk te maken en de beoordeling van de teksten zo veel mogelijk te laten aansluiten bij die in de vorige peiling (Kuhlemeier et al. 2013), waarin de producten zijn beoordeeld op vier aspecten:

- Inhoudelijke kwaliteit (verder kortweg: inhoud)
- Organisatorische en structurele kwaliteit (verder: opbouw)
- Stilistische kwaliteit (verder: formulering)
- Communicatieve kwaliteit (verder: communicatie), door Kuhlemeier et al. (2013) uitgesplitst in doelgerichtheid en publieksgerichtheid.

Hoewel we ervoor hebben gekozen onze beoordelingsmodellen zoveel mogelijk te laten aansluiten bij de vorige peiling, is één van de doelen van deze peiling dat de leerlingresultaten kunnen worden geïnterpreteerd binnen het kader van de referentiekader voor schrijven. Daarom stelt Hoogeveen (2017, p. 26): *“Bij de peiling dienen in principe alle kenmerken van de taakuitvoering [binnen het referentiekader] meegenomen te worden:*

- a. *Samenhang (op verschillende tekstniveaus en opgesplitst in inhoud en organisatie)*
- b. *Afstemming op doel*
- c. *Afstemming op publiek (informeel en formeel taalgebruik)*
- d. *Woordgebruik en woordenschat*
- e. *Spelling, interpunctie en grammatica (zie Taalverzorging: accuraat gebruik van eenvoudige zinsconstructies)*
- f. *Leesbaarheid (titel, opmaak, handschrift, bladspiegel, beeldende elementen en kleur)”*

Met behulp van onze beoordelingsmodellen hebben wij getracht aan te sluiten bij de beoordelingswijze van de vorige peiling én een dusdanige koppeling te maken met de taakuitvoeringsaspecten in het referentiekader dat een standaardbepaling aan de hand van dit referentiekader mogelijk is. Om dit te doen hebben we binnen onze beoordelingsmodellen de kenmerken a, b, c, d en f geprobeerd te vangen onder de koepel van drie elementen – opbouw, formulering en inhoud – in plaats van de dat alle aspecten apart zijn beoordeeld. Hieronder volgt een verantwoording van onze keuzes. Opgemerkt dient te worden dat kenmerk e geen onderdeel uitmaakt van de beoordeling; ook deze keuze wordt hieronder verantwoord.

Communicatiekwaliteit grijpbaar maken via inhoud, opbouw en formulering

Uitgangspunt van onze beoordelingsmodellen is dat de communicatieve effectiviteit van de tekst, het centrale kenmerk van kwalitatief goede leerlingwerken binnen deze peiling, beoordeeld kan worden op basis van drie aspecten, namelijk inhoud, formulering en opbouw. Hiertoe zijn wij gekomen omdat wij bij de ontwikkeling van de beoordelingsmodellen eerst hebben stilgestaan bij een *principiële kwestie*, namelijk: zijn er eigenlijk communicatieve tekstkwaliteiten die *niet* te definiëren zijn als adequaatheid van inhoud, opbouw en formuleringkeuzes? Wanneer we de kwaliteit van een tekst opvatten als effectiviteit voor de doelgroep (zoals beoogd in deze peiling), dan spreken we over zaken zoals geslaagde informatieoverdracht (begrip), attitudeverandering ten aanzien van het bepleite standpunt of van de bepleite actie, positieve attitudes ten aanzien van de gepastheid van de boodschap en van de auteur, of over emotionele effecten van verhalen (de lezer leeft mee). Dit soort effecten komen tot stand via de inhoud, de opbouw en de formulering van de tekst. Onze benadering betekent dat eisen wat betreft inhoud, opbouw en formulering gezien worden als een inschatting van de doel- en publieksgerichtheid van de tekst⁵. En die doel- en publieksgerichtheid zijn tezamen niets anders dan de effectiviteit van de tekst. Het beoogde effect van een tekst is namelijk het bereiken van het tekstdoel bij de lezer. Om deze reden achten wij een extra beoordelingsaspect ‘communicatieve kwaliteit’, zoals gehanteerd in de vorige peiling, niet noodzakelijk en niet gewenst.

De drie aspecten Opbouw, Formulering en Inhoud komen niet als zodanig terug in de taakuitvoeringskenmerken binnen het referentiekader. Echter, ze kunnen wel aan elkaar worden gerelateerd. De beschrijvingen binnen kenmerk a (‘samenhang’) worden binnen onze beoordelingsmodellen geschaard onder de opbouw van een tekst. Het betreft immers aspecten als ordening, samenhang en verwijzingen. Onder *doelgerichtheid* verstaan wij de mate waarin inhoud, opbouw en formulering/stijl dienstbaar zijn aan het gestelde schrijfdoel. Dit betekent dat we willen nagaan in hoeverre de door ons bevraagde kenmerken in de beoordelingsmodellen bijdragen aan de begrijpelijkheid, de informatiewaarde, de overtuigingskracht en de amusementswaarde van de tekst. Wanneer wij dan kijken naar kenmerk b in het referentiekader (‘afstemming op doel’) dan blijkt dat referentieniveau 1S/2F⁶ veronderstelt dat de schrijver in staat is een tekst te schrijven met een eenvoudige lineaire structuur waarin hij/zij trouw blijft aan het schrijfdoel van de tekst. Om het beoogde schrijfdoel te bereiken is het gebruik van inhoudselementen essentieel. Bij het opstellen van een bericht over een vermiste poes is het bijvoorbeeld essentieel dat de lezer geïnformeerd wordt over het feit dat de poes kwijt is, over hoe de verloren poes er uit ziet en dat er een verzoek wordt gedaan over hoe te handelen bij het vinden van de poes (namelijk de eigenaar van de poes informeren en op welke wijze dit kan). Dergelijke inhoudelijke informatie is noodzakelijk om het schrijfdoel te

⁵ In de vorige peiling was ‘stijl’ één van de aspecten die beoordeeld werden. In de huidige peiling hanteren wij ‘formulering’.

⁶ Voor niveau 1F worden geen taakuitvoeringskenmerken beschreven voor ‘afstemming op doel’.

bereiken. Met *publieksgerichtheid* (kenmerk c, afstemming op publiek) zijn wij als volgt omgegaan: het beoogde publiek (de lezer) is een onderdeel van het schrijfdoel, dus is er geen reden om publieksgerichtheid los daarvan te definiëren. Er zijn voor alle in de schrijftaken betrokken genres tekstelementen geïdentificeerd die laten zien hoe rekening wordt gehouden met het publiek. Daarbij zijn verschillende aspecten van belang:

- Het rekening houden met de voorkennis van de lezer, zoals blijkend uit het al of niet toelichten van termen.
- De manier waarop de schrijver omgaat met het gezicht ('face', zie Brown & Levinson, 1987) van de ontvanger, zoals blijkend uit het gebruik van beleefdheidsstrategieën en het gekozen formaliteitsniveau (zie Pander Maat 2002, 208-212 over 'gepastheidsondersteuning').
- Uitingen waarin contact wordt gemaakt met de ontvanger, zoals het aanspreken en verwelkomen van de lezer.
- Uitingen waarmee de lezer wordt gemotiveerd de tekst te lezen doordat de tekstinhoud aantrekkelijk wordt gepresenteerd.
- Meta-communicatieve uitingen die hulp bieden aan de lezer door een vooruitblik of een terugblik op de tekst (zie Pander Maat (2002, p. 202-208) over oriënterende uitingen en uitingen over het belang van of de aanleiding voor de tekst).

Samenvattend gaat het bij publieksgerichtheid om het werken aan een goede relatie met de lezer, gegeven een bepaalde inhoudelijke boodschap. Dit betekent, zoals bovenstaande opsomming laat zien, dat passende formuleringen cruciaal zijn voor een afstemming op publiek. Daarnaast spelen echter ook de aspecten opbouw (denk aan gebruik van aanhef en groet in correspondentie) en inhoud (zoals bijvoorbeeld een inhoudelijke toelichting bij een lastig verzoek) een rol bij de afstemming op publiek.

De manier waarop de kenmerken d en f (resp. 'woordgebruik en woordenschat' en 'leesbaarheid') zijn uitgevoerd zijn evenals de kenmerken a – c van invloed op de communicatieve effectiviteit van de tekst en zijn in beschouwing genomen voor zover ze bijdragen aan het realiseren van a, b & c. Het woordgebruik van de leerling komt daarmee in de modellen aan bod voor zover het bijdraagt aan de kwaliteit van formulering van de tekst, en daarmee het slagen in het communicatieve schrijfdoel. Immers, op het moment dat het woordgebruik onvoldoende is, bereikt een tekst over het algemeen ook niet het communicatieve doel. Ook leesbaarheid is essentieel voor het bereiken van het schrijfdoel. Hieronder vallen aspecten als het gebruik van alinea's en opsommingstekens. Deze scharen wij onder de opbouwweisen. Dit geldt ook voor titels, omdat titels een vast onderdeel zijn van bepaalde genres. Verder is de leesbaarheid van het handschrift als een eliminatiecriterium vooraf geformuleerd (zie hieronder). Er is afgezien van kwalificaties van de overige visuele elementen.

Ook voor het laatste taakuitvoeringskenmerk (e), 'spelling, interpunctie en grammatica', zijn geen aparte beoordelingseisen opgesteld. Op basis van eerdere onderzoeksbevindingen en de resultaten van een raadpleging onder vier schrijfexperts is besloten correctheidsaspecten niet in de vaardigheidsmeting op te nemen. Hiervoor golden enkele zwaarwegende argumenten. Allereerst, en cruciaal, is taalverzorging niet te meten aan de hand van de in de peiling gehanteerde schrijfoopdrachten, aangezien de omschrijving van taalverzorging in de referentieniveaus niet bij onze schrijfoopdrachten aansluit. Voor de kenmerken van de taakuitvoering voor spelling, interpunctie en grammatica wordt in het referentiekader voor het domein schrijfvaardigheid verwezen naar de kenmerken behorend onder Taalbeschouwing/Taalverzorging. Deze richtlijnen zijn voor spelling en

interpunctie zeer specifiek: ze betreffen verschillende categorieën van spellingproblemen en -regels en verschillende leestekens, die door de gehanteerde taken niet per definitie worden uitgelokt. Bovendien laten (vrije) schrijftaken toe dat leerlingen spellings- en grammaticale problemen vermijden, waardoor ze niet geschikt zijn om beheersing van die aspecten te meten. Ten tweede is taalverzorging van een andere aard dan ‘communicatieve effectiviteit’, hetgeen een kernbegrip in deze peiling is. In deze peiling wordt leerlingen gevraagd een verhaaltje te schrijven, een situatie/onderwerp helder uit te leggen, zich in lezers te verplaatsen. Kennis van schrijftaalconventies is een geheel andere kwestie. Bovendien vraagt de veldraadpleging niet om een aparte beoordeling van taalverzorging, maar zouden aspecten van taalverzorging alleen ten dienste van de communicatieve kwaliteit moeten worden beoordeeld. Binnen deze peiling wordt de kwaliteit van de taalverzorging impliciet meegenomen in de andere aspecten waarmee communicatieve effectiviteit gemeten wordt. Net als leesbaarheid en woordgebruik kunnen grove correctheidsfouten de communicatieve effectiviteit in de weg staan op het moment dat ze de navolgbaarheid en begrijpelijkheid van de geschreven tekst onmogelijk maken – verondersteld wordt dat zij op dat moment de aspecten inhoud, opbouw en formulering beïnvloeden. Tenslotte zou het meenemen van taalverzorging in de beoordeling naar verwachting een valide toetsing van de communicatieve vaardigheid van sbo-leerlingen belemmeren. Gezien deze argumenten worden in de huidige peiling correctheidsaspecten (interpunctie, spelling, grammatica) weliswaar gescoord en wordt hierover ook gerapporteerd (zie hoofdstuk 8.4), maar zijn zij niet meegenomen in de bepaling van het schrijfvaardigheidsniveau (m.a.w. ze worden niet meegenomen in de schaling) en niet in de standaardbepaling.

Hoofdpijnen in de beoordelingsmodellen

Zoals hierboven verantwoord, wordt de kwaliteit van de leerlingwerken in deze peiling bepaald aan de hand van inhouds-, formulerings- en opbouwaspecten. Met behulp van vragen wordt de aanwezigheid van bij de taak passende kenmerken van deze drie aspecten bevraagd. Deze vragen vormden de kern van ieder beoordelingsmodel. Daarnaast bevatte ieder beoordelingsmodel enkele vragen vooraf en vragen achteraf. Deze drie onderdelen worden hieronder besproken.

De kern: opbouw, formulering en inhoud

Op basis van bovenstaande uitgangspunten zijn beoordelingsmodellen gemaakt voor de negen nieuw ontwikkelde taken (de beoordeling van de ankertaken volgt hieronder). Bij de ontwikkeling van de beoordelingsmodellen is gebruikgemaakt van schrijfproducten uit pilotonderzoek (begin 2019).

Zoals Tabel 1.5 laat zien verschillen de taken onderling sterk. Enkele belangrijke verschillen zijn:

- Sommige taken vragen reproductie, parafrase of samenvatting van informatie uit de opdracht, andere taken vragen om ‘eigen inhoud’.
- De complexiteit van de thematiek verschilt per taak. Bepaalde taken blijven thematisch dicht bij de belevingswereld van de leerlingen, anderen bieden abstractere onderwerpen aan.
- Er valt te onderscheiden tussen persoonlijk gerichte boodschappen (brieven), teksten voor tijdschriften of kranten, teksten voor leeftijdgenoten (recept, tekst over vlinder) en een verhaal zonder omschreven publiek.
- Er zijn vijf verschillende doelen en zes verschillende teksthandelingen aan de orde.

Per taak verschilde daarom wat als een doel- en publieksgerichte inhoud, opbouw en formulering beschouwd kan worden. Gegeven het uitgangspunt van descriptieve schalen leidde dat ertoe dat de

modellen per taak verschillende kenmerken aan de orde stellen of soms verschillende waarden van eenzelfde kenmerk (bv. het kenmerk formaliteit). In een verhaal (T2) is bijvoorbeeld spanning van belang; in de taken T10, T11 en T12 zijn argumenten van belang; in een recept (T7) is informatie van belang over de duur van de verschillende bereidingsfasen; in brieven (bv. T4) is aandacht voor de relatie met de ontvanger van belang, enzovoort.

De eisen per taak variëren dus, maar er is per aspect wel sprake van terugkerende eisen. Onder **inhoud** bevragen de modellen het gebruik van informatie uit de taakomschrijving (soms gaat het om specifieke informatie, soms gaat het om te bespreken thema's), en de aanwezigheid van voor het genre verplichte elementen (zoals standpunt en argumenten in de betogende brief T11 of ingrediëntenlijst en handelingsbeschrijvingen in het pannenkoekenrecept T7). Het gaat bij inhoud altijd om de *aanwezigheid van elementen*. Onder **opbouw** bevragen de modellen de plaatsing van vaste tekstelementen, zoals aanhef en afsluitende groet in de brief, titels, inleidende zinnen, afsluitende zinnen, markering van relaties en overgangen, en volgordekwesties. Het gaat dus bij opbouw om *elementen aan het begin of aan het eind van de tekst, om relaties en overgangen, en om volgorde*. Onder **formulering** bevragen de modellen de formaliteit van de tekst of tekstelementen, die afhangt van de relatie tussen schrijver en ontvanger. Verder gaat het soms over de aanwezigheid van details die in verhalen de amusementswaarde vergroten en in instructies de uitvoerbaarheid. Ten slotte gaan de modellen in op beleefdheidselementen. Daarbij gaat het om elementen als het ruimte bieden aan de ontvanger en het vriendelijk en positief tegemoet treden van de ontvanger. In Bijlage 1 wordt een overzicht gegeven van de soorten items per taak.

Bij de vaststelling van de eisen is gestreefd naar zo concreet mogelijke formuleringen. Zo zijn vragen met evaluatieve adjectieven zoals 'de tekst is begrijpelijk/ logisch opgebouwd /overtuigend' vermeden. Bij inhoud betreffen alle eisen specifieke subthema's of inhoud, waardoor er geen eisen zijn als: 'de tekst biedt voldoende informatie om de taak uit te voeren'. Bij opbouw gaan de eisen over specifieke tekstcomponenten (bv. aanhef), relaties daartussen (bv. 'link tussen standpunt en argument(en)') of specifieke relaties (bv. chronologie). Bij formulering, tenslotte, zijn altijd specifieke formuleringstrategieën genoemd. Daarnaast zijn voorbeelden toegevoegd om de bevroegde kwaliteiten grijpbaar te maken.

Voor ieder van de negen nieuwe taken is aldus een set criteria vastgesteld, gebaseerd op de schrijfdoelen voor de specifieke taak. Voor iedere taak is een scoreformulier ontworpen bestaande uit de taakgerichte set criteria aangevuld met een aantal algemene vragen die voor alle taken dezelfde waren (zie hieronder onder de kopjes kenmerken die vooraf en achteraf worden bevroegd). In Bijlage 2 is één van deze modellen ter illustratie gegeven. In paragraaf 8.3 wordt nader op de items (elementen) van die taak (Taak 4) ingegaan. Binnen de taakgerichte set criteria werden drie beoordelingsaspecten onderscheiden: aspecten gericht op Opbouw, Formulering en Inhoud (OFI). Per item (OFI-criteria) diende de beoordelaar enkel te noteren of het gevraagde in de leerlingtekst aanwezig was. Om het coderen van de tekstmodellen zo eenvoudig mogelijk te maken zijn de items over eenzelfde tekstonderdeel zo mogelijk bij elkaar gezet. Dat betekent dat er in de uiteindelijke vragenlijsten geen groepering is naar inhoud, opbouw en formulering. Deze termen kwamen ook niet voor in de codeerlijsten.

Kenmerken die vooraf worden bevroegd

Voorafgaand aan iedere beoordeling moest eerst worden nagegaan of het zinvol zou zijn om de tekst op de vastgestelde items te beoordelen. In het geval van onleesbare teksten en teksten zonder zinnen

was het niet mogelijk de tekst te beoordelen op de gehanteerde criteria, waardoor er geen uitspraak over de kwaliteit van de tekst zou kunnen worden gedaan. Alle beoordelingsformulieren startten daarom met twee algemene vragen, die voor alle nieuwe taken dezelfde waren. Deze twee exclusie-/selectievragen bevroegen dus de voorwaarden waaraan teksten moesten voldoen wilden ze beoordeeld worden. Teksten zijn niet beoordeeld als ze:

1. in onleesbaar handschrift geschreven zijn;
2. geen lopende zinnen bevatten.

De derde vraag betrof de vraag of de leerling de taak anders had opgevat dan bedoeld. Dit was geen exclusie-/selectievraag, maar het anders opvatten van de opdracht zou wel invloed kunnen hebben op de score voor de taak. Zoals in hoofdstuk 5 zal blijken is nagegaan hoe hoog het percentage teksten was waarbij de leerlingen de taak anders hadden opgevat. Ook is nagegaan of er verschillen in scores tussen leerlingen waren waarvan de beoordelaar stelde dat de leerling de opdracht anders had opgevat dan bedoeld.

Kenmerken die achteraf worden bevroegd

Na afloop van de beoordeling van de OFI-kenmerken is beoordelaars gevraagd nog enkele vragen beantwoorden. Ook deze vragen achteraf waren voor alle taken dezelfde.

Allereerst werd om een globaal oordeel van de geschreven tekst gevraagd: ‘wat vind je van de kwaliteit van de tekst?’ Beoordelaars konden kiezen uit vier antwoordcategorieën (slecht, onvoldoende, voldoende, goed). Deze globale vraag diende met name ter validering van de descriptieve schalen. We hebben ervoor gekozen om die algemene vraag te laten volgen op de specifieke (OFI-)vragen om te voorkomen dat het algemene oordeel van invloed zou zijn op de scoring van de OFI-criteria.

Een ander kenmerk dat is bevroegd betrof een inschatting van het aantal regels van de (body)tekst. Omdat handmatig tellen van woorden te bewerkelijk is, is er hierbij voor gekozen om de beoordelaar globaal het aantal regels van de bodytekst te laten opgeven. Ook hier kon worden gekozen uit vier categorieën: 1-5 regels, 6-10 regels, 11-15 regels en meer dan 15 regels. Uit eerder onderzoek is gebleken dat de lengte van een tekst positief correleert met tekstkwaliteit. Het meenemen van de lengte van de tekst zou ook ter validering van de descriptieve schalen kunnen dienen⁷. In hoofdstuk 8 (§ 8.4) wordt de relatie tussen tekstlengte en de vaardigheid (op basis van de descriptieve schalen) beschreven.

Een laatste categorie kenmerken die achteraf zijn bevroegd betreft correctheidsaspecten. In elk beoordelingsmodel is aandacht besteed aan fouten in spelling, interpunctie en grammatica. Omdat deze drie soorten fouten nogal verschillen van karakter, zijn deze drie correctheidsaspecten apart bevroegd en gescoord. In de beoordelingsmodellen is de volgende lijn aangehouden:

- Bij de spelfouten zien we af van fouten rondom het aaneen of los schrijven van samenstellingen.
- Bij interpunctie beperken we ons tot het ernstigste probleem, namelijk het ontbreken van het segmenteren van zinnen. Meer concreet gaat het hier om het ontbreken van punten tussen zinnen of deelzinnen.

⁷ Dit is uiteindelijk niet gebeurd.

- Bij grammatica spreken we van ‘niet-lopende constructies’. We hebben ons binnen deze categorie beperkt tot incomplete zinnen (met ontbrekende woorden, zoals *In deze wereld wonen allemaal heksen tovenaars enge beesten*), verkeerd (evt. onconventioneel) taalgebruik (bijvoorbeeld *En dat opa bij de begraafplaats een hartstilstand kreeg was van die map*) en ontspoorde zinnen (zoals *Toen kwamen er allemaal portalen met allemaal zombies er uit*).

Er is gekozen voor deze minimale werkwijze om te voorkomen dat het scoren op correctheid te veel tijd en aandacht van de beoordelaars zou kosten, waardoor de focus op de drie inhoudelijke aspecten mogelijk zou afnemen. Per correctheidsvraag golden dezelfde drie antwoordcategorieën: (bij) meer dan de helft van de zinnen, (bij) minder dan de helft van de zinnen, (bij) geen enkele zin. Naar verwachting zouden deze globale inschattingen voldoende indicatie van de mate van correctheid van de teksten geven. We hebben ervoor gekozen de beoordelaars aan het eind van het beoordelingsformulier een inschatting voor deze correctheidsaspecten te laten geven onder de veronderstelling dat deze inschattingen over correctheid de codering van de OFI-items zo min mogelijk zouden beïnvloeden.

De ankertaken

Met het oog op de vergelijkbaarheid van de resultaten zijn drie taken uit de vorige peiling overgenomen. Het gaat om T1 (‘Wim’), een vertel-taak, T3 (‘Unicef’), een uitleg-taak en T8 (‘Poekie, de verdwenen kat’), een instruerende taak. Voor de scoring van de drie ankertaken wordt gebruikgemaakt van de beoordelingsmodellen die daarvoor eerder zijn gebruikt. Zoals eerder beschreven waren de items in deze beoordelingsmodellen eerder ondergebracht in vier categorieën: inhoud, opbouw, stijl en communicatieve kwaliteit. In de analyse van de resultaten zijn de items uit deze vragenlijsten ingedeeld onder inhoud, opbouw en formulering. De meeste items konden daarbij één op één worden geclassificeerd onder inhoud, opbouw en formulering⁸. Gezien onze opvatting dat er eigenlijk geen communicatieve tekstkwaliteiten zijn die *niet* te definiëren zijn als adequaatheid van inhoud, opbouw en formuleringkeuzes, zijn de items, die in 2009 onder communicatieve kwaliteit vielen, in de meest geëigende OFI-categorieën geplaatst. Ook de ankertaken kenden beoordelingsmodellen met vragen vooraf en achteraf. Vooraf was er één selectievraag (‘Is de tekst leesbaar?’). Daarnaast was één vraag gericht op het af zijn van de taak (‘Is de tekst af?’) en één vraag gericht op een andere taakopvatting dan bedoeld (‘Heeft de leerling zich aan de opdracht gehouden?’). Beide vragen golden ook in 2009 niet als exclusievragen. De vragen achteraf waren gericht op algemene indruk, een inschatting van het aantal woorden (tekstlengte) en een indicatie van het aantal schrijffouten (veel; niet veel maar ook niet weinig; weinig of geen).

⁸ Deze formulering was gevat onder ‘stijl’ in de vorige peiling.

1.4.3 Meting leerlingkenmerken

Voor het verkrijgen van leerlingkenmerken is gebruik gemaakt van een vragenlijst, interview en beschikbare informatie uit de schooladministratie (leerlinggewicht⁹, leeftijd, sekse, dyslexieverklaring). De leerlingvragenlijst bevatte vragen over de achtergrond van de leerling (land van herkomst en thuistaal), vragen over schrijven op de mobiele telefoon, over hoe vaak de leerling thuis schrijft (en wat voor teksten dan) en over hoe leuk en moeilijk de leerling schrijven vindt (Bijlage 3A). In een verdiepend onderzoek zijn daarnaast interviews bij leerlingen afgenomen – zie paragraaf 1.6.2 voor meer informatie. De leidraad staat weergegeven in Bijlage 3B.

1.5 Opzet huidige peiling: onderwijsleerproces

1.5.1 Kenmerken van het schrijfonderwijs

Al decennialang is er onvrede over het schrijfniveau van leerlingen en zijn er twijfels over de effectiviteit van de schrijflessen. Van Gelderen en Blok (1991) hebben een eerste studie naar de lespraktijk bij stellingen op Nederlandse basisscholen uitgevoerd, waarbij onder andere door middel van observaties de tijdsverdeling in schrijflessen is gemeten. Uit de observatieresultaten bleek dat bijna 2/3 van de lessen gericht was op voorbereiding en het schrijven van een eerste tekstversie. De overige 1/3 was gewijd aan het voorlezen van de teksten voor en door leerlingen en het geven van commentaar op deze teksten (16%), het (vaak klassikaal) evalueren van de opdracht zelf en van de gemaakte teksten (10%) en aan organisatorische aspecten, die niet gerelateerd waren aan de schrijfoopdrachten (7%). Opmerkelijk was dat slechts 2% van de tijd werd besteed aan revisie, waarbij enkel werd ingegaan op spelling, grammatica, woordkeus en interpunctie. Nadien zijn nog enkele (kleinschalige) studies naar de praktijk van stellingen in het bo gedaan, zoals bijvoorbeeld door de Inspectie van het Onderwijs (2012) en Franssen & Aarnoutse (2013), maar inzicht in onderwijsleerproces was slechts fragmentarisch (Rietdijk, Van Weijen, Janssen, Van den Bergh & Rijlaarsdam, 2018). Vanwege dit fragmentarische karakter van de kennis over het Nederlandse schrijfonderwijs hebben Rietdijk en collega's een baseline studie uitgevoerd met als doel grondig inzicht te krijgen in de praktijk binnen de schrijflessen in de bovenbouw van het po, met gebruikmaking van een variëteit aan instrumenten (vragenlijsten, stimulated recall interviews en lesobservaties).

Het beeld dat uit de verschillende onderzoeken naar voren komt is dat het Nederlandse schrijfonderwijs drie problemen kent: 1) beperkte onderwijstijd voor schrijven, 2) onvoldoende voorbereiding voor toekomstige leerkrachten door lerarenopleidingen en 3) onvoldoende gebruik van evidence-based instructiemethodieken (Rietdijk et al., 2018). De geringe tijdsbesteding, het eerste probleem, blijkt bijvoorbeeld uit de resultaten van de vorige peiling. Kuhlemeijer en collega's (2013) vonden bijvoorbeeld dat slechts 17-18% van de tijd besteed aan taal in de bovenbouw gericht was op schrijven, terwijl er 26% en 28% van de tijd werd besteed aan respectievelijk lezen en spelling. Hoogeveen (2017) noemt daarbij dat er, naast tijdgebrek, ook sprake kan zijn van een gebrek aan prioriteit. Verder blijken leerkrachten tijdens de schrijflessen weinig effectieve (schrijf)instructie te geven (Inspectie van het Onderwijs, 2012; Hoogeveen, 2017). Ook de recente studie van Rietdijk en

⁹ Voor de definitie van de gewichten wordt onder andere verwezen naar <https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2006/04/26/nieuwe-gewichtenregeling-basisonderwijs/brochuregewichtenregeling.pdf>

collega's (2018) laat zien dat de ingeroosterde tijd voor schrijfonderwijs beperkt is, want 75% van de leerkrachten zegt in ieder geval twee lessen per maand aan schrijven te besteden – het minimum voorgeschreven aantal schrijflessen volgens de Inspectie van het Onderwijs. Het door de leerkrachten genoemde gemiddelde aantal lessen per maand kwam neer op 2.6, met een range van 0.3 tot 7 keer per maand. De observaties lieten echter wel zien dat leerlingen tijdens de schrijfles voldoende taakgericht bezig waren. Het tweede probleem is dat het twijfelachtig is of basisschoolleerkrachten dusdanig worden opgeleid dat zij in staat zijn goede schrijflessen te verzorgen. Op de Pabo wordt weinig tijd besteed aan schrijven, de didactiek is onvoldoende en aankomende leerkrachten ontvangen amper feedback op hun eigen schrijfvaardigheid (Koster et al., 2017; Van der Leeuw, Pauw, Smits & Van de Ven, 2012). Het derde probleem in het schrijfonderwijs is dat er onvoldoende gebruik wordt gemaakt van bewezen effectieve interventies. Dit ondanks dat er nationaal en internationaal tal van meta-analyses zijn verschenen, die consistente bevindingen laten zien over wat effectieve interventies zijn. Zo vinden Janssen en Van Weijen (2017) in hun reviewstudie naar de resultaten van 26 meta-analyses over effectieve schrijfinterventies (vanaf 1984) vier benaderingen, die de schrijfvaardigheid van leerlingen (po en vo) aantoonbaar bevorderen: instructie in schrijfstrategieën, het stellen van product- en/of procesdoelen, het samen schrijven en de feedback van de leerkracht of anderen. De meta-analyse van Koster en collega's (2015) richtte zich specifiek op effectieve schrijfinterventies voor leerlingen in de bovenbouw van het po. Zij vinden vergelijkbare resultaten: de genoemde vier interventies, aangevuld met instructie in tekststructuur. Het stellen van product- en procesdoelen kwam in deze meta-analyse als meest effectieve interventie naar voren.

Kenmerkend voor het Nederlandse schrijfonderwijs is overigens dat de taalmethode leidend is voor de leerkrachten en dat is al jaren het geval (Van Gelderen & Blok, 1991; Hoogeveen, 2017). Deze methodes sluiten weliswaar aan bij de kerndoelen, maar tot op heden is onduidelijk in hoeverre de methodes een vertaalslag naar de referentieniveaus maken. Hiertoe staan leerkrachten wel de in 1.3.1 genoemde praktische vertalingen van kerndoelen naar leerlijnen ter beschikking. Desondanks laat het onderzoek van Regioplan zien dat leerkrachten niet goed weten hoe ze met de referentieniveaus kunnen werken (Dekker et al., 2012).

1.5.2 Richtlijnen voor de meting van het onderwijsleerproces

Om zicht te krijgen op effectieve kenmerken van het schrijfonderwijs hebben Janssen en Van Weijen (2017) de eerdergenoemde review uitgevoerd. De review leverde een overzicht van kenmerken en benaderingen op die de schrijfvaardigheid van leerlingen in het basisonderwijs bevorderen. Daarbij maakten Janssen en Van Weijen een onderscheid tussen relevante algemene en domeinspecifieke aspecten van het onderwijsleerproces. In de domeinbeschrijving van Hoogeveen (2017), met daarin richtlijnen voor de peiling, wordt dit onderscheid ook aangehouden. Daarnaast beveelt Hoogeveen aan minimaal de volgende achtergrondgegevens van de school te verzamelen: het opleidingsniveau van de ouders (SES, te meten als percentage leerlingen met een formatiegewicht op een school) en de thuistaalsituatie. Deze thuistaalsituatie wordt in deze peiling bij de leerlingen bevraagd. Op schoolniveau worden tevens schoolgrootte en regio in kaart gebracht.

De experts stellen voor minimaal vier *algemene factoren*¹⁰ te meten: 1) directe instructie (werken met en delen van expliciete en heldere leerdoelen), 2) activering van leerlingen (bijvoorbeeld door samenwerking, het stimuleren van leerlingen om actief met de les mee te doen, bevordering van

¹⁰ Hierin volgen wij de domeinbeschrijving die een onderscheid maakt tussen algemene en domeinspecifieke factoren. Opgemerkt dient te worden dat het in beide gevallen gaat om een toepassing in het schrijfonderwijs.

leesplezier), 3) opbrengstgericht werken (systematische monitoring van voortgang in schrijfvaardigheid en formatieve evaluatie) en 4) opvattingen van leerkrachten.

Voorgesteld wordt om minimaal de volgende domeinspecifieke factoren te meten: 1) onderwijstijd voor schrijven, 2) het gebruik van proces- en productdoelen, 3) leerstof (aandacht voor aspecten van de taakuitvoering en aandacht voor het schrijfproces), 4) functies van teksten (aandacht voor communicatief doel, publiek), 5) modeling, 6) feedback, 7) het leren van anderen, 8) integratie met andere vakken en 9) ondersteuning binnen de school en professionalisering.

Genoemde factoren zijn bevraagd middels school-/leerkrachtvragenlijsten en aanvullend door het houden van interviews met leerkrachten (verdiepend onderzoek, zie paragraaf 1.6). De schoolvragenlijst bevatte vragen over het taalbeleid van de school, over schoolbrede aandachtspunten binnen het schrijfonderwijs en de visie op schrijfonderwijs (Bijlage 4). Directeuren of taalcoördinatoren werden verzocht deze vragenlijst in te vullen. De leerkrachtvragenlijst ('aanbodvragenlijst') richtte zich op vier onderdelen: algemene achtergrondinformatie, onderwijstijd voor schrijven en gebruikte methoden en materialen, de schrijfproducten waar leerlingen aan werken en de manier waarop het schrijfonderwijs in de klas wordt vormgegeven (zie Bijlage 5A). De interviews kenden een vaste opbouw, waarin gevraagd werd naar de visie van de leerkracht op schrijfonderwijs, het leerplan, de praktijk van de schrijflessen en het volgen en beoordelen van de schrijfvaardigheid van de leerlingen (zie Bijlage 5B). Daarnaast werd leerkrachten gevraagd naar hun sterke kanten waar het gaat om het lesgeven in schrijfvaardigheid en belemmeringen die zij ervaren bij het geven van schrijfonderwijs. In paragraaf 1.6.2 wordt verder ingegaan op de inhoud van het interview. Tenslotte is met behulp van observaties, aan de hand van het instrument van Van Gelderen en Blok (1991), de tijdsverdeling in de schrijflessen in kaart gebracht. In hoofdstuk 6 worden deze resultaten beschreven. Voorts zijn aspecten van het onderwijsleerproces meegenomen in de meerniveau-analyses, waarmee de relaties tussen verschillende leerling-, school-, leerkracht- en instructiekenmerken en de leerlingprestaties zijn geschat (zie hoofdstuk 9).

1.6 Verdiepend onderzoek

Naast het hoofdonderzoek heeft een verdiepend onderzoek plaatsgevonden met als doel om via observaties en interviews met leerkrachten het onderwijsleerproces op de deelnemende scholen – aanvullend op de ontvangen school- en leerkrachtvragenlijsten - nader in kaart te brengen. Streven was om in iedere deelnemende klas een observatie van een schrijfles uit te voeren, een interview met de betreffende leerkracht te houden en twee leerlingen individueel te interviewen. In het bo is het onderzoek uitgevoerd door onderwijsinspecteurs die in de rol van onderzoeker met behulp van het door de onderzoekers aangeleverde instrumentarium informatie over het schrijfonderwijs hebben verzameld. In het sbo hebben studenten orthopedagogiek de schoolbezoeken uitgevoerd.

Uiteindelijk zijn 85 van de 95 bo-scholen bezocht (89%), terwijl er 22 van de 38 aangemelde sbo-scholen zijn bezocht (58%).¹¹ De sbo-scholen zijn bezocht door (master)studenten orthopedagogiek. Het moment van afname viel voor deze studenten in een periode van afronding van hun studie, waardoor het zeer moeilijk was geschikte interviewers/observatoren te vinden. Dit heeft

¹¹ In hoofdstuk 6 (paragraaf 6.2) gaan we nader in op de definitieve aantallen geïnterviewde leerkrachten en leerlingen en de aantallen geobserveerde lessen die beschikbaar waren voor analyses.

ertoe geleid dat niet alle scholen bezocht konden worden. Wel is ernaar gestreefd zorg te dragen voor een representatieve afspiegeling van de sbo-scholen naar regio.

Het onderzoek binnen het bo diende te voldoen aan de volgende randvoorwaarden:

- Volledige scheiding van *regulier inspectietoezicht op de scholen en het kwalitatieve onderzoek*. De betrokken inspecteurs zijn niet verbonden aan het reguliere toezicht op de betreffende school én er heeft geen communicatie plaatsgevonden met de ‘regulier toezichthoudende’ inspecteur over de bevindingen tijdens het schoolbezoek.
- Er wordt zowel tijdens/na de lesobservatie als ook tijdens het interview *geen oordeel* uitgesproken. De inspecteur bezoekt de school enkel als onderzoeker.
- Schoolbezoeken nemen in principe één dagdeel in beslag. Bij het maken van de afspraken op een school met één groep 8 is ernaar gestreefd de maximumtijd van 4 uur te handhaven. Op scholen met parallelklassen diende het onderzoek zo mogelijk in alle groepen 8 plaats te vinden. Hier duurt het schoolbezoek langer dan een dagdeel.
- Het onderzoek moest worden uitgevoerd in de periode april – juli 2019, min of meer gelijktijdig met de hoofdafname. Het verdiepende onderzoek vond plaats ná de hoofdafname.

1.6.1 Het observatie-instrument

Het doel van de observaties was om na te gaan welke (leerkracht)activiteiten in de schrijflessen plaatsvonden, waardoor getracht is een dekkend overzicht van deze activiteiten in de schrijfles te geven. Hiertoe is gebruik gemaakt van het observatie-instrument dat is ontwikkeld en gebruikt bij de studie van Van Gelderen en Blok (1991). Er is gebruik gemaakt van de ***Predominant Activity Sampling (PAS)*** observatiemethodiek: gedurende een vooraf gespecificeerde periode werd geobserveerd en daarna werd de meest voorkomende (leerkracht)activiteit gedurende de observatieperiode gescoord. In het huidige onderzoek is elke **20 seconden** geobserveerd welk leerkrachthandelen in deze periode het meest voorkomt, gevolgd door een periode van 20 seconden om te scoren. Per observatieperiode is informatie genoteerd over ‘werkvorm en taakkenmerken’ en over de ‘fasering van de les’. Onder ‘werkvorm en taakkenmerken’ is informatie verzameld over *gerichtheid* (richt de leerkracht zich tot de hele groep, een subgroep, een individuele leerling?), *werkvorm* (samenwerken of individueel), *opdrachtkenmerk* (algemene kenmerken van de taak), *tekstsoort*, *publieksgerichtheid* en *materiaal*. De kern van het observatie-instrument richtte zich op de geobserveerde fasen in de schrijfles en de activiteiten die de leerkracht daarin uitvoert. De observator stelde per observatie vast in welke fase de les zich bevond en bepaalde welke activiteit dominant was gedurende de 20 seconden waarin is geobserveerd. Er werden vijf fasen onderscheiden: voorbereidende fase, de schrijffase, de revisiefase, de evaluatiefase en de publicatiefase. Iedere fase kende gedefinieerde categorieën van leerkrachtactiviteiten. Voor alle lesfasen gold dat ook kon worden genoteerd wanneer er sprake was van transitie (overgang van de ene naar de andere activiteit, zoals het uitdelen van materiaal), of van non-instructie (de les werd verstoord of de leerkracht was bezig met andere activiteiten dan de schrijfactiviteiten). Bij het scoren van de leerkrachtactiviteiten was de *intentie van de leerkracht* (wat wil de leerkracht met zijn/haar actie bereiken?) leidend. De observator is gevraagd *het feitelijke gedrag én de bedoeling van de leerkracht* te interpreteren (en te scoren). Observatoren hebben bij de observaties gebruik gemaakt van een digitale tool. In bijlage 6 staat het scorings scherm (met de betreffende scoringscategorieën) weergegeven. Een verdere toelichting op de (uitvoering van de) scoring wordt gegeven in paragraaf 6.5.2.

1.6.2 De interviewleidraden

Interviewleidraad leerkrachten

Voor het interview met de leerkrachten is een leidraad opgesteld, bestaand uit 22 vragen over vijf thema's (onderdeel A-E). Leerkrachten werden uitgenodigd te vertellen en de interviewer rubriceerde de antwoorden in gecodeerde categorieën op een in te vullen formulier. In onderdeel A is de visie op schrijfonderwijs bevraagd: wat verstaat de leerkracht onder goed schrijfonderwijs? De inspecteur kon daarbij aangeven waar de leerkracht *nadruk op legt* in de beschrijving van zijn/haar visie: de communicatieve functie van schrijven (belang van schrijven om te communiceren met anderen), de conceptualiserende functie van schrijven (belang van schrijven om te leren/om de werkelijkheid te ordenen), de expressieve functie van schrijven (belang van schrijven om persoonlijke gedachten/emoties te uiten) of het schrijven als middel om goed te leren spellen/correcte zinnen te leren bouwen. Ook registreerde de interviewer *centrale begrippen* die voorkwamen in de omschrijving van de visie. In onderdeel B (het leerplan) is leerkrachten gevraagd in hoeverre zij expliciet vastgelegde leerdoelen hanteren en evalueren. Met behulp van onderdeel C (de praktijk van schrijflessen) werd gevraagd naar het volgen van een bepaalde taakaanpak en of (en hoe) leerkrachten proberen zicht te krijgen op het schrijfproces. In onderdeel D (volgen en beoordelen) is gevraagd naar summatieve en formatieve beoordeling van schrijfproducten van leerlingen. Onderdeel E was gericht op ervaren problemen bij het geven van de schrijflessen en (additionele) scholing.

Interviewleidraden leerlingen

In iedere bezochte klas zijn twee leerlingen willekeurig door een onderzoeker geselecteerd, met wie individueel een kort interview is afgenomen. Deze interviews waren gericht op de taakaanpak van de leerlingen. Aan de hand van een voorbeeldschrijftaak is gevraagd hoe zij aan die taak zouden beginnen. Ook is gevraagd wat ze moeilijk vinden tijdens het schrijven en of en hoe zij teksten reviseren (aan de hand van een gefingeerde voorbeeldtekst). Daarnaast werd de wijze van instrueren door de leerkracht en de mate van samenwerking tijdens de schrijflessen bevraagd. De interviewer selecteerde zelf twee leerlingen, zonder inmenging van de leerkracht.

1.6.3 Training

Eind maart zijn de inspecteurs en twee orthopedagogiekstudenten getraind. De training richtte zich met name op het observatie-instrument (inclusief het werken met de digitale tool). De observatiecategorieën zijn uitgelegd en aan de hand van videomateriaal geoefend en besproken. Ook de interviewleidraden zijn toegelicht en de werkwijze tijdens de interviews is besproken. Een protocol met daarin een uitgebreide beschrijving van de inhoud van de instrumenten en de beoogde werkwijze is opgesteld en verstrekt aan de betrokken inspecteurs.

1.7 Verschillen ten opzichte van de vorige peiling

De huidige peiling onderscheidt zich op de volgende punten van de vorige peiling:

- Aan de peiling in 2009 namen leerlingen van groep 5 en 8 in het reguliere basisonderwijs (bo) en schoolverlaters in het speciaal basisonderwijs deel. Aan de huidige peiling namen enkel leerlingen van groep 8 (bo) en schoolverlaters (sbo) deel.
- Evenals in 2009 maakten leerlingen in groep 8 in de huidige peiling vier schrijftaken (afnameduur in totaal 90 minuten), terwijl schoolverlaters in het sbo twee schrijftaken maakten (afnameduur 45 minuten). Afnames waren niet helemaal vergelijkbaar: in 2009 vond afname in het sbo plaats onder begeleiding van een testleider, terwijl in groep 8 de helft van de taken onder begeleiding van een testleider en de helft door de eigen leerkracht werd afgenomen. In de huidige peiling werden alle taken in het regulier en het speciaal basisonderwijs door een getrainde testleider afgenomen, op één dag.
- In de vorige peiling is gebruik gemaakt van schrijftaken en vragenlijstdata om de leerlingprestaties en het onderwijsleerproces te beschrijven. In de huidige peiling is daarnaast een verdiepend onderzoek uitgevoerd, bestaande uit interviews met leerlingen en leerkrachten van groep 8 en lesobservaties van schrijflessen in groep 8. Informatie over het onderwijsleerproces is daarnaast in het bo bij leerkrachten van groep 6, 7 en 8 verzameld middels vragenlijsten (en bij leerkrachten van de schoolverlaters).
- In 2019 zijn niet dezelfde vragenlijsten gebruikt als in 2009. Wel was een deel van de vragen vergelijkbaar waardoor beschrijvende vergelijkingen daarvan mogelijk zijn. De drie ankertaken die gebruikt zijn ten behoeve van de trendvergelijking over tien jaar zijn wel ongewijzigd overgenomen (schrijfopdrachten en beoordelingsformulieren).

Hoofdstuk 2. Steekproef, werving en onderzoeksgroep

De peiling schrijfvaardigheid einde (speciaal) basisonderwijs heeft tot doel de schrijfvaardigheid van leerlingen van groep 8 in het reguliere basisonderwijs (bo) en de schoolverlaters in het speciaal basisonderwijs (sbo) in kaart te brengen. Hiervoor zijn voorafgaand aan de peiling twee steekproeven uit de twee populaties (bo en sbo) getrokken. De steekproeftrekking is door de Inspectie van het onderwijs uitgevoerd. In dit hoofdstuk wordt een verantwoording gegeven van de opzet en uitvoering van de twee steekproeftrekkingen en de daarbij spelende overwegingen (paragrafen 2.1 en 2.2). De opzet en het verloop van de werving van scholen is, inclusief de non-respons, beschreven in paragraaf 2.3. We besluiten het hoofdstuk met een overzicht en een representativiteitsanalyse van de gerealiseerde steekproef op school- en leerlingniveau (paragrafen 2.4 en 2.5).

2.1 Steekproeftrekking bo

Ten behoeve van de peiling heeft de Inspectie van het onderwijs voor het bo een steekproef getrokken van leerlingen die in het schooljaar 2018-2019 in groep 8 van het reguliere basisonderwijs zaten. De steekproef van leerlingen is indirect getrokken door een steekproef van bo-scholen te trekken, en (in principe) alle leerlingen in groep 8 op een geselecteerde school op te nemen in de steekproef. De steekproeftrekking diende aan een aantal, vooraf opgestelde vereisten te voldoen:

1. Er diende één hoofdsteekproef te worden getrokken - de eerste tranche (t1) -, alsmede twee reservetranches (t2, t3). Iedere tranche diende uit 120 scholen te bestaan;
2. De steekproef diende getrokken te worden binnen de groep schoolbesturen primair onderwijs die in schooljaar 2018/2019 geselecteerd waren voor een inspectiebezoek. Deze groep van schoolbesturen is representatief voor alle schoolbesturen primair onderwijs in Nederland.
3. Er diende rekening te worden gehouden met deelname aan het lopende onderzoek TIMSS/Peil. onderwijs rekenen-wiskunde 2018/19 dat in hetzelfde schooljaar plaatsvond als de peiling schrijfvaardigheid.¹²

2.1.1 Stratificatiecriteria

Daarnaast diende de steekproef te worden getrokken aan de hand van twee stratificatiecriteria: schoolgewicht en schoolgrootte. Schoolgewicht was een expliciet stratificatiecriterium. De stratificatie vond plaats op basis van het percentage leerlingen met een formatiegewicht ('gewichtenleerlingen') op een school. Dit criterium is voor alle scholen in de onderzoekspopulatie (het regulier basisonderwijs) bekend. Hierbij zijn vier gewichtenstrata¹³ gebruikt. De definitie van deze gewichtenstrata is gegeven in Tabel 2.1.

¹² Deze scholen zijn niet uit het steekproefkader verwijderd, maar verwisseld met een vergelijkbare school in 2^e of 3^e tranche als ze in de 1^e tranche terecht kwamen.

¹³ De afkorting S4o verwijst hier naar de gewichtenstratum vierdeling van de onderwijsinspectie

Tabel 2.1: Betekenis van de gewichtenstrata

Gewichtenstratum		Omschrijving	
Label	Afkorting	Volledig	Verkort
Stratum 1	S1	0% gewichtenleerlingen (geheel geen gewichtenleerlingen)	0% gew.
Stratum 2	S2	>0%-10% gewichtenleerlingen	>0-10%
Stratum 3	S3	>10%-25% gewichtenleerlingen	>10-25%
Stratum 4	S4	> 25% gewichtenleerlingen	>25%

In deze verdeling wordt er geen verschil gemaakt tussen leerlingen met een laag gewicht (0,30) en een hoog gewicht (1,20). Dit is een gebruikelijke keuze. Bij het trekken van de steekproef moest tevens rekening worden gehouden met een impliciet stratificatiecriterium, te weten schoolgrootte.

2.1.2. Exclusiecriteria

Exclusiecriteria op bestuurs-/schoolniveau

Op voorhand zijn de volgende exclusiecriteria op bestuurs-/schoolniveau geformuleerd.

- Eenpitterbesturen dienden te worden uitgesloten van de steekproef;¹⁴
- Een aantal specifieke scholen diende op voorhand van deelname te worden uitgesloten: nieuwkomerscholen, scholen in het buitenland, niet bekostigd onderwijs (nbo), rijdende scholen, azc-scholen, internationaal onderwijs, scholen die niet opgenomen waren in de kennisanalyse van de Inspectie van het Onderwijs (dat wil zeggen, scholen waarvan de opbrengsten in groep 8 ontbreken) en scholen met een voorziene opheffingsdatum.

Exclusiecriteria op leerlingniveau

Daarnaast zijn er in het peilingsonderzoek twee exclusiecriteria op leerlingniveau gehanteerd:

- Leerlingen die op grond van op de school geregistreerde kenmerken van cognitieve, emotionele, motorische of functionele aard niet konden deelnemen aan de peiling, werden uitgesloten van deelname. Het gaat daarbij om leerlingen die ook binnen de reguliere lessen (schrijven, taal, zaakvakken) in het Nederlands geen tekst kunnen produceren. Cruciaal was dus dan ook dat de genoemde kenmerken eveneens verhinderden dat de betreffende leerlingen op school normaliter schrijftaken uitvoerden;
- Leerlingen die, ten tijde van de afname, korter dan 1 jaar in Nederland waren *en* het Nederlands onvoldoende beheersten.

¹⁴ Om redenen van belasting van de onder deze eenpitters ressorterende scholen. Er is zowel nationaal als internationaal relatief weinig empirisch onderzoek gedaan naar de relatie tussen schoolbesturen en onderwijskwaliteit (zie rapport Kwaliteit van sturing in het primair onderwijs van Reezigt, Rekers-Mombarg en Spithoff, 2019). In het onderzoek 'De kwaliteit van schoolleiders in het basisonderwijs, speciaal onderwijs en voortgezet onderwijs' (Inspectie van het Onderwijs, 2014) vond de inspectie dat er geen significante samenhang is tussen de grootte van een bestuur (eenpitter of meerpitter) en de mate van kwaliteitsbewaking.

Leerlingen die voldeden aan één of beide criteria konden niet meedoen aan de afnames. In paragraaf 2.5 (uitval op leerlingniveau) wordt beschreven hoeveel leerlingen vanwege deze reden(en) niet hebben deelgenomen. Aan leerlingen met motorische beperkingen is de mogelijkheid geboden om de taken op de computer uit te voeren. Leerlingen met dyslexie konden gewoon met de afname meedoen.

2.1.3 Gebruikte bronnen

Voor de steekproeftrekking is een vijftal bronnen gebruikt¹⁵. Dit zijn bestanden die binnen de Inspectie van het Onderwijs worden gehanteerd. Op 23 maart 2018 waren 6612 scholen (clusters) voor regulier basisonderwijs (bo) actief, exclusief de scholen die voldeden aan de uitsluitingskenmerken.

2.1.4 Procedure

In totaal zijn er vier stappen ondernomen. De eerste stap was de uitsluiting van éénpitter besturen, waarna nog 6240 scholen resteerden. Hiervan stonden 1617 scholen vanwege het bestuursgerichte toezicht gepland voor schoolbezoek in 2018/19. De tweede stap was de bepaling van de benodigde aantallen in de steekproef, evenredig aan de populatie. De vooraf bepaalde stratificatiegroepen zijn hiertoe ingedeeld op basis van percentage gewichtenleerlingen op de school. In Tabel 2.2 staat per gespecificeerde categorie schoolgewicht het aantal scholen in de populatie en in de steekproef weergegeven, alsmede het aantal scholen dat in bestuursjaar 2018/19 bezocht zou gaan worden.

Tabel 2.2: Stratificatietabel basisonderwijs

		Populatie (N)	Populatie (%)	Steekproef (n)	Beoogd schoolbezoek Bestuursjaar 2018/19 (N)
Categorie school- gewicht	1- 0%	416	6.7	8	92
	2- max. 10%	3.824	61.3	73	932
	3- 10-25%	1.395	22.4	27	398
	4- >25%	604	9.7	12	195
	Onbekend	1			
	Totaal	6240	100.0	120	1617

Als derde stap zijn de scholen met gepland schoolbezoek in bestuursjaar 2018/19 op basis van schoolgrootte geordend binnen de stratificatiegroepen. Tenslotte, stap 4, zijn per stratificatiegroep vanaf een random beginpunt met gelijke intervallen scholen getrokken voor de eerste tranche van de steekproef (t1). Voor respectievelijk tranche 2 en 3 (t2 en t3) is de wat betreft schoolgrootte direct onder- en bovengelegen school getrokken. Per school in de eerste tranche waren qua schoolgewicht en schoolgrootte overeenkomstige scholen als reserve beschikbaar in de tweede en de derde tranche.

Steekproeven TIMSS en peiling rekenen-wiskunde

Naast de peiling schrijfvaardigheid werd er in het basisonderwijs ook een peilingsonderzoek uitgevoerd naar rekenen-wiskunde binnen de steekproef van het TIMSS-onderzoek (n=200 in eerste tranche) en een afzonderlijke aanvullende steekproef (n=40 in eerste tranche). In tabel 2.3 staat de overlap tussen

¹⁵ De volgende bronnen zijn gebruikt: Basisbestand 23112017; Kennisanalyse oordelen 17012018; (multi)sectorale bestuursverdeling 23032018; Elementenbestand 23032018; Opgave deelname TIMMS 10042018.

de tranches van de getrokken steekproef schrijfvaardigheid en de steekproef TIMSS/peiling rekenen-wiskunde weergeven.

Tabel 2.3: Overlap tranches met hoofsteekproef TIMMS en peil rekenen/wiskunde 2018/19.

		t1	t2	t3
TIMSS/Peil rekenen-wiskunde	Nee	117	119	113
	Ja	3	1	10

Om te voorkomen dat dezelfde scholen in de hoofsteekproeven (eerste tranche, t1) van beide peilingsonderzoeken terecht zouden komen, zijn drie scholen van de steekproef TIMSS/rekenen-wiskunde die in de eerste tranche zaten, geruild met de reserveschool uit de derde tranche. Hierdoor zaten er in de eerste tranche van dit onderzoek naar schrijfvaardigheid geen scholen meer die ook in de hoofsteekproef (1^e tranche) van TIMSS of de extra steekproef rekenen-wiskunde waren opgenomen.

2.2 Steekproeftrekking in het speciaal basisonderwijs (sbo)

Voor het speciaal basisonderwijs is een aparte steekproef getrokken van leerlingen die in het schooljaar 2018-2019 als schoolverlater in het sbo zaten. Ook deze steekproef is door de Inspectie van het onderwijs getrokken. De steekproef van leerlingen is indirect getrokken door een steekproef van sbo-scholen te trekken, en (in principe) alle schoolverlaters op een geselecteerde school op te nemen in de steekproef. Ook deze steekproeftrekking diende aan een aantal vooraf opgestelde vereisten te voldoen:

- Er dienden voor het peilingsonderzoek in schooljaar 2018/19 twee steekproeven te worden getrokken, één voor de peiling Rekenen en Wiskunde/TIMSS en één voor de peiling schrijfvaardigheid. Ieder van de twee steekproeven bestond uit twee tranches (hoofsteekproef t1 en reservesteekproef t2). Alle tranches dienden uit 60 scholen te bestaan (dus 4*60 scholen in totaal);
- Er diende rekening te worden gehouden met deelname aan eerder peilonderzoek mondelinge taalvaardigheid (mtv 2017/18): streven was om zo min mogelijk scholen die aan dit onderzoek hadden deelgenomen te selecteren voor de eerste tranche.

2.2.1 Stratificatiecriteria

De steekproef diende te worden getrokken aan de hand van twee stratificatiecriteria: regio als expliciet stratificatiecriterium en schoolgrootte (het aantal leerlingen per school) als impliciet stratificatiecriterium.

2.2.2. Exclusiecriteria

Exclusiecriteria op leerlingniveau

Voor het sbo golden dezelfde twee exclusiecriteria op leerlingniveau als in het bo (zie p. X). Leerlingen die aan de uitsluitingskenmerken voldeden konden niet deelnemen aan de afnames. Voor leerlingen

met motorische beperkingen was er de mogelijkheid om de taken op de computer uit te voeren. Leerlingen met dyslexie konden ook in het sbo gewoon met de afname meedoen. Voor leerlingen met een zeer laag technisch leesniveau (lager dan E3) waren speciale voorzieningen (zie paragraaf 3.1).

2.2.3 Gebruikte bronnen

Voor de steekproeftrekking is een drietal bronnen gebruikt¹⁶. Dit zijn bestanden die binnen de Inspectie van het Onderwijs worden gehanteerd. Op 7 maart 2018 zijn 280 scholen voor sbo actief.

2.2.4 Procedure

Deze steekproeftrekking is ook weer uitgevoerd in afzonderlijke stappen. Allereerst zijn de stratificatiegroepen ingedeeld op basis van regio van de school (Noord, Oost, Midden, Zuid). Ten tweede zijn benodigde aantallen in de steekproef bepaald, evenredig aan de populatie.

Tabel 2.4: Stratificatietabel sbo-scholen

		Populatie (N)	Populatie (%)	Steekproef (n)
Regio op basis van provincie	Noord (Drenthe, Friesland, Groningen)	28	10.0	6
	Oost (Gelderland, Overijssel)	55	19.6	12
	Midden (Flevoland, Noord-, Zuid- Holland, Utrecht)	138	49.3	29
	Zuid (Limburg, Noord-Brabant, Zeeland)	59	21.1	13
Totaal		280	100	60

In totaal waren er in schooljaar 2018-2019 280 sbo-scholen actief. Van deze 280 sbo-scholen moesten er 240 in één van de vier tranches terechtkomen voor de werving van de peilingsonderzoeken Rekenen en wiskunde/TIMSS en schrijfvaardigheid 2018/19. In het schooljaar 2017/18 namen er van de 280 actieve sbo-scholen 65 scholen deel aan het peilingsonderzoek mondelinge taalvaardigheid (mtv).

Besloten is om 40 van deze 65 sbo-scholen niet in de tranches voor beide peilingen op te nemen en de overige 25 sbo-scholen zoveel mogelijk in de tweede tranches op te nemen. Daartoe zijn de zogenaamde 'mtv-sbo-scholen' per regio random buiten de steekproef gehouden. Een voorbeeld: voor de regio Noord waren 24 scholen nodig. Immers, er dienden in totaal 4 tranches voor beide peilingen te worden getrokken, bestaande uit ieder zes sbo-scholen. Dit betekende dat er vier van de 28 sbo-scholen in Noord buiten de tranches konden blijven. Van de 8 scholen die in 2017-18 aan de peiling mtv hadden deelgenomen, konden daardoor willekeurig vier scholen uit deze regio (Noord) worden geselecteerd, die buiten de tranches zijn gehouden. Tabel 2.5 laat per regio zien hoeveel scholen aan mtv hebben deelgenomen.

¹⁶ De volgende bronnen zijn gebruikt: Basisbestand 07032018; Elementenbestand 07032018; Opgave deelname peiling mtv 1802018.

Tabel 2.5: Deelname aan peiling mtv (sbo) 2017/18, per regio

	Regio	Mondelinge taalvaardigheid
Regio op basis van provincie	Noord	8
	Oost	15
	Midden	29
	Zuid	13
Totaal		65

De volgende stap was om de scholen per regio te ordenen naar schoolgrootte. Elke vier opeenvolgende scholen kregen uitsluitend een van de volgende vier getallen toegewezen.

1. Schrijven tranche 1
2. Schrijven tranche 2
3. Rekenen tranche 1
4. Rekenen tranche 2

De volgende acties zijn uitgevoerd om de kans op een mtv-sbo-school in de eerste tranches zo klein mogelijk te maken:

- Wanneer er bij de vier opeenvolgende scholen geen enkele mtv-school zat zijn de nummers random toegewezen;
- Wanneer er bij de vier opeenvolgende scholen één mtv-school zat, is deze random aan een van beide tweede tranches toegewezen en de overige (niet mtv-)scholen zijn vervolgens random aan de overige drie tranches toegewezen;
- Wanneer er bij de vier opeenvolgende scholen twee mtv-scholen bij zaten, zijn deze random aan de tweede tranches toegewezen, en de overige twee scholen random aan de eerste tranches.

Meer dan twee mtv-scholen kwamen niet voor per set van vier opeenvolgende sbo-scholen. Deze procedure resulteerde in evenredige steekproeven wat betreft regio en schoolgrootte. Geen enkele school die aan de mtv-peiling heeft deelgenomen zit in de eerste tranche. Per school in de eerste tranche is een qua regio en schoolgrootte overeenkomstige school als reserve beschikbaar in de tweede tranche.

2.3 Opzet en uitvoering van de werving

Een goed steekproefkader en een correcte steekproeftrekking zijn van belang voor een goed beeld van de schrijfvaardigheid van leerlingen in het (s)bo. Daarnaast is het natuurlijk noodzakelijk dat er voldoende scholen aan de peiling willen deelnemen. Om aan deze voorwaarde te voldoen is vanaf eind april 2018 een gerichte wervingscampagne uitgezet. Hieronder wordt een toelichting gegeven op de aankondiging van de peiling bij scholen, de inzet van ambassadeurs en het verloop van de wervingsperiode.

2.3.1 Informeren beoogde scholen

Aan de schoolleiders van bo- en sbo-scholen uit de hoofdstreekproef - de eerste tranche-scholen - is op 23 april 2018 een brief verstuurd met daarin informatie over doel en belang van de peiling schrijfvaardigheid, de verwachte inspanning voor de school en de te verwachten opbrengsten, alsmede de beoogde afnameperiode en organisatorische informatie. Tevens is een bijsluiter toegevoegd, waarin bondig en overzichtelijk is weergegeven wat deelname de school oplevert, van de school vraagt, hoe de schrijfvaardigheid wordt bepaald en hoe met de verstrekte gegevens wordt omgegaan. In de brief is tevens aangekondigd dat de scholen persoonlijk benaderd zouden worden door zogenaamde ambassadeurs. Een e-mail met dezelfde inhoud is op 24 april 2018 aan de schooldirecties verstuurd. Indien een school uit de eerste tranche weigerde, werd door de onderzoekers de reserveschool per post en per e-mail benaderd (uit t2) en nadien door de ambassadeur benaderd. Deze procedure werd ook gevolgd met reservescholen uit de derde tranche, in het geval van weigering door scholen in de tweede tranche.

Daarnaast is op 23 april een brief vanuit de Inspectie van het Onderwijs (ondertekend door de plaatsvervangend Inspecteur-Generaal) verstuurd aan de besturen van de betreffende scholen in de eerste tranche. Naast een inhoudelijke toelichting van de peiling en een aankondiging dat we de directeurs van de betreffende scholen zouden benaderen, had de brief vooral tot doel dat het bestuur de scholen zou motiveren aan de peiling deel te nemen. Voorafgaand aan de informatieronde is een website (www.peil-schrijfvaardigheid.nl) ingericht met alle informatie over het onderzoek en eventuele deelname. In de e-mails en brieven werd naar deze website verwezen.

2.3.2 Inzet ambassadeurs

Om scholen te motiveren voor deelname aan onderzoek, lijkt niets beter te werken dan een persoonlijke benadering. De 120 scholen uit de eerste steekproeftranche zijn daarom vanaf half mei benaderd en geworven door een team van 14 ambassadeurs. Er is met name gekozen voor de inzet van taalspecialisten en/of (ex-) schoolleiders met affiniteit voor schrijfonderwijs als ambassadeur. Voor de werving van deze ambassadeurs is gezocht binnen het eigen netwerk van het consortium. Daarnaast hebben we gebruik kunnen maken van het netwerk van Herman Franssen (lid van de begeleidingscommissie van de peiling schrijfvaardigheid, Inspecteur Primair Onderwijs, Lid Platform Onderwijs Nederlands en bestuurslid van de Stichting Lezen). In het contact met de scholen fungeerden de ambassadeurs als een soort 'visitekaartje' voor het onderzoek. In eerste instantie verliep de benadering via telefonisch contact en e-mail. Indien de school dat wenste, kon de ambassadeur een persoonlijk bezoek aan de school brengen om het onderzoek nader toe te lichten. Ter ondersteuning van de werving is een 'handleiding ambassadeurs' opgesteld en gedeeld met de ambassadeurs. Indien nodig, bijvoorbeeld bij specifieke vragen of afwijkende wensen voor afname van een schoolleider, werden zij begeleid door de onderzoekers.

De ambassadeurs ontvingen verder per toegewezen school een Excel formulier waarin de gegevens van de beoogde eerste tranche scholen waren opgenomen. De toegewezen scholen waren zoveel mogelijk verdeeld o.b.v. de voorkeursregio van de ambassadeurs. Bij weigering door een school ontving de ambassadeur de benodigde informatie voor de reserveschool. Ingeval van weigering werden ambassadeurs verzocht de reden(en) tot weigering te mailen. Deze zijn geregistreerd door de onderzoekers. De meeste correspondentie vond op verzoek van de scholen plaats via telefoon en e-mail.

2.3.3 Verloop van de werving

De telefonische werving door de ambassadeurs startte half mei 2018 en liep door tot eind oktober 2018. Op 11 oktober 2018 is in overleg met de begeleidingscommissie besloten dat het toenmalige toegezegde aantal basisscholen (93) en sbo-scholen (40) voldoende was om een representatieve peiling uit te kunnen voeren. In Tabel 2.6 staat per schooltype (bo en sbo) en tranche weergegeven hoeveel scholen eind oktober 2018 waren geworven. Uitgaande van een ondergrens van 102 deelnemende basisscholen (85% van de steekproef van 120 scholen) moet gesteld worden dat dit aantal niet gehaald is: 78% van de steekproef was geworven. Wel was het percentage scholen dat afkomstig was uit de hoofdstekproef hoger dan beoogd: met 67 scholen uit de eerste tranche (56%) werd voldaan aan de minimumeis van 50% deelname uit de eerste tranche. Voor het sbo gold ook dat de ondergrens, namelijk 51 scholen (=85% van 60 scholen) niet was behaald. Met 40 scholen lag het responspercentage op 66% in het sbo. Met 26 sbo-scholen uit de eerste tranche (43%) werd in het sbo niet voldaan aan de minimumeis van 50%. De algehele bereidwilligheid tot deelname lag bij de benaderde bo-scholen op 48.5% (=95/196), terwijl die bij de sbo-scholen op 43% (=40/93) lag.

Tabel 2.6: Bruto wervingsresultaat eind oktober 2018

		Benaderd	Toezegging	Weigering
Basisonderwijs (bo)	1 ^e tranche	120	65	55
	2 ^e tranche	51	23	28
	3 ^e tranche	25	5	20
Totaal		196	93	103
Speciaal basisonderwijs (sbo)	1 ^e tranche	60	26	34
	2 ^e tranche	33	14	19
	Totaal		93	40

2.3.4 Aanpassing planning: gevolgen voor de werving en voor de gerealiseerde steekproef

Door wisselingen in het onderzoeksteam, waardoor taken moesten worden overgedragen, is de voortgang van het peilingsonderzoek enige maanden vertraagd. De deelnemende scholen zijn hierover via e-mail geïnformeerd in december 2018. De vertraging heeft tot gevolg gehad dat de planning van het project moest worden verschoven van de beoogde – en aan scholen gecommuniceerde - periode januari t/m maart 2019 naar de periode april t/m juni 2019. Deze wijziging is eind januari 2019 definitief aan de scholen gecommuniceerd middels een e-mail, waarin zij geïnformeerd zijn over de gewijzigde afnameperiode en een herzien stappenplan ontvingen. Ook ontvingen de scholen een herziene brief die zij aan de ouders konden sturen in het kader van de ouderlijk toestemming (voor het geval de scholen de eerder verstuurde brief nog niet hadden verspreid).

De gewijzigde planning had gevolgen voor het wervingsresultaat. Van de (uiteindelijk) 96 bo-scholen die begin december 2018 waren geworven, bleken vijf scholen uit de eerste tranche en één school uit de tweede tranche definitief af te zeggen, omdat zij ongelukkig waren met de gewijzigde afnameperiode die zich kenmerkt door de eindtoetsen (april) en afsluitende evenementen in groep 8. Afgesproken werd om te proberen deze zes bo-scholen door reservescholen te vervangen. Een nieuwe wervingsronde bij reservescholen is uitgevoerd door de onderzoekers. Aan het eind van deze wervingsronde kwamen we uit op een definitief aantal bo-scholen van 95. Het uiteindelijk percentage

scholen uit de eerste tranche is door de wijzigingen verlaagd naar 51%, waarmee nog steeds voldaan wordt aan de minimumeis. Van de 40 geworven sbo-scholen zegde één school af vanwege de gewijzigde afnameperiode. Tabel 2.7 laat het bruto wervingsresultaat zien, met daarin per tranche en schooltype ook het definitieve aantal aangemelde scholen. In totaal hadden in het voorjaar van 2019 134 scholen deelname aan de peiling toegezegd. Het uiteindelijke deelnemerspercentage in bo bedraagt daarmee 44% (95/215) terwijl dat in het sbo 41% is (39/94).

Tabel 2.7: Bruto wervingsresultaat voorjaar 2019 (gerealiseerde steekproef)

		Benaderd	Toezegging	Weigering
Basisscholen	1 ^e tranche	120	61	59
	2 ^e tranche	60	26	34
	3 ^e tranche	35	8	27
Totaal		215	95	120
Sbo-scholen	1 ^e tranche	60	24	36
	2 ^e tranche	34	15	19
	Totaal	94	39	55

Genoemde redenen tot weigeringen

Zoals in Tabel 2.9 is weergegeven hebben in totaal 175 scholen deelname geweigerd. Bij weigering of afmelding is deze scholen gevraagd waarom zij hiertoe besloten. De ambassadeurs hielden de genoemde redenen tot weigering bij. Deze gegevens zijn achteraf geanalyseerd om de non-respons in kaart te brengen. Er zijn verschillende redenen naar voren gebracht. Het is lastig deze terug te brengen tot categorieën: soms spelen er hele specifieke zaken (fusie, ander gebouw) en soms worden meerdere redenen naar voren gebracht. Vaak zijn dat meerdere redenen om aan te geven dat de school 'te druk' is met allerlei activiteiten en ontwikkelingen. We hebben dat als een reden gescoord, ook omdat de genoemde redenen vaak terug te brengen zijn tot een reden: een overladen programma, veel personele wisselingen, etc. We hebben per school de belangrijkste reden gekozen en gepoogd voldoende onderscheid te maken tussen soorten redenen. 'Te druk' door fusie, overladen programma, werkdruk etc. hebben we gescoord als 'overladen programma'.

Druk kan betekenen: drukke periode op school tijdens de afnameperiode, soms vanwege genoemde verschuiving van de afnameperiode (code 1). Te druk kan ook betekenen: te veel activiteiten of ontwikkelingen (1 keer een fusie), te grote belasting/werkdruk. We hebben dat gescoord met code 2. Soms zit de school in een periode waarin onderzoek geen prioriteit heeft, omdat andere, soms dringender onderwerpen aan de orde zijn (code 8) of omdat er al andere onderzoeken lopen of liepen (code 3). Het onderscheid tussen code 2 en code 8 is vooral dat code 8 aangeeft dat de school duidelijk andere prioriteiten (genoemd) heeft of met specifieke ontwikkelingen bezig is (niet zijnde verbetertrajecten vanwege een zwakke beoordeling), terwijl code 2 gereserveerd is voor scholen die 'een overladen programma', niet nader gespecificeerd, aangeven als reden. Soms is als reden genoemd dat schrijven überhaupt geen prioriteit heeft, omdat er niet aan schrijfonderwijs gedaan wordt (code 5). Er worden ook personele zaken genoemd als reden: veel personele wisselingen of een personeelstekort hebben we onderscheiden van een langdurig zieke leerkracht of een vervanger in groep 8 (resp. code 9 en code 4). Wanneer de school in een verbetertraject zit, soms met nieuwe directeur, vaak vanwege een zwakke inspectiebeoordeling, hebben we dit apart gescoord met code 6. Het kwam ook voor dat een directeur wel wilde meedoen aan het onderzoek maar zijn leerkrachten

niet meekreeg. We hebben dat met code 7 gescoord. In een enkel geval wilde een bestuur niet dat de school mee zou doen (code 10), in een enkel geval wilden de ouders niet dat hun kinderen deelnamen (vanwege een klein aantal leerlingen in een klas, code 14). Ook is een keer aangegeven dat het leerlingenaantal in groep 8 te klein werd gevonden voor deelname (code 11) en dat de school opgeheven was (code 13). Een aantal scholen wilde geen reden noemen (code 0). Specifiek door sbo-scholen zijn de onderwerpen behorend bij code 12 genoemd. Hieronder vallen: leerlingen die schoolverlater zijn, zijn a) te kort op school voor een evaluatie, hebben b) te lage IQ's en/of kunnen c) nauwelijks schrijven of er lopen d) veel onderzoeken bij de schoolverlater waardoor dit onderzoek er niet meer bij kan. De aantallen per genoemde redenen zijn opgenomen in Tabel 2.8.

Van alle genoemde redenen vallen codes 5 en 6 onder selectieve non-respons; iets wat om representativiteitsredenen niet gewenst is in peilingsonderzoeken. Gelukkig gaat het om geringe percentages (1.6% & 7.5% in het bo, 1.8% & 5.5% in sbo), met naar alle waarschijnlijkheid geringe impact op de resultaten. Ook non-respons vanwege het feit dat schoolverlaters 'nauwelijks schrijven' (sbo) kan onder selectieve non-respons worden geschaard.¹⁷

¹⁷ Het is onduidelijk bij hoeveel sbo-scholen (van 6) dit speelde.

Tabel 2.8: Redenen voor weigeren deelname bij bo- en sbo-scholen

	bo (n)	sbo (n)	Totaal (n)
Onbekend (code 0)	6	0	6
Afnameperiode is te drukke periode op school (code 1)	13	2	15
Overladen programma (te veel activiteiten, ontwikkelingen, fusie, te grote belasting/tijdnood/te druk/veel werkdruk) (code 2)	30	13	43
Teveel onderzoeken (gehad of nog lopend) (code 3)	8	4	12
Langdurig zieke leerkracht/vervanger gr 8 (code 4)	1	0	1
Schrijven is geen prioriteit op school (code 5)	2	1	3
Zwakke school/school in verbetertraject, soms geen of nieuwe directeur (code 6)	9	3	12
Leerkrachten willen niet: te druk, volgen een opleiding, zien geen meerwaarde (code 7)	8	1	9
Dringender onderwerpen, andere behoeften, schoolontwikkeling, past niet in jaarplan (code 8)	19	10	29
Veel personele problemen of wisselingen/veel ziekte, geen of net nieuwe directeur of directeur gaat weg (code 9)	21	14	35
Mag niet van bestuur (code 10)	1	0	1
Te klein aantal leerlingen groep 8 (code 11)	1	0	1
Sbo: veel onderzoeken bij schoolverlaters, te kort op school voor een evaluatie, meer nadruk op mondelinge taalvaardigheid: schriftelijke taalvaardigheid is te moeilijk, (te) lage IQ's (code 12)	0	6	6
School opgeheven (code 13)	0	1	1
Ouders weigeren (code 14)	1	0	1
Totaal	120	55	175

Voor het regulier basisonderwijs is nagegaan hoe de aanmeldingen en weigeringen verdeeld zijn over de vier strata. Uit Tabel 2.9 kan worden afgeleid dat het percentage weigeringen zowel in S1 als S4 hoger dan 60% is. Het lijkt daarom niet waarschijnlijk dat het percentage leerlingen met een formatiegewicht op school leidend is geweest bij de beslissing tot deelname of weigering, temeer omdat het percentage weigeringen in S3 (10-25% gewichtenleerlingen) niet heel sterk afwijkt van de overige categorieën.

Tabel 2.9: Non-respons naar stratum

Cel	Totaal benaderd	Geen aanmelding	Aangemeld	Percentage Non-respons
S1.0%	16	10	6	63%
S2. 0-10%	127	66	61	52%
S3. 10-25%	47	27	20	57%
S4. >25%	25	17	8	68%
Totaal	215	120	95	56%

De non-respons in het basisonderwijs heeft met deze verdeling van 95 scholen geen consequenties voor de generaliseerbaarheid van de resultaten van dit peilingsonderzoek naar de expliciete stratificatiecriteria, aangezien de reservescholen altijd behoorden tot hetzelfde stratum en type school als de beoogde 'eerste' school.

Netto Respons

Tussen april en juni 2019 hebben de afnames plaatsgevonden. Tabellen 2.10 en 2.11 bevatten de aantallen scholen voor het reguliere basisonderwijs en het sbo dat uiteindelijk heeft deelgenomen. Alle 95 aangemelde basisscholen hebben uiteindelijk aan de peiling deelgenomen. Er is in de afnameperiode geen school meer uitgevallen.

Tabel 2.10: Aanmelding en deelname naar stratum - bo

%		Steek Proef	Beoogd percentage	Gerealiseerde deelname	Percentage realisatie
Gewichtenleerlingen	0%	8	6.7 %	6	6.3%
	0-10%	73	60.8 %	61	64.2%
	10-25%	27	22.5 %	20	21.0%
	>25%	12	10.0 %	8	8.4%
Totaal		120	100 %	95	100%

In de afnameperiode bleek dat één sbo-school het niet voor elkaar kreeg de afname daadwerkelijk te realiseren. Het netto gerealiseerde aantal sbo-scholen betreft derhalve 38 (Tabel 2.11).

Tabel 2.11: Daadwerkelijke deelname naar stratum - sbo

		Steekproef	Beoogd percentage	Gerealiseerde deelname	Percentagerealisatie
Regio	Noord	6	10%	5	13.2%
	Oost	12	20%	9	23.7%
	Midden	29	48%	15	39.5%
	Zuid	13	22%	9	23.7%
Totaal		60	100%	38	100%
			Aantal	Gemiddeld aantal	SD
Schoolgrootte	1 ^e		23	133	64.1
	2 ^e		15	120	50.7
Totaal			38	128	59.0

2.3.5 Representativiteit op schoolniveau

Om de representativiteit van de gerealiseerde steekproef na te gaan is een aantal achtergrondkenmerken van de totale set deelnemende scholen (bo en sbo) vergeleken met die van de landelijke populaties (speciale) basisscholen.

Basisonderwijs

Tabel 2.12 bevat de verdeling naar regio, mate van stedelijkheid, schoolgrootte, aantal groep 8 leerlingen, percentage gewichtenleerlingen op schoolniveau, en denominatie binnen de steekproef vergeleken met de totale populatie basisscholen. De eenheid 'school' is in deze peiling gedefinieerd door de combinatie van BRIN-code en vestigingsnummer. Ook nevenvestigingen zijn dus, net als bij de steekproef, betrokken in de vergelijking. In tegenstelling tot de steekproef (schooljaar 2017-2018) kon bij deze vergelijking wel gebruik gemaakt worden van de landelijke DUO-schoolgegevens van het schooljaar van de afname (2018-2019). Er is gebruikt gemaakt van de DUO-bestanden '01.leerlingen-po-soort-po-cluster-leeftijd-2018-2019.csv', en '03.leerlingen-bo-sbo-leerjaar-geslacht-2018-2019.csv' die verkrijgbaar zijn via de website van DUO. Beide bestanden kennen 6333 schoolvestigingen basisonderwijs (brinnummers). Het DUO-bestand bevat geen directe gegevens over de mate van stedelijkheid. Hiervoor is gebruik gemaakt van het CBS-bestand '161010-Kenmerken-postcode-mw.xlsx'. Daarin is de mate van stedelijkheid gekoppeld aan alle Nederlandse postcodes. Omdat sinds eind 2018 geen BRON gegevens van leerlingen meer worden gepubliceerd is voor de variabele 'percentage gewichtenleerlingen' gebruik gemaakt van het DUO-bestand van het voorgaande schooljaar ('02.leerlingen-bo-in-bron-2017-2018.csv').

De scholen zijn ingedeeld naar percentage gewichtenleerlingen op schoolniveau met dezelfde gewichtenstrata zoals gehanteerd als bij het vaststellen van de steekproef (zie paragraaf 2.1). Omdat we in de steekproef met opzet rekening hielden met een strikte indeling van het percentage gewichtenleerlingen in vier strata (0%, >0-10%, >10-25%, >25%), blijkt de verdeling van de groepen in de steekproef (95 scholen) bijna volledig overeen te komen met de populatie. Er is geen sprake van significante verschillen in de verdeling van de scholen over de vier strata ten opzichte van de populatie: $\chi^2(3; N = 6302) = 0,34, p = 0,952$.

Voor het criterium 'stedelijkheid' geldt dat er geen statistisch significante verschillen bestaan wanneer de steekproef wordt vergeleken met de totale populatie basisscholen $\chi^2(4; N = 6332) = 2,31, p = 0,679$. Ook voor het impliciete criterium 'schoolgrootte' geldt dat er geen significante verschillen tussen de steekproef (95 scholen) en de populatie zijn: $\chi^2(4; N = 6319) = 0,60, p = 0,963$.

Ten opzichte van de gehele populatie zien we ten aanzien van het aantal leerlingen in leerjaar 8 alleen een afwijking in de onderste groep (<6 leerlingen), één school in de steekproef, ten opzichte van 4.0% in de populatie). Bij de overige groepen, die qua aantal leerlingen globaal kunnen worden omschreven als 'maximaal halve klas (6 tot 15)', 'normale klas' (16-30), 'twee klassen (31-60)' en 'meer dan twee klassen', zien we dat er geen significante verschillen zijn ($\chi^2(4; N = 6333) = 2,37, p=0,669$).

Wanneer gekeken wordt naar de geografische ligging van de scholen, dan worden vier regio's onderscheiden: Noord (provincies Groningen, Drenthe en Friesland), Oost (provincies Gelderland en Overijssel), Midden (provincies Flevoland, Noord-Holland, Zuid-Holland en Utrecht) en Zuid (provincies Noord-Brabant, Limburg en Zeeland). Er lijkt een oververtegenwoordiging van de steekproefscholen uit de oostelijke provincies aan het onderzoek te hebben meegedaan. Voor regio Noord lijkt er juist

sprake te zijn van een ondervertegenwoordiging. De gevonden verschillen tussen steekproefscholen en populatie zijn echter niet significant met een drempel van $p < 0,05$: $\chi^2(3; N = 6333) = 5,05, p = 0,168$.

Het DUO-bestand bevat zeventien categorieën als het gaat om denominatie. Voor de overzichtelijkheid zijn deze groepen gereduceerd tot vier hoofdgroepen. De categorieën 'algemeen bijzonder', 'openbaar' en 'rooms-katholiek' zijn integraal overgenomen uit het DUO-bestand. Algemeen bijzonder omvat scholen die niet op basis van levensbeschouwing zijn in te delen, maar bijvoorbeeld behoren tot het traditioneel vernieuwingsonderwijs (Jenaplan, Montessori et cetera). Aan de groep protestants-christelijke scholen zijn naast de scholen die zich zo afficheren ook de Evangelische, de Gereformeerd vrijgemaakte en de Reformatorische scholen toegevoegd. In de categorie 'overig levensbeschouwelijk' vallen alle scholen van religieuze minderheden (islamitische, joodse, hindoeïstische scholen), confessioneel samenwerkende scholen (bijvoorbeeld rooms-katholiek en protestants-christelijk), 'interconfessionele' scholen en antroposofische scholen. De vergelijking laat zien dat er geen sprake is van een statistisch significant verschil tussen de gerealiseerde steekproef en populatie: $\chi^2(4, N = 6333) = 5,72, p = 0,221$.

Ten aanzien van de achtergrondvariabelen kan geconcludeerd worden dat de basisscholen uit de steekproef (95 scholen) voldoende representatief zijn voor de populatie. Bij gebruik van de data (op schoolniveau) hoeft er dus niet voor deze variabelen gecorrigeerd te worden.

Tabel 2.12: Deelname scholen naar achtergrondkenmerken t.o.v. landelijke populatie

	Steekproef (N=95)		Landelijke populatie	
	n	%	n	%
Regio (N=6333)				
Noord	31	32.6	920	42.0
Oost	20	21.1	1381	14.5
Midden	22	23.2	2663	21.8
Zuid	22	23.2	1369	21.6
Stedelijkheid (N=6332)				
Zeer sterk stedelijk	14	14.7	875	13.8
Sterk stedelijk	17	17.9	1349	21.3
Matig stedelijk	16	16.8	1172	18.5
Weinig stedelijk	17	17.9	1260	19.9
Niet stedelijk	31	32.6	1676	26.5
Schoolgrootte (N=6319)				
1-100 lln.	17	17.9	1136	18.0
101-200 lln.	33	34.7	2104	33.3
201-300 lln.	24	25.3	1640	26.0
301-400 lln.	10	10.5	797	12.6
401 en meer lln.	11	11.6	642	10.2

Tabel 2.12: Deelname scholen naar achtergrondkenmerken t.o.v. landelijke populatie (vervolg)

	Steekproef (N=95)		Landelijke populatie	
	n	%	n	%
Aantal leerlingen in leerjaar 8 (N=6333)				
<i>Minder dan 6</i>	1	1.1	255	4.0
<i>6 tot 15</i>	23	24.2	1426	22.5
<i>16 tot 30</i>	39	41.1	2621	41.4
<i>31 tot 60</i>	27	28.4	1739	27.5
<i>Meer dan 60</i>	5	5.3	292	4.6
Percentage gewichtenleerlingen (N=6302)				
0% leerlingen	36	37.9	2494	39.6
> 0%-10% leerlingen	31	32.6	1889	30.0
>10%-25% leerlingen	20	21.1	1391	22.1
> 25% leerlingen	8	8.4	528	8.4
Denominatie (N=6333)				
Algemeen bijzonder	2	2.1	306	4.8
Openbaar	37	38.9	1878	29.7
Rooms-katholiek	29	30.5	1907	30.1
Protestants-christelijk	25	26.3	2033	32.1
Overig levensbeschouwelijk	2	2.1	209	3.3

Speciaal basisonderwijs

Tabel 2.14 bevat de verdeling van de sbo-scholen naar achtergrondvariabelen. De variabelen regio, stedelijkheid en denominatie zijn op dezelfde wijze samengesteld als de bo-scholen (zie hierboven). Daarbij is gebruik gemaakt van het DUO-bestand '02.-hoofdvestigingen-speciaal-(basis)onderwijs.xlsx'. Voor de variabele stedelijkheid is hetzelfde bestand van CBS gebruikt als bij de reguliere bo-scholen (zie boven). De variabele 'percentage gewichtenleerlingen' is niet opgenomen aangezien leerlinggewichten in het sbo geen rol spelen.

Uit Tabel 2.13 valt op te maken dat er in deze peiling sprake lijkt van ondervertegenwoordiging van sbo-scholen uit de regio Midden, terwijl er een oververtegenwoordiging van de overige regio's lijkt te zijn. De verschillen zijn echter niet significant $\chi^2(3; N = 280) = 1,55, p=0,670$. Het aantal scholen uit de regio Midden lijkt ervoor te zorgen dat er in de steekproef voor het achtergrondkenmerk stedelijkheid sprake is van een oververtegenwoordiging van scholen in 'sterk stedelijke' gebieden, terwijl er minder scholen dan zou mogen worden verwacht uit de groep 'zeer sterk stedelijk' hebben deelgenomen. Statistisch gezien is dit echter (net) niet significant $\chi^2(3; N = 267) = 8,71, p=0,069$. Ten aanzien van het aantal schoolverlaters (leerjaar 8) per school is de steekproef zeer vergelijkbaar met de populatie en dus statistisch gezien niet significant afwijkend $\chi^2(3; N =) = 0,172, p=0,997$. Ten slotte is ook de verdeling naar denominatie niet statistisch significant afwijkend van de populatie, $\chi^2(3; N = 280) = 9,77, p=0,202$. Ten aanzien van alle achtergrondvariabelen kan geconcludeerd worden dat de groep sbo-scholen uit de steekproef (38 scholen) voldoende representatief is voor de populatie sbo-scholen.

Tabel 2.13: *Deelname sbo-scholen naar achtergrondkenmerken ten opzichte van landelijke populatie*

	Deelnemende scholen		Landelijke populatie	
	n	%	n	%
Regio (N=280)				
noord	5	13.2	29	10.4
oost	9	23.7	55	19.6
midden	15	39.5	136	48.6
zuid	9	23.7	60	21.4
Stedelijkheid (N =267)				
zeer	1	2.6	49	18.4
sterk	20	52.6	109	40.8
matig	12	31.6	72	27.0
weinig	5	13.2	32	12.0
niet	0	0	5	1.9
Aantal lln in leerjaar 8 (n=280)				
<i>Minder dan 6</i>	1	2.6	8	2.9
<i>6 tot 15</i>	6	15.8	44	15.7
<i>16 tot 30</i>	22	57.9	157	56.1
<i>31 tot 60</i>	8	21.1	65	23.2
<i>Meer dan 60</i>	1	2.6	6	2.1
Denominatie (N=280)				
Algemeen bijzonder	11	28.9	74	26.4
Openbaar	8	21.1	81	28.9
Rooms-katholiek	10	25.3	68	24.3
Protestants-christelijk	1	2.6	3	1.1
Overig levensbeschouwelijk	8	21.1	54	19.3

2.4 Gerealiseerde steekproef: leerlingen

In de opzet van de peiling was het de bedoeling dat per deelnemende school alle leerlingen in leerjaar 8 en alle schoolverlaters aan het onderzoek zouden deelnemen. Dit had tot gevolg dat er 23 basisscholen in de gerealiseerde steekproef zitten met twee of meer groepen 8, en 27 scholen in het sbo met twee of meerdere (kleinere) groepen schoolverlaters. In de voorbereiding van de afname hebben de scholen de gegevens van 3560 leerlingen aangeleverd. In Tabel 2.14 zijn de aantallen *aangemelde* leerlingen uitgesplitst naar stratum. Aan elke aangemelde leerling is een uniek volgnummer toegewezen. Voor elke klas of deel van de klas (bij combinatieklassen) zijn telkens twee reservevragenlijsten en twee extra boekjes met schrijftaken aangemaakt.

Tabel 2.14: Aanmeldingen leerlingen naar stratum

Cel	Scholen	Aantal leerjaar 8	Percentage
S4o1 (0%)	6	157	6.2%
S4o2 (>0-10%)	61	1782	70.2%
S4o3 (>10-25%)	20	372	14.7%
S4o4 (>25%)	8	226	8.9%
Totaal bo	95	2540	100%
		Aantal schoolverlaters	
sbo	39	1020	100%
Totaal peiling	134	3560	

Uitval op leerlingniveau

Tussen begin april en begin juli hebben de afnames plaatsgevonden. Per school is/zijn in de beoogde klas(en) de beoogde taken afgenomen, op één sbo-school (klas) na. Op de 95 bo-scholen en 38 sbo-scholen heeft het overgrote deel van de beoogde leerlingen daadwerkelijk deelgenomen aan de schrijftaken. Op leerlingniveau is er uitval vanwege a) ziekte of afwezigheid vanwege een andere reden, b) het ontbreken van ouderlijke toestemming voor het gehele onderzoek, c) vanwege uitsluiting voorafgaande aan de afname a.d.h.v. de uitsluitingscriteria (zie paragraaf 2.1.2), of d) vanwege het geheel niet realiseren van de beoogde afname (één sbo-school).

Scholen hebben bij de digitale aanlevering van gevraagde leerlinggegevens doorgegeven welke leerlingen aan de exclusiecriteria voldeden. Omdat op basis van deze informatie niet eenduidig kon worden vastgesteld welke leerlingen moesten worden uitgesloten, is voor de zekerheid op de afnamedag nogmaals bij de leerkracht/contactpersoon nagegaan welke leerlingen aan bovenstaande kenmerken voldeden. Op basis van deze informatie zijn de betreffende leerlingen uit het databestand verwijderd. De uitval op basis van de exclusiecriteria is 25 leerlingen (0,7%).

De dataset bevat eveneens geen gegevens van leerlingen die op de afname dag afwezig waren vanwege ziekte, andere (medische of belangrijke) reden of vanwege het ontbreken van de ouderlijke toestemming om deel te nemen. Eén aangemelde sbo-school (15 leerlingen) kon uiteindelijk de afname niet realiseren.

Binnen het peilingsonderzoek is gebruik gemaakt van informed-consent van de ouders. Ouders hebben hiertoe een informatiebrief ontvangen met de mogelijkheid om van deelname af te zien. In totaal 11 leerlingen hadden geen ouderlijke toestemming (0.31% van de aangemelde leerlingen). Van de groep leerlingen mét toestemming was op de afnamedag ongeveer vier procent afwezig vanwege ziekte of andere redenen.

Voor leerlingen met motorische beperkingen was er de mogelijkheid om de taken op de computer uit te voeren. In totaal 28 leerlingen leken hiervoor in aanmerking te komen, waarvan er 5 op de dag van afname niet aanwezig waren. Zeventien leerlingen uit het basisonderwijs en zes leerlingen uit het sbo waren gecategoriseerd als leerling met motorische beperking en hebben wel deelgenomen aan het peilingsonderzoek. Verder was er zowel in bo en sbo sprake van deelname van leerlingen die ten tijde van de aanmelding (verstrekking van de leerlinggegevens) nog niet aanwezig waren op school. Zij hebben gebruik gemaakt van het reservemateriaal.

Tabel 2.15: Uitval van aangemelde leerlingen

	bo (N=2540)		sbo (N=1020)		Totaal	
	n	%	n	%	N	%
Ziek/Afwezig/van school	91	3.6%	73	7.2%	164	4.6%
Geen toestemming	6	0.2%	5	0.5%	11	0.3%
Exclusiecriteria	13	0.5%	12	1.2%	25	0.7%
Geen realisatie	0	-	15	1.5%	15	0.4%
Totale uitval	110	4.3%	105	10.3%	225	6.0%
Extra leerlingen (reservemateriaal)	16		20			
Netto aantal deelnemers	2446		935		3381	

Op de deelnemende reguliere bo-scholen was er in 18 gevallen sprake van twee klassen (groepen 8), in vier gevallen namen 3 klassen deel. Voorts was er één school die deelnam met vier kleine(re) stamgroepen (Montessori onderwijs).¹⁸ In het sbo is het gebruikelijk dat de leerlingen in vergelijking met het regulier basisonderwijs in kleinere groepen onderwijs ontvangen. Scholen kennen daarom vaak meerdere klassen/groepen met schoolverlaters. Elf van de 39 aangemelde sbo-scholen hadden één groep (unit), terwijl de andere 27 twee, drie of zelfs vier (kleinere) groepen hadden.¹⁹

2.4.1 Representativiteit op leerlingniveau

Om de representativiteit van de deelnemende bo en sbo-leerlingen na te gaan is een aantal van hun achtergrondkenmerken vergeleken met die van de landelijke populatie basisscholieren uit groep 8 en voor de schoolverlaters met die van de landelijke populatie scholieren die op het sbo zitten. Voor die vergelijking is gebruik gemaakt van landelijke DUO-gegevens die in bewerkte vorm via de Inspectie van het Onderwijs zijn verkregen. De via de Inspectie van het Onderwijs verkregen gegevens hebben betrekking op het schooljaar 2018/2019, het jaar van de afname van de peiling²⁰.

De Tabellen 2.16 en 2.17 bevatten de verdeling van de leerlingen naar sekse, leeftijd, en leerlinggewicht voor de steekproefscholen en de populatie. De variabele leerlinggewicht is voor het sbo niet opgenomen, aangezien leerlinggewichten in het sbo niet van toepassing zijn.

Analoog aan Tabel 2.15 zou verwacht kunnen worden dat er voor 3381 leerlingen (nummers) één of meerdere schrijftaken (en score) beschikbaar zouden moeten zijn. Analyse van de beoordelingen van de taken laat echter zien dat er voor 3319 leerlingnummers een score beschikbaar is (hoofdstuk 8). Daarnaast zijn voor 39 leerlingen alleen gegevens uit de vragenlijst beschikbaar. Het kan gaan om leerlingen die a) aanwezig waren bij de afname van de vragenlijst, maar die om een medische reden of andere geoorloofde afwezigheid niet (meer) konden deelnemen aan de schrijftaken, b) leerlingen die wel aanwezig waren bij de schrijfoverdrachten maar niets concreets op papier hebben gezet (ontbrekende motivatie), of c) leerlingen waarvan de schrijfproducten geheel niet

¹⁸ Het aantal units 'groep8' in de dataset bedraagt daarom 126 (voor het bo), aangegeven met de identificatievariabele 'Combi_klas_ID' (beginnend met duizendtallen 1,2 of 3).

¹⁹ Het aantal units 'klas schoolverlaters' in de dataset bedraagt daarom 79, eveneens aangegeven met de identificatievariabele 'Combi_klas_ID' (beginnend met duizendtallen 4 en 5). Vanwege de kleine groepen (10-15 leerlingen) is de afname op sbo-scholen soms in één gezamenlijke sessie verlopen (en begeleid door de testleider).

²⁰ Het gaat om landelijke cijfers met als peildatum 1 oktober 2018.

te beoordelen waren. Voor de representativiteitsanalyse (Tabel 2.16a en b en Tabel 2.17) gaan we uit van de 3358 unieke leerlingen die in het totaalbestand zijn opgenomen.

Tabel 2.16a: Deelname bo-leerlingen ten opzichte van landelijke populatie

	Steekproef (N=2438)		Landelijke populatie (N=172798)	
	n	%	n	%
Sekse				
Jongen	1083	49.3%	85800	49.7%
Meisje	1113	50.7%	86998	50.3%
Totaal²¹	2196		172798	
Leeftijd op 1-10-2018				
9 jaar oud	0	-	209	0.1%
10 jaar oud	286	13.0%	20371	11.8%
11 jaar oud	1664	75.4%	129562	75.0%
12 jaar oud	250	11.3%	21991	12.7%
13 jaar oud	6	0.3%	665	0.4%
Totaal⁵	2206		172798	
Gemiddelde leeftijd	10,99 jaar		11,01 jaar	
Leerlinggewicht				
Geen gewicht ²²	2283	94.5%	157483	91.1%
Gewicht 0,3	61	2.5%	7716	4.5%
Gewicht 1,2	72	3.0%	7599	4.4%
Totaal	2416		172798	

Zoals uit Tabel 2.16a kan worden afgeleid, is het percentage jongens en meisjes in het bo-deel van de peiling redelijk in lijn met de verwachtingen. Het verschil met de populatie is niet significant $\chi^2(1, N = 172798) = 0.098, p = 0,754$. Wat de verdeling naar leeftijd betreft geldt dat voor de bepaling van de representativiteit is uitgegaan van de leeftijd van de leerlingen uit de steekproef op 01-10-2018. De gegevens van de populatie hebben betrekking op groep 8 leerlingen in het schooljaar 2018/2019 en daarbij is dezelfde peildatum van de leeftijd gehanteerd. De verschillen zijn niet significant (bij een gehanteerde waarde van $p < 0,01$), $\chi^2(4, N = 172798) = 9.93, p = 0,053$. Bij analyses naar geslacht en leeftijd (op een specifieke peildatum) hoeft dus niet te worden teruggewogen.

²¹ Geboortedatum en geslacht zijn voorafgaande aan de afnamedatum opgevraagd. Het hoge percentage leerlingen in het bo waarvoor geen geslacht en/of leeftijd bekend is heeft o.a. te maken met het feit dat scholen alleen geanonimiseerd aan de peiling wilden deelnemen. Geslacht en leeftijd zijn voor respectievelijk 242 (9.9%) en 232 (9,5%) leerlingen niet bekend. In de vergelijking met de populatie gaan we uit van het deel waar wel gegevens voorhanden zijn.

²² Leerlinggewicht (als ook schooladvies) zijn pas op de afnamedatum opgevraagd. Van slechts één school is zeker dat het leerlinggewicht onbekend is (22 leerlingen, 1% van het totaal), omdat de aanvullende lijst ontbreekt. Voor alle andere scholen geldt dat bij het ontbreken van een notitie ervan uitgegaan wordt dat er aan de leerling geen formatiegewicht is toegekend. Bij de meerniveau-analyses is echter wel gebruik gemaakt van imputaties van de missende waarden. Ter controle van de imputaties per school zijn DUO-data aangaande percentage leerlingen met formatiegewicht per jaargroep gebruikt.

Voor de variabele formatiegewicht lijkt het er in eerste instantie op dat er in de peiling sprake is van een ondervertegenwoordiging van de 0.3 en 1.2 leerlingen. Het verschil is significant, $\chi^2(2, N = 172798) = 34.095, p = 0,000$. Voor analyses tussen groepen leerlingen met en zonder gewicht zou dus in principe gewogen moeten worden.

Een kruistabel van de leerlingen met de gewichtenstrata op schoolniveau laat echter zien dat bij 11 (55%) van de 20 S4o3 scholen (scholen met >10-25% leerlingen met een leerlinggewicht) aan geen enkele leerling een gewicht is toegekend. Een visuele inspectie van de administratieformulieren laat zien dat er inderdaad niets omtrent gewicht is genoteerd. Voor stratum S4o4 (>25%) geldt dat voor twee van de acht scholen (25%). Naar alle waarschijnlijkheid is er dus sprake van een onderschatting van het aantal leerlingen met een gewicht in de peiling.²³ Door het aantal gewichtenleerlingen te schatten voor de S4o3 en S4o4-scholen kan een alternatieve representativiteitsanalyse worden gedaan (Tabel 2.16b).²⁴ In de tabel is aantal gewichtenleerlingen aangevuld met een schatting voor de ontbrekende scholen op basis van de gemiddelde aantallen 0.3 en 1.2 leerlingen in de (9) S4o3 en (6) S4o4 scholen waarvoor wel gegevens beschikbaar zijn. Er is dan nog steeds sprake van een kleine ondervertegenwoordiging ten opzichte van de populatie, echter is verschil niet significant, $\chi^2(2, N = 172798) = 4.206, p = 0,122$.

Tabel 2.16b: Deelname bo-leerlingen met geschatte aantallen gewichtenleerlingen

Leerlinggewicht				
Geen gewicht	2228	92.2%	157483	91.1%
Gewicht 0,3	89	3.7%	7716	4.5%
Gewicht 1,2	98	4.1%	7599	4.4%
Totaal	2416		172798	

Uit Tabel 2.17 blijkt dat ook de percentages jongens en meisjes in het sbo-deel van de peiling redelijk in lijn der verwachting liggen. Het verschil met de populatie is niet significant $\chi^2(1, N = 6979) = 0.200, p = 0,655$. Wat de verdeling naar leeftijd betreft geldt ook hier peildatum 01-10-2018. De sbo-leerlingen zijn gemiddeld gezien slechts enkele dagen ouder dan in de landelijke populatie. De verschillen zijn niet significant (bij een gehanteerde waarde van $p < 0,01$), $\chi^2(4, N = 6979) = 3.84, p = 0,429$. Bij analyses naar geslacht en leeftijd (op een specifieke peildatum) hoeft dus ook voor het sbo deel niet te worden gewogen.

²³ Niet in alle gevallen zal het leerlinggewicht bij de betreffende leerkracht bekend zijn of was de schooladministratie op de dag toegankelijk. Verder is het – gezien het ontbreken van gegevens bij geslacht en geboortedatum – denkbaar dat scholen de informatie over gewicht vanwege privacy redenen niet wilden geven. Voor die gevallen geldt derhalve dat het onzeker is of de leerling daadwerkelijk geen formatiegewicht heeft of dat formatiegewichten niet zijn genoteerd.

²⁴ Ook op een groot deel van de S4o2 scholen (max 10%) zijn volgens de data geen gewichtenleerlingen aanwezig in groep 8. Het is hier nog meer de vraag of op deze scholen de gegevens niet voorhanden waren, de gegevens niet zijn genoteerd of dat er daadwerkelijk geen leerling met een gewicht in groep 8 zat. Het schatten van eventuele extra gewichtenleerlingen voor deze groep scholen is uiterst complex. De school behoort immers al met slechts één leerling met formatiegewicht - in één van de leerjaren – tot dit stratum. Om die reden hebben we keuze gemaakt alleen te schatten voor de S4o3 en S4o4 scholen.

Tabel 2.17. : Deelname sbo-leerlingen ten opzichte van landelijke populatie

	Steekproef		Landelijke populatie	
	n	%	n	%
Sekse				
Jongen	595	67.0%	4629	69.9%
Meisje	292	33.0%	2350	30.1%
Totaal²⁵	887	-	6979	
Leeftijd op 1-10-18				
9 jaar oud	0	-	2	0%
10 jaar oud	25	2.8%	140	2.0%
11 jaar oud	426	48.2%	3506	50.2%
12 jaar oud	423	47.4%	3230	46.3%
13 jaar oud	13	1.5%	101	1.4%
Totaal²³	877		6979	
Gemiddelde leeftijd	11,48 jaar		11,47 jaar	

²⁵ Voor 33 sbo-leerlingen (2,9%) is het geslacht niet bekend. Het gaat om twee scholen die zonder gebruik van achtergrondgegevens wensten deel te nemen. Voor nog eens 10 extra leerlingen is ook geen geboortedatum bekend (4.7% van het totaal).

Hoofdstuk 3. Design en afname

In dit hoofdstuk worden het design en de afname van de peiling beschreven. In paragraaf 3.1 gaan we allereerst in op de uitgangspunten voor het design. In paragraaf 3.2 wordt in algemene termen besproken welke taken de leerlingen hebben uitgevoerd. In paragraaf 3.3 worden de twee gebruikte toetsdesigns weergegeven, waarna in paragraaf 3.4 de aantallen observaties ter sprake komen. De opzet, voorbereiding en uitvoering van de hoofdafnames worden beschreven in de paragrafen 3.5 tot en met 3.6.

3.1 Uitgangspunten

De peiling schrijfvaardigheid had tot doel om een beeld te geven van de schrijfvaardigheid van groep 8 leerlingen in het reguliere basisonderwijs en van schoolverlaters in het speciaal basisonderwijs. Dit is gedaan door bij de leerlingen een vragenlijst af te nemen en vier (bo) respectievelijk twee (sbo) schrijftaken af te nemen. Om ervoor te zorgen dat de schrijfvaardigheid met behulp van een diversiteit aan schrijftaken kon worden gemeten, is er gewerkt met een design voor afname en toetsing waarin iedere leerling slechts een deel van alle opdrachten hoefde uit te voeren. In dit hoofdstuk geven we een verantwoording van de overwegingen en keuzes die gemaakt zijn wat betreft het onderzoeksdesign, de selectie van de taken en het toetsdesign. Bij het opstellen van dit design is rekening gehouden met ervaringen die zijn opgedaan in een pilot.

Het streven was om voor de peiling schrijven 12 schrijftaken te ontwikkelen, die 6 teksthandelingen meten (zie hoofdstuk 1). Vaststaand gegeven was dat er 3 ankertaken uit de vorige peiling in het huidige instrumentarium moesten worden opgenomen, zodat een trendvergelijking kan worden gemaakt. Er dienden dus nog negen nieuwe taken te worden ingezet, zodat per teksthandeling twee verschillende schrijftaken konden worden opgenomen (waaronder dus de drie ankertaken). De zes teksthandelingen zijn:

- 1) vertellen
- 2) beschrijven
- 3) uitleggen
- 4) instrueren
- 5) verzoeken
- 6) argumenteren

Om te komen tot 12 schrijftaken die goed aansluiten bij het schrijf- en belevingsniveau van leerlingen zijn per taak twee alternatieven ontwikkeld, die vooraf getest zijn in een pilot. Deze twee alternatieven waren bedoeld om parallel te zijn qua moeilijkheid, maar verschillend qua thematiek of onderwerp. Deze pilot vond plaats eind januari/februari 2019 en had gevolgen voor het design.

Resultaten pilot

Voorafgaand aan de hoofdafname heeft een pilot op zes bo-scholen (zes klassen) en op twee sbo-scholen (schoolverlaters uit drie klassen per school) in Noord-Nederland plaatsgevonden (eind januari/begin februari 2019). Gezien hun bewezen bruikbaarheid zijn de ankertaken niet betrokken in

de pilots. Voor elk van de overige negen taken zijn twee varianten uitgetoetst. Deze taken waren vergelijkbaar maar hadden een ander onderwerp. In totaal zijn er dus 18 taken uitgetoetst. Doel was om na te gaan of de taken in de voorgestelde vorm geschikt waren, welke van de alternatieven het meest geschikt zou zijn, tegen welke problemen er aan gelopen zou worden en hoe lang leerlingen per taak nodig zouden hebben. Daarnaast kon worden nagegaan in hoeverre alle taken bruikbaar waren voor de leerlingen in het sbo.

Om de taken zo breed mogelijk uit te zetten, is ervoor gekozen om in het bo zes taken per leerling af te nemen – iedere leerling maakte dus één van de twee alternatieve taken per teksthandeling. In verband met de belastbaarheid van leerlingen in het sbo zijn hier vier taken per leerling afgenomen. De afname was bij beide schooltypen verspreid over twee blokken, van ieder 45 minuten. Daarbij werd het eerste blok in de ochtend en het tweede in de middag afgenomen. Na afname van het eerste blok werd nog een vragenlijst afgenomen. Alle aan de pilot deelnemende leerlingen (bo en sbo) konden deze vragenlijst binnen de tijd (15-20 minuten) invullen en er waren geen vragen die aangepast dienden te worden (i.v.m. begrijpelijkheid of iets dergelijks). De pilot werd tevens benut om de zogenaamde laptop-conditie uit te proberen. Beoogd werd om in elke klas één leerling op een laptop de schrijftaken te laten maken. Dit zou nadere analyses over het schrijfproces toelaten. Per blok van drie taken werd in het bo één leerling per klas random aangewezen om de taken op de laptop te maken.

De pilot had relevante resultaten voor het op te stellen design. Ten eerste bleek het maken van zes (bo), respectievelijk vier (sbo) taken te vermoeiend. De indruk bestond tevens dat leerlingen de neiging hadden weinig op te schrijven bij de laatste taken (bijvoorbeeld door vermoeidheid of motivatieproblemen of omdat ze in tijdnood kwamen), waardoor een betrouwbare meting van hun schrijfvaardigheid in het geding kwam. Ten tweede is zowel op basis van observaties, maar met name op basis van de gerealiseerde schrijfproducten, besloten dat het voor alle leerlingen in het sbo mogelijk zou moeten zijn om alle taken te maken. Ook bij de taken die op voorhand moeilijk waren ingeschat bleken leerlingen met het laagste uitstroomprofiel voldoende tekst te schrijven. Het niet meenemen van deze ‘moeilijke’ taken in het sbo zou een onderschatting van de mogelijkheden van deze leerlingen betekenen, hetgeen onwenselijk zou zijn. Wel bleek uit de observaties dat het lezen van de taak zelf bij leerlingen met een lage leesvaardigheid onevenredig veel tijd in beslag nam, zeker bij de taken waar leerlingen op basis van teksten moesten samenvatten of argumenteren. Ten derde was de uitvoering van de laptopconditie enigszins complex. Hoewel de testleiders getraind waren op omgang met het programma op de laptop en daarbij schriftelijke instructie (een stappenplan) konden gebruiken, bleken de testleiders in de praktijk de taken voor de gehele klas moeilijk te kunnen combineren met het begeleiden van de leerling die met de laptop werkte – zeker in het geval van haperingen bij de laptop of het invulprogramma.

Na de pilot is het volgende besloten:

- Iedere leerling in het bo maakt vier schrijftaken (per taak maximaal 20 minuten, twee taken per schrijfsessie van 45 minuten);
- Iedere leerling in het sbo maakt twee schrijftaken (per taak maximaal 20 minuten, twee taken in een schrijfsessie van 45 minuten);
- Aan de sbo-leerlingen worden dezelfde 12 schrijftaken voorgelegd als aan de bo-leerlingen.
- Leerlingen in het sbo met een zeer laag leesniveau (lager dan AVI E3-niveau) krijgen de opdrachten voorgelezen door de eigen leerkracht;
- Iedere leerling, in beide schooltypen, vult de vragenlijst in (afnametijd 15-20 minuten);

- Om een zorgvuldige en vergelijkbare afname in alle klassen te kunnen waarborgen is afgezien van het gebruik van laptops door één of twee leerlingen in de klas. Alleen leerlingen met motorische problemen mochten op een computer werken (zie hoofdstuk 2).

De uiteindelijke afname heeft plaatsgevonden tijdens een ochtendsessie van ongeveer één uur (45 minuten voor twee taken en 15 minuten voor de vragenlijst) en, alleen in het bo, een middagsessie van 45 minuten (opnieuw twee taken).

3.1.1 Uitgangspunten design

Voor het op te stellen design stond een aantal uitgangspunten centraal, waarbij rekening is gehouden met ervaringen uit de pilot. Het voorgestelde design moest recht doen aan de volgende eisen:

- De beschikbare afnametijd is beperkt. Voor een peiling is in principe maximaal twee uur afnametijd bij leerlingen beschikbaar. Dit heeft tot gevolg dat niet alle taken door alle leerlingen konden worden uitgevoerd en dat er gekozen moest worden voor een onvolledig gelinkt design;
- Er moesten voldoende waarnemingen per taak zijn om generalisatie naar de populaties toe te staan. Gezien de tijdens de pilot ervaren vermoeidheidsverschijnselen bij de leerlingen is nagegaan of er voldoende waarnemingen zouden zijn op het moment dat de leerlingen minder taken zouden uitvoeren. Dit bleek het geval. Als gevolg daarvan hebben leerlingen in het bo vier schrijftaken gemaakt en in het sbo twee schrijftaken;
- Taken moesten in verschillende volgorde worden aangeboden om volgorde-effecten of eventuele morgen-middag-verschillen op het spoor te kunnen komen;
- Ook in het sbo moest het design zodanig zijn dat we *Differential Item Functioning (DIF)* op het spoor zouden kunnen komen;
- Het design moest praktisch uitvoerbaar zijn. Omdat er werd gewerkt met schrijfboekjes, waarop unieke leerlingnummers bedrukt zouden worden en niet alle scholen leerlingnamen wilden delen, wilden we niet het risico lopen dat schrijfboekjes door niet-beoogde leerlingen zouden worden gemaakt. Gekozen is daarom om niet te roteren binnen klassen (alle versies van de boekjes worden verdeeld over het aantal leerlingen in de klas), maar wel over klassen (elke leerling in de klas ontvangt een boekje met dezelfde schrijfp opdrachten).

3.2 Algemene opzet design

Binnen de peiling zijn in totaal twee hoofdkenmerken gemeten. Het eerste hoofdkenmerk is de schrijfvaardigheid, gemeten aan de hand van 12 verschillende schrijftaken. Leerlingen hadden per schrijftaak 20 minuten de tijd, afgenomen in blokken van twee schrijftaken (met 5 minuten transitietijd om van de ene naar de andere taak over te gaan). In het bo zijn twee blokken van twee schrijftaken afgenomen (ochtend- en middagsessie), terwijl er in het sbo één blok bestaande uit twee schrijftaken werd afgenomen. De secundaire hoofdkenmerken waren de frequentie van (het zelf) schrijven en de attitude en self-efficacy ten opzichte van (het vak) schrijven zoals deze door de leerlingen zelf wordt gerapporteerd. Dit is gemeten met een vragenlijst, die in principe door alle leerlingen is gemaakt. Het invullen van de vragenlijst nam ongeveer 15-20 minuten in beslag.

3.2.1 Toetsdesign voor de vragenlijst

Voor de vragenlijst is een eenvoudig design opgesteld: alle leerlingen, ongeacht schooltype, krijgen alle vragen voorgelegd. De vergelijkbaarheid van de kandidaten is op de (sub)testscores bepaald.

3.2.2 Design voor de schrijftaken (bo)

Tijdens het onderzoek werd het afnamesdesign gehanteerd dat is weergegeven in Tabel 3.1. Elke kolom beschrijft een taak, waarbij bijvoorbeeld '3-2' staat voor taak 3 die teksthandeling 2 (=Informatief beschrijven) betreft. Het gaat hier om de Unicef taak. Kolom 4-2 beschrijft de alternatieve taak behorende bij dezelfde teksthandeling (taak 4, 'brief aan Annika'). Elke rij beschrijft de vier schrijftaken in een boekje. In de cellen staat de positie van de taken in het boekje. Boekje 1 en 2 bevatten bijvoorbeeld dezelfde vier schrijftaken, maar in twee verschillende volgordes van het aanbieden van de taken. Er zijn voor de zes combinaties van vier schrijftaken steeds twee varianten met een verschillende volgorde van schrijftaken samengesteld. Alle leerlingen in een klas maakten een van de volgende twaalf boekjes met elk vier schrijftaken.

Tabel 3.1: Afnamesdesign taken

Taak/Teksthandeling	1-1	3-2	5-3	7-4	9-5	11-6	2-1	4-2	6-3	8-4	10-5	12-6
Boekje 1	1	2	3	4								
Boekje 2	4	3	2	1								
Boekje 3			1	2	3	4						
Boekje 4			4	3	2	1						
Boekje 5					1	2	3	4				
Boekje 6					4	3	2	1				
Boekje 7							1	2	3	4		
Boekje 8							4	3	2	1		
Boekje 9									1	2	3	4
Boekje 10									4	3	2	1
Boekje 11	3	4									1	2
Boekje 12	2	1									4	3

Dit design voldoet aan de volgende, relevant geachte randvoorwaarden:

- Elke taak komt precies één keer op iedere positie (1,2,3,4) voor;
- De combinatie teksthandelingen f1-f2-f3-f4, f1-f2-f5-f6 en f3-f4-f5-f6²⁶ komt ieder vier keer voor;
- Het is mogelijk om voor iedere taak ongeveer even veel observaties krijgen. Op voorhand werd uitgegaan van circa 720 observaties per taak;
- Indien een taak op de 4^e positie aanmerkelijk minder schrijfvaardigheid zou uitlokken vanwege bijvoorbeeld vermoeidheidsverschijnselen of motivatieproblemen bij de leerlingen, dan

²⁶ Hierbij staat f1-f2-f3-f4 voor de combinatie teksthandeling 1, teksthandeling 2, teksthandeling 3, teksthandeling 4. Hetzelfde geldt voor de overig genoemde combinaties.

zouden deze observaties eventueel kunnen worden verwijderd en zouden er toch nog 540 observaties per taak overblijven. Dit aantal zou voldoende zijn.

Overwogen is om de ankertaken op een speciale positie in het design te plaatsen (bijvoorbeeld altijd als eerste in een boekje). Hier is van afgezien, omdat dit a) het design zou compliceren en b) de ankertaken bij de vorige peiling ook niet op een speciale positie stonden, waardoor de vergelijkbaarheid in het geding zou komen.

3.2.3 Design voor de schrijftaken (sbo)

Zoals hierboven beschreven liet de pilot zien dat de afname van vier taken op één afnamedag (te) veeleisend was voor de sbo-leerlingen. Evenals in de vorige peiling is er daarom voor gekozen om bij de sbo-leerlingen twee taken af te nemen, gedurende één ochtendsessie van 45 minuten. De afnametijd en –procedure was wel vergelijkbaar aan die in het bo: leerlingen zouden 2 keer 20 minuten krijgen om twee schrijfoopdrachten te maken. Voor het sbo is het afnamedesign gehanteerd dat is weergegeven in Tabel 3.2. De tabel is op dezelfde manier te lezen als Tabel 3.1 (zie de beschrijving op de vorige pagina).

Tabel 3.2: Design voor afname taken in het sbo (voor legenda zie Tabel 3.1)

Teksthandeling	1-1	3-2	5-3	7-4	9-5	11-6	2-1	4-2	6-3	8-4	10-5	12-6
Boekje 1	1	2										
Boekje 2		1	2									
Boekje 3			1	2								
Boekje 4				1	2							
Boekje 5					1	2						
Boekje 6						1	2					
Boekje 7							1	2				
Boekje 8								1	2			
Boekje 9									1	2		
Boekje 10										1	2	
Boekje 11											1	2
Boekje 12	2											1

Ook in het sbo kwam iedere taak op de eerste en laatste positie voor. Met dit design waren er naar verwachting circa 150 observaties per taak in het sbo. Ook in dit design maakte iedere leerling taken met verschillende teksthandelingen en was de volgorde in taken per boekje verschillend: iedere taak werd een keer als eerste en een keer als laatste aangeboden.

3.3 Toetsdesign over vaardigheden heen

In Tabel 3.3 is het toetsdesign weergegeven over de vaardigheden heen, in volgorde van afname. Op alle scholen kende de afname dezelfde volgorde. Er werd gestart met de afname van de vragenlijst bij alle leerlingen, waarna respectievelijk schrijfsessie 1 ('s ochtends) en – in het bo - schrijfsessie 2 ('s middags) werd afgenomen. De vragenlijst is op alle scholen als eerste afgenomen, aangezien het de bedoeling was dat de leerlingen, ongeacht hun gepercipieerde prestaties op de toetsen, een mening

zouden geven over hun attitude ten opzichte van en ervaring (buitenschools) met schrijven. Omdat het schrijven van teksten vaak veeleisend is voor leerlingen, is besloten om de schrijfsessies in het bo te verdelen over een ochtend- en een middagsessie.

Tabel 3.3: Samenstelling van de afnamen (bo)

Versie	Vragenlijst	Schrijfsessie 1	pauze	Schrijfsessie 2
Versie 1	VL	1-1 en 3-2		5-3 en 7-4
Versie 2	VL	7-4 en 5-3		3-2 en 1-1
.....				
Versie 12	VL	3-2 en 1-1		12-6 en 10-5

Een school heeft in principe altijd één versie/boekje met schrijfopdrachten toegewezen gekregen. In het geval van meerdere klassen zijn verschillende versies/schrijfboekjes bij de verschillende parallelklassen afgenomen.

3.4 Aantallen Leerlingen

Binnen de voorgestelde designs zou iedere bo-/sbo-school één van de twaalf versies/schrijfboekjes toegewezen krijgen. Voor het bo en het sbo is separaat een verdeling van versies naar scholen gemaakt. Op basis van de 93 bo-scholen die zich in eerste instantie hadden aangemeld betekende dit dat iedere versie/schrijfboekje op 9 tot 11 scholen aangeboden zou worden. Daarbij is rekening gehouden met het expliciete stratificatiecriterium ‘percentage leerlingen met een formatiegewicht op school’. Bij de toewijzing van de versies is ernaar gestreefd binnen ieder stratum alle twaalf versies af te nemen, waardoor iedere versie dus in ieder geval waarnemingen bevatte uit de vier strata die de basis voor de streekproeftrekking voor het bo werden onderscheiden. De proportionele verdeling van scholen over de strata is daarbij per versie zoveel mogelijk aangehouden. Er is daarnaast zoveel mogelijk geprobeerd om bij de toewijzing van de versies aan de scholen ervoor te zorgen dat het aantal waarnemingen per versie vrijwel even groot zou zijn (op basis van school-/klasgrootte). Voor het sbo was het expliciete stratificatiecriterium schoolgrootte. Bij de verdeling van de versies over de 39 sbo-scholen is hiermee (en met klasgrootte) rekening gehouden. De versies zijn aan 5 tot 8 klassen toegewezen, op een dusdanige wijze dat er ongeveer evenveel waarnemingen per versie zouden zijn.

Door de gewijzigde afnameperiode van de peiling kon bij het opstellen van het design worden uitgegaan van reeds door de scholen aangeleverde leerlinggegevens om schattingen te maken over aantallen te toetsen leerlingen. In het bo zouden de twaalf versies volgens verdeling worden afgenomen bij 801 tot 933 leerlingen per taak, terwijl dit in het sbo het geval was bij 148 tot 200 leerlingen. Het gemiddelde beoogde aantal observaties per schrijftaak is voor de meest toepassingen ruim te noemen, zeker bij het toepassen van een IRT-model (zie de COTAN richtlijnen²⁷).

²⁷ COTAN Beoordelingssysteem voor de kwaliteit van tests (Evers, Lucassen, Meijer & Sijtsma, 2010; <https://www.psynip.nl/wp-content/uploads/2016/07/COTAN-Beoordelingssysteem-2010.pdf>)

Hoofdstuk 4. Voorbereiding en uitvoering van de peiling

Een grootschalige nationale peiling in het basisonderwijs vraagt om een gedegen voorbereiding. In dit hoofdstuk beschrijven we de verschillende stappen in de voorbereiding van de afnames op de scholen. De voorbereiding heeft plaatsgevonden van najaar 2018 tot en met maart 2019. In de eerste plaats wordt in dit hoofdstuk een beschrijving gegeven van hoe de testleiders zijn geworven en getraind en welke taken zij hadden. Daarna wordt een beschrijving gegeven van de opzet en globale aanpassingen naar aanleiding van de pilotafnames die eind januari/begin februari 2019 hebben plaatsgevonden. Ook de logistieke voorbereiding wordt beschreven.

4.1 Werving en selectie van testleiders

De uitvoering van de peiling in de praktijk is uitgevoerd door zeventien testleiders die allen werden aangestuurd door de onderzoekers. De testleiders hadden een centrale rol in de peiling als het aanspreekpunt voor schoolleiders en leerkrachten en als verantwoordelijke voor een betrouwbare en systematische afname van de schrijftaken en de vragenlijst. Voor de rol van testleider zijn personen geselecteerd met een ruime ervaring in het onderwijs en/of in het afnemen van toetsen/taken op (po-)scholen. Zij zijn afkomstig uit het netwerk van de onderzoekers en hebben veelal meegeholpen bij de uitvoering van eerdere nationale/internationale peilingen. Er was sprake van een ruime en goede ervaring met deze testleiders.

4.2 Training van (hoofd)testleiders

In de voorbereiding op de training is voor alle betrokken testleiders een handleiding ontwikkeld. In deze handleiding werd achtergrondinformatie over het peilingsonderzoek gegeven (doel en opzet van het onderzoek) en informatie over deelname aan de huidige peiling Schrijfvaardigheid. Het tweede deel van de handleiding bestond uit protocollen en beschrijvingen van de manier waarop de testleider zich voorafgaand aan de afname diende voor te bereiden, alsmede van algemene richtlijnen voor, tijdens en na de afname. Deze richtlijnen waren met name gericht op het maken van afspraken voor de afname, op de nog benodigde informatie van de scholen en op logistieke aspecten. Daarna werd een overzicht van de gebruikte instrumenten gegeven en een uitgebreide richtlijn per af te nemen opdracht. Per opdracht is beschreven welke materialen nodig waren voor de opdracht, welke procedure gevolgd moest worden, welke instructies aan de leerlingen moesten worden gegeven en hoe testleiders om moesten gaan met (veelvoorkomende) vragen. Dit geheel was aangevuld met een overzicht van aandachtspunten en tips.

Om de betrouwbaarheid van de metingen te vergroten zijn alle testleiders, voorafgaand aan de (pilot)metingen, getraind zodat zij de taken volgens de opgestelde protocollen zouden afnemen. Deze training nam één dag in beslag. De training startte met algemene voorlichting omtrent de peiling schrijfvaardigheid, waarna de instrumenten, het design en de richtlijnen voor de (voorbereiding van de) afnames uitgebreid werden besproken. Omdat de materialen via de post werden verzonden naar de scholen en nog niet alle benodigde informatie van de scholen bekend was, was er tevens speciale

aandacht voor de logistieke afhandeling en het verzamelen van extra achtergrondgegevens van de deelnemende leerlingen.

In aanloop naar de pilots zijn drie noordelijke testleiders gedurende een middag getraind in de afname van de schrijftaken, de vragenlijst en een laptopconditie (zie paragraaf 4.3). Deze training vond eind januari 2019 plaats. Voorafgaand aan de hoofdafnames hebben twee volledige trainingdagen plaatsgevonden, waaraan alle testleiders gedurende één dag hebben deelgenomen. Twee testleiders die reeds ten behoeve van de pilot waren getraind hebben een telefonische update van de groepstraining gehad.

4.3 Pilotonderzoek

De pilotafnames hadden tot doel om alle 18 nieuwe schrijfoopdrachten uit te testen op doorlooptijd en geschiktheid (ook voor sbo). Voor bo en sbo werden vier toetsversies geconstrueerd, in het bo bestaande uit 6 opdrachten en in het sbo uit vier opdrachten. Omdat er in het sbo minder taken konden worden gemaakt en de pilot hier ook tot doel had na te gaan in hoeverre sbo-leerlingen in staat waren om de moeilijker geachte taken te maken, is ervoor gekozen om de amuserende taak (T1) in het sbo niet af te nemen en de instructieve taak (T7) slechts in twee versies op te nemen. Ook de leerlingvragenlijst is afgenomen om na te gaan of de beschikbare tijd toereikend was en of er geen lastige of multi-interpretabele items in de vragenlijst zaten. Tijdens een klassikale bespreking van de vragenlijst konden leerlingen aangeven welke vragen ze niet snapten. Ten slotte is geoefend met de rol van testleider. Ook werd de pilot gebruikt om de zogenaamde laptop-conditie uit te proberen: in het bo heeft per blok van 3 taken één leerling per klas de taken op de laptop gemaakt.

Na afloop van de pilot hebben er evaluaties plaatsgevonden. Allereerst moest er worden besloten welke van de alternatieve opdrachten in het peilingsinstrumentarium zou worden opgenomen. Primair beslissingscriterium was daarbij de lengte van de geschreven teksten, omdat deze een indicator is van de kwaliteit van teksten (Hacquebord & Lenting-Haan, 2012; Espin, 1999) en van de aansluiting bij de belevingswereld van de leerlingen. Van de twee alternatieven werd de taak met het meeste aantal woorden geselecteerd. Andere consequenties van de pilot voor het design staan beschreven in hoofdstuk 3.

4.4 Voorbereiding van de scholen

Voorafgaand aan de daadwerkelijke peilingsafnames zijn enkele leerlinggegevens door de onderzoekers bij de contactpersonen op de scholen opgevraagd en hebben de testleiders contact met hen opgenomen om de afnamedatum af te spreken en eventuele vragen te beantwoorden. Deze afspraken zijn in februari en maart 2019 gemaakt. De testleiders bespraken telefonisch de volgende onderwerpen met de contactpersoon: gebruik van de aangeleverde toestemmingsbrief aan de ouders en check of deze was meegegeven/verstuurd aan ouders, informatie over de aanlevering van materialen via de post (en handelingswijze bij ontvangst), afspraken over het meedoen van groep-7 leerlingen in combinatieklassen en eventuele groep-8 leerlingen die niet mee konden doen (als gevolg van uitsluitingscriteria of geen toestemming) en de verplichte aanwezigheid in de klas van de leerkracht tijdens de afnames.

Toelichting op de ouderlijke toestemming

De ouders van leerlingen is gevraagd naar toestemming voor deelname aan het peilingsonderzoek via 'informed consent'. Hiertoe is reeds in het najaar een brief door de onderzoekers aan de scholen ter beschikking gesteld, met het verzoek deze met de ouders te delen. In verband met de gewijzigde planning van het onderzoek hebben de deelnemende scholen nogmaals een vergelijkbare brief ontvangen, waarin deze gewijzigde planning was opgenomen. Beide brieven hadden, afgezien van de genoemde periode van onderzoek, dezelfde inhoud. Er werd beknopt informatie gegeven over het onderzoek, over de manier waarop met gegevens zou worden omgegaan (conform AVG) en hoe ouders bezwaar konden maken. Bezwaar dienden ouders kenbaar te maken bij de leerkracht. De testleiders zijn in het voorafgaande telefoongesprek nagegaan of er leerlingen waren die geen toestemming hadden en hebben met de contactpersoon gesproken over eventuele alternatieve taken/oplossingen (zoals aanschuiven bij een andere klas). Op de dag van afname is nogmaals nagegaan welke leerlingen geen ouderlijke toestemming hadden. Uit paragraaf 2.4 blijkt dat het gaat om geringe aantallen: 6 in het bo en 5 in het sbo.

Vorbereiding materialen leerlingen

Scholen hebben voorafgaand aan het onderzoek informatie aangeleverd. Het ging daarbij om klassenlijsten met daarin de klasnaam, naam van de leerkracht en de leerlingnamen of -nummers van de groep 8-leerlingen en schoolverlaters, hun geboortedatum en geslacht. Ook bevatte deze klassenlijst een kolom waarin kon worden genoteerd welke leerlingen aan de vastgestelde uitsluitingscriteria voldeden of dat er sprake was van motorische beperkingen, zodat de afname op de klassencomputer moest plaatsvinden. In de laatste kolom was ruimte om per leerling bijzonderheden te noteren. Met de verkregen gegevens zijn leerlingnamen gekoppeld aan barcodes ten behoeve van de schrijfboekjes en de vragenlijst. De leerlingnummers (in de vorm van barcodes) bestonden uit zeven cijfers: de eerste vier corresponderen met het schoolvolgnummer, de vijfde met een klascode en de laatste twee met het volgnummer in de leerlingenlijst. Ook zijn per klas twee reserveleerlingen (met leerlingnummers eindigend op 98 en 99) toegevoegd. Op zowel de vragenlijst als de boekjes met schrijftaken (waarin leerlingen ruimte hadden om hun teksten te schrijven) zijn de leerlingcodes gedrukt. Met deze nummers werd het tevens mogelijk voor de school, wanneer zij dat wenste, op geanonimiseerde manier deel te nemen aan de peiling. Aan de leerlingnummers werd dan in plaats van voor- en achternaam een voor de school herkenbare code gekoppeld (bijvoorbeeld het interne leerlingnummer van de school), in combinatie met de achtergrondkenmerken zoals hierboven genoemd.²⁸

4.5 Organisatie van de afnames

Iedere school was toebedeeld aan een testleider. De planning en de organisatie van de afnames op de scholen werden door de testleiders zelf gedaan. Testleiders hadden hiertoe contactinformatie van de hen toebedeelde scholen ontvangen. Op één afnamedag konden maximaal twee klassen op één school worden bezocht (een planningssuggestie stond beschreven in de handleiding). In het geval van drie of

²⁸ De (voor)namen van leerlingen zijn alleen op de schrijfboekjes geprint om het uitdelen gemakkelijker te maken, zodat het beoogde boekje bij de juiste leerling (geboortedatum, geslacht) terecht kwam. De namen werden na de afname van de boekjes geknipt. In alle bestanden zijn geen namen of andere naar unieke personen of scholen herleidbare informatie opgenomen.

meer klassen op een school zou de afname over meerdere dagen plaatsvinden. In het sbo heeft de afname regelmatig niet in de eigen klas plaatsgevonden. De samenstelling van groepen is hier vaak “voorgesorteerd” op het vervolgonderwijs en klassen tellen soms slechts enkele schoolverlaters. In dergelijke gevallen was er sprake van gezamenlijke afname van leerlingen uit meerdere klassen.

Ten behoeve van de afnamen hebben testleiders het volgende ontvangen: een lege usb-stick die kon worden gebruikt voor leerlingen die hun schrijftaken op een computer in de klas moesten maken (vanwege een motorische beperking) en een usb-stick met daarop een PowerPoint presentatie die tijdens de afname kon worden gebruikt (met een overzicht van de dag en beknopte informatie over afnameprocedure). Ook stond er een timer op de usb-stick, die kon worden gebruikt, zodat leerlingen konden zien hoeveel tijd ze nog hadden voor het maken van een schrijftaak.

De benodigde materialen werden in dozen via de post aan de scholen gestuurd, met een op de doos geplakte brief aan de contactpersoon. De dozen bevatten:

- > Schrijfboekjes op naam en nummer
- > Vragenlijsten voor de leerlingen op naam en nummer
- > Vragenlijsten voor de leerkrachten van groep 6, 7, 8
 - Antwoordenvelop voor retourzending leerkrachten groep 6 en 7
- > Retoursticker
- > Per klas 2 leerlinglijsten
 - Eén om additionele informatie te noteren t.b.v. het onderzoek
 - Eén om als sleutel voor de terugrapportage te dienen (voor de school)²⁹

4.6 Verloop van een gemiddelde afnamedag

De regie op de afnamedag lag bij de testleider, die zorgde voor een goed verloop van de peilingstaken en een goede communicatie met de school. Na aankomst op de school zocht de testleider contact met de contactpersoon. Gevraagd werd naar de locatie van de materialendoos en naar specifieke informatie over leerlingen die nog niet was opgevraagd of niet eenduidig was ontvangen. Dit kon worden aangegeven op de in de doos aanwezige leerlinglijst, met daarin per leerling ruimte om informatie te noteren over:

- uitsluiting/motorische problemen (eenduidige codering)
- ontbreken ouderlijke toestemming
- leerlinggewicht (0, 0.3 en 1.2)
- schooladvies vo/uitstroomprofiel (voor gedefinieerde categorieën)
- eventuele afwezigheid + reden:
 - Code 1: niet meer op school; Code 2: ziek
- sbo: Leesniveau leerlingen indien <AVI E3
- Leerlingen uit groep 7³⁰

Na korte afstemming met de groepsleerkracht over de afname, startte de afname. Na afloop zorgde de testleider ervoor dat de namen van zowel de leerlingboekjes als de leerlinglijsten werden geknipt, zoals afgesproken met de scholen. De testleider verzamelde alle boekjes en ingevulde vragenlijsten

²⁹ Uiteindelijk heeft de terugkoppeling plaatsgevonden op klasniveau, waardoor de sleutel niet nodig was.

³⁰ Dit om te voorkomen dat er niet traceerbare leerlingen uit groep 7 zouden deelnemen. Hoewel duidelijk aan de scholen was gecommuniceerd dat deelname enkel voor leerlingen van groep 8 was, leert de ervaring dat er in combinatieklassen nog wel eens een leerling uit groep 7 kan deelnemen. In dat geval zou deze leerling uit de dataset moeten worden gehaald.

(inclusief die van de leerkracht in groep 8, en zo mogelijk van de andere leerkrachten in de bovenbouw) en droeg zorg voor retourzending van alle materialen.

Alle afnamedagen verliepen in principe via hetzelfde patroon. Dit is weergegeven in het Handboek voor testleiders³¹.

4.7 Verdiepend onderzoek

Naast het hoofdonderzoek heeft een verdiepend onderzoek plaatsgevonden met als doel het onderwijsleerproces op de deelnemende scholen nader in kaart te brengen. Streven was om in iedere deelnemende klas een observatie van een schrijfles uit te voeren, een interview met de betreffende leerkracht uit te voeren en twee leerlingen individueel te interviewen. Om leerkrachten zo min mogelijk te belasten is leerkrachten gevraagd een gangbare schrijfles te verzorgen, zoals deze op het programma staat. In het bo is het onderzoek uitgevoerd door onderwijsinspecteurs die in de rol van onderzoeker met behulp van door de onderzoekers aangeleverd instrumentarium informatie over het schrijfonderwijs hebben verzameld. In het sbo hebben studenten orthopedagogiek de schoolbezoeken uitgevoerd, met gebruikmaking van hetzelfde instrumentarium.

Werving

Reeds tijdens de werving voor het peilingsonderzoek zijn de scholen geïnformeerd over het verdiepende onderzoek dat onderdeel vormde van dit totale peilingsonderzoek. De werving voor het kwalitatieve onderzoek is half maart gestart. Half/eind maart zijn de schooldirecteuren en contactpersonen op de scholen hierover benaderd via een mailbericht, waarin een datum voor het schoolbezoek is voorgesteld. Na deze mail wordt telefonisch contact met de contactpersonen/leerkrachten opgenomen om concrete afspraken voor de afnamedag en een dagplanning te maken. Scholen zijn vanzelfsprekend geïnformeerd over de zorgvuldigheid waarmee met de data wordt omgegaan en dat er sprake zou zijn van een geanonimiseerde rapportage. Voorwaarde voor de planning van het schoolbezoek was dat de hoofdafname op de school plaatsvond vóór de afname van het verdiepende onderzoek. De basisscholen zouden worden bezocht door inspecteurs (in de rol van onderzoekers), terwijl de sbo-scholen dienden te worden bezocht door studenten orthopedagogiek. De werving en uitvoering van het verdiepende onderzoek in het sbo is later van start gegaan, omdat er sprake was van beperkte beschikbaarheid in april vanwege problemen om geschikte, ter zake kundige kandidaten te vinden. De sbo-scholen zijn op eenzelfde manier benaderd voor het verdiepende onderzoek (via mail en telefoon). Omdat gaandeweg bekend werd dat niet alle sbo-scholen bezocht konden worden, is besloten dat bij de selectie van scholen rekening zou worden gehouden met regio, het expliciete stratificatiecriterium.

De periode voor afnames van het hoofdonderzoek (april-half juni 2019) en het kwalitatieve onderzoek (april-begin juli 2019) was grotendeels dezelfde, als gevolg van de gewijzigde planning (zie paragraaf 2.3). De genoemde volgorde in afname van hoofd- en kwalitatief onderzoek heeft daarbij de planning van de schoolbezoeken bemoeilijkt. Zeker aan het eind van de afnameperiode bleek het bij sommige scholen niet meer mogelijk om een afspraak voor de verdieping in te roosteren. In de helft van de gevallen ging het om scholen die het te druk hadden met de eindmusical, het op kamp gaan en andere afrondende activiteiten in groep 8. In de andere helft van de gevallen bleek het niet mogelijk

³¹ Dit handboek is beschikbaar bij de onderzoekers.

om voor de zomervakantie nog een match te vinden tussen de beschikbaarheid van de inspecteurs en de school. Uiteindelijk zijn 85 van de 95 reguliere basisscholen bezocht (89%), terwijl 22 van de 38 deelnemende sbo-scholen zijn bezocht (56%).

Training

De inspecteurs zijn eind maart gedurende een dag getraind in het gebruik van de instrumenten (zie paragraaf 1.6). Twee studenten orthopedagogiek namen deel aan deze training. Alle getrainden ontvingen na afloop een handleiding met daarin richtlijnen voor afname.

Afnames

In het geval er één groep 8 op een school was nam het schoolbezoek 1 dagdeel (4 uren) in beslag, waarin – bij voorkeur – eerst een schrijfles werd geobserveerd en daarna de leerkracht werd geïnterviewd. Het interviewmoment met de leerlingen vond plaats in overleg met de leerkracht. Leerlingen werden geselecteerd door de inspecteur, zonder inmenging van de leerkracht. Afspraken werden zoveel mogelijk afgestemd op de wensen van de school. In het geval van meerdere groepen 8 of groepen schoolverlaters duurde de afname langer dan het beoogde dagdeel, omdat er in die gevallen zo mogelijk in alle groepen acht/schoolverlaters moest worden geobserveerd en er extra interviews moesten worden gehouden.

In totaal zijn 85 van de 95 reguliere basisscholen bezocht. Veertien daarvan participeerden met twee groepen 8 aan de hoofdstudie. Van 70 leerkrachten zijn data van geobserveerde schrijflessen beschikbaar en bruikbaar voor analyses.³² Van 99 leerkrachten zijn interviewdata beschikbaar, op 14 scholen is ook de leerkracht van de tweede deelnemende klas geïnterviewd. Per aangemelde klas zijn in principe steeds twee leerlingen geïnterviewd. Twee vanwege de beperkte tijd die de inspecteur (per klas) op één schooldag tot zijn/haar beschikbaarheid had. In totaal zijn op de reguliere basisscholen van 199 leerlingen interviewdata beschikbaar.

In het speciaal basisonderwijs zijn in totaal 22 scholen bezocht. Van deze scholen zijn van 36 leerkrachten (van 36 groepen schoolverlaters) interviewdata beschikbaar, waarvan op 12 scholen van twee leerkrachten (twee groepen) en van één school van drie leerkrachten (drie groepen). Van 30 leerkrachten is één les geobserveerd. In totaal zijn er van 72 sbo-leerlingen interviewdata beschikbaar (net als in het bo 2 leerlingen per groep).

³² Er waren observatiedata van 79 bo-leerkrachten beschikbaar, maar door het hanteren van een exclusie criterium van maximaal 15% ontbrekende waarden zijn 9 lesobservaties uit de analyses verwijderd.

Hoofdstuk 5: Kwaliteit van de instrumenten

In hoofdstuk 1 is beschreven aan de hand van welke schrijftaken de schrijfvaardigheid van leerlingen is gemeten en aan de hand van welke beoordelingscriteria de schrijfproducten zijn beoordeeld. Voor elke schrijftaak werd hiertoe een analytisch beoordelingsmodel opgesteld, waarin beoordelaars hebben gecodeerd of specifieke kenmerken in het schrijfproduct aanwezig waren of niet. De hierdoor ontstane codes zijn middels vooraf vastgestelde scoringsregels gescoord, waarna is nagegaan of de items geschaald konden worden op een vaardigheidsschaal over de taken heen. Het proces om te komen tot de schaling verliep in een viertal stappen. Allereerst wordt beschreven hoe de observatieformulieren zijn omgezet naar scores (paragraaf 5.1). Stap 2 bevatte de analyses per taak (paragraaf 5.2). In paragraaf 5.3 wordt gerapporteerd over de manier waarop we van taak naar schaal zijn gekomen (stap 3) en in paragraaf 5.4 worden de resultaten van de uitgevoerde OPLM-analyses beschreven (stap 4). Vervolgens bevat paragraaf 5.5 informatie over de gemaakte schalen in de gebruikte school-, aanbod- en leerlingvragenlijsten.

5.1 Stap 1: Data - van observatieformulier naar scores

De door de beoordelaars ingevulde scoreformulieren zijn omgezet naar data³³. De tabbladen per taak bevatten per record de gegevens van een beoordeling van een leerling. Dubbel beoordeelde leerlingen bij een taak stonden hierdoor tweemaal in deze tabbladen. Per record was er een identificatie van een beoordelaar, van de beoordeelde leerling en waarden op de observatieformulieren. Het aantal kolommen verschilde per taak. Ook de observatie die per kolom gecodeerd is verschilde per taak. In Tabel 5.1 staat het overzicht van de taken en de aantallen ingevulde kolommen die de leerlingwerken betroffen. Welke vraag bij welke kolom per taak hoorde kon worden achterhaald via de observatieformulieren. In Tabel 5.1 is tevens weergegeven hoeveel vragen per taak als Opbouw-, Formulerings- en Inhoudsvragen (OFI-vragen) zijn geïdentificeerd. De observatieformulieren bevatten namelijk OFI-vragen en additionele vragen: voor de anker taken bevatten zij ieder zes selectie- en algemene vragen en voor de nieuwe taken acht. In totaal waren er 253 OFI-vragen.

De in het Excel-bestand weergegeven observaties dienden allereerst te worden omgezet naar scores, omdat de waarden per observatie niet oplopend indicatief waren voor een goede schrijfvaardigheid, maar vooral kwalitatief konden worden geïnterpreteerd. De inhoudsexperts en de methodologen hebben gezamenlijk een overzicht gemaakt hoe iedere observatiewaarde ("observatiemogelijkheid") bij iedere observatievraag ("item") per taak omgezet kon worden naar een score. Daarnaast is bij iedere observatievraag aangegeven welke functie in termen van opbouw, formulering en inhoud (OFI) die observatie had³⁴.

³³ Deze zijn aangeleverd als een Excel-bestand met een tabblad met het codeboek (beschrijving van de data), twaalf tabbladen met de data per taak (iedere taak een eigen tabblad), en vier tabbladen met aanvullende informatie. De aanvullende gegevens betreffen overzichten van totale aantallen, twee tabbladen ("recodes" en "herverdeeld") waarin beschreven is hoe onduidelijkheden in de ingevulde formulieren zijn aangepakt, en een tabblad met een overzicht van combinaties van taken die de leerlingen hebben gemaakt.

³⁴ Een deel van de taken is door twee beoordelaars gescoord. De toewijzing van een beoordelaar als eerste dan wel tweede beoordelaar is door middel van een random procedure gedaan. Het oordeel van de eerste beoordelaar is uiteindelijk in de schaling meegenomen.

Tabel 5.1: Overzicht van taken en observaties per taak

Taak nummer	Taak naam	Anker	Aantal vragen/kolommen	Aantal OFI-vragen
01	Wim	Anker	25	19
02	Verboden boek	Nieuw	21	13
03	Unicef	Anker	30	24
04	Brief aan Annika	Nieuw	33	25
05	Nepnieuws	Nieuw	27*	19
06	Vlinder	Nieuw	25	17
07	Pannenkoeken	Nieuw	50	42
08	Poekie	Anker	29	23
09	Naima	Nieuw	32	24
10	Speelveld	Nieuw	28	20
11	Korte zomervakantie	Nieuw	27	19
12	Frisdrank en snoep	Nieuw	16	8

* Bij Taak 5 nummert het aantal vragen tot 25, maar vraag 6 bestaat uit drie delen (a,b,c), waardoor er 2 extra vragen zijn

5.1.1 Soorten vragen

De observatielijsten van de nieuwe opgaven startten met twee “**selectievragen**” die aangaven of het te beoordelen product te beoordelen was (bij de ankeropgaven waren dat er drie). Als een van die vragen negatief beoordeeld werd stopte de verdere beoordeling van het product. In paragraaf 5.2 worden resultaten gegeven van controles op het daadwerkelijk stoppen van de beoordelingen door de beoordelaars³⁵.

Na de selectievragen volgden vragen die in drie groepen op te delen waren:

- O Opbouw (samenhang)
- F Formulering
- I Inhoud

Deze vragen worden verder **OFI-vragen** genoemd. Het aantal vragen van deze soort varieerde per taak. Ook het totaal over deze drie typen varieerde per taak. De kern van de schrijfvaardigheidsmeting zijn deze OFI-vragen. Na toepassing van de scoreregels zou de schatting van de schrijfvaardigheid gemaakt worden op basis de scores op deze vragen. De samenhang met de algemene vragen is ook onderzocht.

Iedere taak eindigde met een aantal algemene vragen. Deze vragen betreffen deels ook de “**correctheid**” van de tekst. Bij de *Ankertaken* werden de volgende vragen gebruikt:

- Geef een indicatie van het aantal schrijffouten?

³⁵ Daarnaast werd nog gevraagd of de leerling de taak anders had opgevat. Dit was geen reden om de beoordeling te stoppen; de vraag was geen selectievraag.

- Antwoordcategorieën: veel schrijffouten; niet veel, maar ook niet weinig schrijffouten; weinig of geen schrijffouten
- Geef een schatting van het aantal woorden? (getal tussen 1 en 999)
- Wat is uw algemene indruk van het taalgebruik van de leerling?
 - Antwoordcategorieën: goed; voldoende; onvoldoende

Bij de nieuwe taken betrof dit een vijftal vragen:

- Totaaloordeel: Wat vind je van de kwaliteit van de tekst?
- Hoeveel geschreven regels telt de bodytekst?
- Bij hoeveel zinnen ontbreekt de hoofdletter aan het begin of de punt aan het einde van een zin?
- Hoeveel zinnen bevatten niet-lopende constructies?
- Bij hoeveel zinnen bevat de tekst spelfouten?

Bij de eerste twee van deze vragen waren er vier codeer-categorieën. Bij algemene indruk waren de categorieën: 1) slecht, 2) onvoldoende, 3) voldoende en 4) goed. De inschatting van het aantal zinnen van de bodytekst kon worden genoteerd als 1) 1-5 regels, 2) 6-10 regels, 3) 11-15 regels en 4) meer dan 15 regels. Bij de laatste drie vragen over correctheid waren dat de categorieën: 1) meer dan de helft van de zinnen, 2) minder dan de helft van de zinnen en 3) geen enkele zin.

5.1.2 Scores op de vragen - aandachtspunten

Allereerst is nagegaan of de oorspronkelijk beoogde scores voor alle vragen ook daadwerkelijk een indicatie voor de schrijfvaardigheid weergaven. De observaties zijn dus omgezet naar scores. Voor iedere OFI-vraag gold dat de minimumscore 0 was. De maximumscore per OFI-vraag varieerde, bijvoorbeeld omdat antwoorden gesommeerd werden, maar ook omdat er soms meerdere antwoordmogelijkheden konden worden aangekruist. De vraag was of een hogere score, bij meerdere bolletjes, ook daadwerkelijk indicatief was voor een betere schrijfvaardigheid. Hierbij waren drie situaties relevant:

- 1) Situaties waarbij beoordelaars één of meerdere bolletjes bij een vraag konden aankruisen. De vraag was of meer aangekruiste bolletjes bij een item duiden op een betere schrijfvaardigheid dan een enkel bolletje.
- 2) Situaties waarbij twee verschillende observaties zouden leiden tot dezelfde score. De vraag was of het voor de schrijfvaardigheid relevant was of bij een item bolletje a of bolletje b was aangekruist.
- 3) De aanwezigheid van ontbrekende waarden. De vraag was hoe deze ontbrekende waarden te duiden: als niet aanwezig of over het hoofd gezien (en daarmee niet indicatief voor schrijfvaardigheid).

Bij de omzetting van de observaties naar scores was er een twintigtal opgaven die speciale aandacht verdienen omdat beoordelaars zowel één als twee bolletjes konden aankruisen om hun observatie weer te geven. Een voorbeeld hiervan was Taak 10 (Speelveld – verzoek), item 7, waarbij gold dat het invullen van zowel optie b als c indicatief was voor een betere schrijfvaardigheid dan het invullen van een enkele optie (b of c of d).

- | | |
|--|--|
| <p>7. Wat staat er in de inleiding van de brief?
(Maak twee rondjes zwart, indien van toepassing.)</p> | <p><input type="radio"/> Er is geen inleiding, de tekst begint direct met het verzoek</p> <p><input type="radio"/> een zin die het verzoek inleidt zonder de aanleiding te noemen</p> <p><input type="radio"/> een of meer inleidende zinnen die de situatie en/of het speelveldje beschrijven</p> <p><input type="radio"/> iets anders dan b of c</p> |
|--|--|

Bij deze twintig opgaven konden meer mogelijke observaties gedaan worden: er waren meer mogelijke scorecategorieën dan op het formulier vermeld. Nadeel hiervan was dat het aantal leerlingen waarbij een bepaalde mogelijke observatie in de tekst gevonden werd soms laag of zelfs afwezig is. Bij vier vragen (T06_v04, T06_v17, T07_v44 en T12_v10) waren er scorecategorieën met minder dan 10 leerlingen. Hierbij zijn we overgegaan tot samenvoeging van categorieën. Zo werd bij de eerste twee opgaven (T06_v04, T06_v17) zeer zelden de beoogde hoogste scorecategorie gevonden in de teksten. Deze score-categorie is samengevoegd met de beoogde een-na-hoogste scorecategorie, die nu feitelijk de hoogste score bij deze vragen werd. Bij de andere twee items (T07_v44 en T12_v10) was de scorecategorie ‘precies een onder de maximum score’ zelden geobserveerd. Ook hier zijn de hoogste en de een-na-hoogste score categorie samengevoegd, waardoor ook hier de beoogde een-na-hoogste scorecategorie uiteindelijk de hoogste score categorie werd.

Bij twee nieuwe items moest de scoring nogmaals nagegaan worden, omdat het veelvuldig (ieder meer dan 150 observaties) voorkwam dat twee mogelijke observaties dezelfde score op zouden leveren. In dat geval zou er geen onderscheid meer kunnen worden gevonden tussen deze twee veel voorkomende observaties, hetgeen alleen verantwoord is als de twee verschillende observaties ook werkelijk met de vaardigheid samenhangen. Aangezien de vaardigheid nog niet bekend was, zijn als proxy de vijf eerdergenoemde algemene vragen gebruikt, waarbij met name de eerste relevant was (algemene indruk van het leerlingwerk). In één geval leverden de verschillende observaties geen verschillen op die variabelen op. Hier is de origineel beoogde omzetting van observatie naar score toegepast. Bij het andere item was dat wel het geval: de twee observaties –met ieder meer dan 250 observaties– leverden significante verschillen op wat betreft algemene tekstkwaliteit. Om die reden hebben de twee observaties verschillende scores gekregen (de observatie met het hogere gemiddelde op de algemene vraag kreeg de hogere score, de andere de lager score). Na bepaling van vaardigheid met de OFI-items is onderzocht of dit onderscheid nog steeds gerechtvaardigd was. Dit was het geval bij item T06_v20, waarbij twee beoogde verschillende scores (de twee hoogste scores) ieder zeer weinig observaties hadden. Deze twee scorecategorieën zijn alsnog samengevoegd.

5.1.3 Ontbrekende waarden

Een laatste punt van aandacht bij de omzetting van beoordelingen naar scores betrof het scoren van de ontbrekende waarden bij de OFI-vragen. Bij meerkeuze-opgaven is het gebruikelijk om een waarde 0 te imputeren voor ontbrekende waarden: de kandidaat heeft het goede antwoord immers niet gegeven. In het geval van beoordelingsformulieren is dat niet noodzakelijk het geval, omdat het ook een fout van een beoordelaar kan betreffen. Het beoordelaarsformulier was echter dusdanig ingericht dat het wel het meest aannemelijk was dat een ontbrekende waarde gelijk stond aan het niet vinden van een bepaald kenmerk in een tekst. Van tevoren is achterhaald hoe vaak ontbrekende waarden voorkwamen en of dat bij sommige items veel vaker het geval was dan bij andere. Het kon zijn dat er

ontbrekende waarden waren omdat een tekst niet voldeed, zoals aangegeven met de eerste twee (of drie, bij ankers) selectievragen. Bij de evaluatie per opgave is daar rekening mee gehouden.

In totaal waren er 253 OFI-vragen. Bij 41 van de 43 opgaven waarbij het percentage ontbrekende waarden boven de 2% uitkwam³⁶ zijn de hoge percentages – tot en met 78% ontbrekend – te wijten aan afhankelijkheden in de vraagstelling. Deze vragen moesten overgeslagen worden wanneer een bepaald antwoord gegeven was op een andere vraag die ervoor gesteld werd. Te denken valt hierbij aan een voorbeeld als “Is er een titel?” waarna de volgende vragen, die over de titel gaan, overgeslagen kunnen worden als er geen titel is. Hierover volgt later meer. Daarnaast waren er nog twee andere vragen met relatief veel ontbrekende waarden (T07_v11 en T07_v38). Hier was de ontbrekende waarde een eigen inhoudelijke betekenis en (daarmee) ook een eigen scorecategorie, zoals blijkt uit het voorbeeld hieronder (T07_v11):

11. Gaan de ingrediënten vooraf aan de bereidingsstappen? (alleen beantwoorden als de tekst zowel ingrediënten als stappen bevat)	<input type="radio"/> Ja	<input type="radio"/> Nee
---	--------------------------	---------------------------

Bij de overige 210 OFI-vragen kwam het percentage ontbrekende waarden nooit boven 2,5% uit. Van deze 210 OFI-vragen kwam het percentage bij 206 vragen niet boven de 1,5% uit en bij 160 vragen niet boven een half procent. Geconcludeerd kan worden dat de impact van de ontbrekende waarden klein is. Ontbrekende waarden zijn uiteindelijk opgevat als niet aanwezig (nee). Bij de algemene vragen worden ontbrekende waarden niet geïnterpreteerd en blijven ze een aparte categorie. Deze opgaven zijn geen onderdeel van de vaardigheidsschaal.

5.2 Stap 2.1: Analyses per taak – een analyse van de selectievragen

De eenheden binnen de metingen van de schrijfpeiling zijn de schrijftaken en de daaruit volgende schrijfproducten. Hoewel we over taken heen willen meten, was het verstandig eerst de resultaten per taak te bekijken om opmerkelijke zaken te verwerken alvorens de resultaten van de verschillende taken samen te voegen en na te gaan of er een of meer onderliggende vaardigheidsschalen zou(den) bestaan.

5.2.1 Selectievragen en ontbrekende waarden – nieuwe taken

Een eerste punt van aandacht was of een schrijfproduct meegenomen moest worden in de analyses. Dit hing ten eerste af van de selectievragen. Hiervoor moest bij de verwerking een onderscheid worden gemaakt tussen de aanpak bij de nieuwe taken en de ankertaken. Bij de nieuwe taken was expliciet opgenomen dat de beoordeling zou stoppen als op een van de twee selectievragen een “nee” geantwoord zou worden; de rest van de observaties zou dan conform verwachting ontbreken. Als een selectievraag was overgeslagen (ontbrekende waarde), is er in eerste instantie van uitgegaan dat dit vergelijkbaar was met een positief antwoord (namelijk dat de tekst voldoende leesbaar danwel voldoende bodytekst bevatte). De eerste vraag luidde of de tekst leesbaar is. Was dat niet het geval dan stopte de beoordeling en ontbrak ook een antwoord op de tweede vraag. Dat is in

³⁶ Alle genoemde percentages zijn gecorrigeerd voor de kandidaten waarbij alle opgaven niet gescoord zijn.

alle gevallen ook toegepast door de beoordelaars. In één geval ontbrak een antwoord op deze eerste vraag. Bij vraag twee is gevraagd naar de bodytekst. Als er geen bodytekst was dan hoefde er verder niet beoordeeld te worden.

In Tabel 5.2 wordt het overzicht gegeven van de antwoordcombinaties op de twee selectievragen bij verschillende beoordelingen van de nieuwe taken. Het gaat daarbij om beoordelingen, niet om leerlingen. De kopregel maakt het onderscheid in het aantal beoordelingen per taak dat al dan niet geselecteerd is. De twee kolommen onder de kopregel 'nee' geven de beoordelingen weer die niet geselecteerd zijn.

Tabel 5.2: Antwoordcombinatie op de twee selectievragen bij de verschillende beoordelingen

Geselecteerd	Naam taak	Nee			Ja	
		NEE	JA	JA	JA	JA
Leesbaar		NEE	JA	JA	JA	JA
Body tekst		OW*	NEE	OW*	SOMS	WEL
Taak 02	Verboden Boek	26	0	5 ³⁷	15	948
Taak 04	Brief aan Annika	21	22	3	86	875
Taak 05	Nepnieuws	10	6 ³⁸	5	31	1096
Taak 06	Vlinder	6	8	1	27	940
Taak 07	Pannenkoek	16	19 ³⁹	5	34	1082
Taak 09	Naima	12	1	1	4	1009
Taak 10	Speelveld	10	3 ⁴⁰	0	11	1024
Taak 11	Korte zomervakantie	10	0	5	6	994
Taak 12	Frisdrank en snoep	4	0	4	7	1019

* OW = ontbrekende waarde bij body tekst

Er is ook gekeken naar het percentage gegeven antwoorden. Bij de niet-geselecteerde beoordelingen lag het percentage gegeven antwoorden altijd onder de 10%. Bij de wel-geselecteerde beoordelingen ontbreekt minder dan 10%, lopende van 3% (Pannenkoek) tot 6.3% (Vlinder en Frisdrank en Snoep).

5.2.2 Selectievragen en ontbrekende waarden - ankertaken

Ook bij de ankertaken is eerst nagegaan of alle schrijfproducten konden worden meegenomen in de schaling. Er waren drie redenen om schrijfproducten niet op te nemen: 1. een negatief antwoord bij selectievragen (met name vraag 1), 2. een hoog percentage ontbrekende waarden in de beoordeling en 3. meerdere beoordelingen van dezelfde beoordelaar bij dezelfde taak.

³⁷ Van de 5 beoordelingen was één leesbaar geacht omdat hier niets ingevuld was, het bleek dat het beoordelingsformulier geheel niet ingevuld was. Die beoordeling is alsnog niet geselecteerd.

³⁸ Bij 1 van de deze 6 beoordelingen waren toch scores ingevuld – de beoordeling is geselecteerd

³⁹ Bij 2 van de deze 19 beoordelingen waren toch scores ingevuld – de beoordelingen zijn geselecteerd

⁴⁰ Bij 2 van de deze 3 beoordelingen waren toch scores ingevuld – de beoordelingen zijn geselecteerd

Het was relevant om het percentage ontbrekende waarden nader te bestuderen, omdat er bij de ankertaken niet expliciet gevraagd wordt om te stoppen met de beoordelingen. Bij deze ankertaken is er een drietal vragen dat als selectievraag opgevat kan worden: (1) *Is de tekst leesbaar?*; (2) *Is de tekst af?* (3) *Heeft de leerling zich aan de opdracht gehouden?* Van deze vragen bleek de eerste vraag als een knock-out criterium te functioneren. Wanneer die vraag negatief beantwoord werd of wanneer een beoordeling op de eerste vraag ontbrak, dan is het percentage antwoorden in het geheel van de taak bij alle drie de ankertaken altijd onder de 10%. Deze teksten waren ook nooit beoordeeld als "af" of "voldoet aan opdracht". Deze beoordelingen zijn niet meegenomen in de verdere analyses.

Er bleken 14 beoordeelde werken met relatief veel ontbrekende waarden te zijn. Bij deze 14 werken ontbrak meer dan 60%⁴¹. Deze zijn verwijderd omdat die werken niet bruikbaar zijn⁴².

Een laatste reden om beoordeelde werken uit te sluiten van verdere analyses was het aantal beoordelingen. Er was een drietal werken dat drie beoordelingen had: in twee gevallen (in Taak 01 en Taak 02) had dezelfde beoordelaar tweemaal dezelfde kandidaat beoordeeld. Deze beoordelingen waren vrijwel duplicaten van elkaar, waardoor er een verwijderd is. In het derde geval (Taak 07) waren de beoordelingen van dezelfde kandidaat door de beoordelaars dusdanig verschillend dat dit een administratieve fout lijkt: aangezien niet te achterhalen was wat de correcte beoordeling was zijn beide beoordelingen verwijderd.

5.2.3 Definitief verwijderde producten (over alle taken heen)

In Tabel 5.3 is het overzicht gegeven van de definitief uit de analyses verwijderde producten. Uit Taak 04 zijn relatief veel beoordelingen verwijderd. Deze taak had met name bij de vraag aangaande de bodytekst veel uitval. Dat geldt ook voor Taak 07, maar daar waren in totaal meer producten beoordeeld. Taak 02 had relatief veel onleesbare producten. Bij de overige taken is er een uitval van 2% of minder.

Het onderscheid tussen de eerste en de tweede beoordelaar is gemaakt bij producten die door twee beoordelaars beoordeeld waren. De toewijzing van een beoordelaar als eerste dan wel tweede beoordelaar is door middel van een random procedure gedaan.⁴³ Meer over de beoordelaars volgt in de beschrijving van het beoordelaarsonderzoek.

⁴¹ Bij 12 van deze werken ontbrak meer dan 80% en bij 8 van de producten 90% of meer.

⁴² Het betreft één beoordeling bij taak 01, zeven beoordelingen bij taak 03 en zes bij taak 08.

⁴³ Iedere product heeft naast een unieke codering ook een randomgetal gekregen. Als er twee beoordelaars waren, is de beoordelaar met het laagste random getal de eerste beoordelaar. Als de eerste en tweede beoordelaar vergeleken is met de random ID dan is de verdeling dat het laagste beoordelaars-ID dan wel het hoogste beoordelaars ID de eerste beoordelaar is, ongeveer 50%-50%.

Tabel 5.3: Overzicht van de van de producten en resultaat van selectie per taak en beoordelaar

Selectie	Naam taak	Start	In: geselecteerd			Uit: niet in schaal-analyses			
			Aantal		%	Aantal			%
Definitief			1e Beo	2e Beo	2 Beo's/in	1e Beo	2e Beo	3e Beo	Uit/Totaal
Taak 01	Wim	1172	1028	135	11.6%	8	0	1	0.8%
Taak 02	Verboden boek	994	854	112	11.6%	26	1	1	2.8%
Taak 03	Unicef	1195	1043	128	10.9%	23	1	0	2.0%
Taak 04	Brief aan Annika	1007	838	126	13.1%	42	1	0	4.3%
Taak 05	Nepnieuws	1148	1005	128	11.3%	12	3	0	1.3%
Taak 06	Vlinder	982	858	110	11.4%	14	0	0	1.4%
Taak 07	Pannenkoek	1156	992	129	11.5%	31	2	2	3.0%
Taak 08	Poekie	964	840	106	11.2%	17	1	0	1.9%
Taak 09	Naima	1027	901	113	11.1%	13	0	0	1.3%
Taak 10	Speelveld	1048	921	116	11.2%	11	0	0	1.0%
Taak 11	Korte zomervakantie	1015	894	111	11.0%	9	1	0	1.0%
Taak 12	Frisdrank en snoep	1034	909	121	11.7%	4	0	0	0.4%
Totaal		12742	11083	1435	11.5%	210	10	4	1.8%⁴⁴

5.2.4 Nieuwe taken: selectievraag 3

Een laatste vooronderzoek betrof de derde vraag van de nieuwe taken. Deze vraag is enigszins vergelijkbaar met de derde vraag bij de ankertaken: "Heeft de leerling de opdracht anders opgevat dan bedoeld?". Deze derde vraag was in tegenstelling tot de eerste twee vragen bij de nieuwe taken geen exclusie criterium, net zomin als dat bij de ankertaken het geval was. De derde vraag is echter ook geen onderdeel van de schaal.

⁴⁴ In dit hoofdstuk presenteren wij de resultaten die we per stap hebben gevonden. Zoals in paragraaf 5.4 blijkt, bleken acht dubbele beoordelingen geen daadwerkelijk dubbele beoordelingen te zijn. Wanneer we rekening houden met dit gegeven, dan zien we dat er 1427 tweede beoordelingen geselecteerd zijn (11.2%) en 17 dubbele oordelen niet (1.8%).

Tabel 5.4 geeft per taak het percentage beoordelingen, waarbij de taak anders is opgevat dan bedoeld. Dit percentage loopt bij de nieuwe taken van 1.7% (bij Taak 6) tot en met 5.5% (bij Taak 09), met een gemiddelde van 3.3%. Bij de ankertaken zijn percentages bij de vergelijkbare vraag iets hoger met een gemiddelde van 4.2%, maar veel scheelt het niet.

Tabel 5.4: Percentage beoordelingen per taak waar het product anders is opgevat door de leerling

Nieuwe taken						Anker taken							
Taak	% anders			Taak	% anders			Taak	% anders			Taak	% anders
T02	4.7%			T06	1.7%			T10	2.3%			T01	4.0%
T04	2.6%			T07	5.1%			T11	1.9%			T03	2.6%
T05	2.3%			T09	5.5%			T12	4.0%			T08	5.9%

Nadat de scores verkregen waren, is deze vraag naar een andere taakopvatting nog eens bekeken om na te gaan of het niet begrijpen van de taak impact had op prestatie. Daartoe is gekeken naar verschillen in scores tussen de leerlingen waarvan de beoordelaar stelde dat de leerling de opdracht wel/niet anders heeft opgevat dan bedoeld. Er bleken wel enige verschillen te zijn tussen de scores van leerlingen die de taak wel en leerlingen die de taak niet goed hadden begrepen. De eerste groep scoorde over het algemeen hoger op de taak. De vraag is hoe dit verschil moet worden geïnterpreteerd. De lagere score kon komen door de specifieke opdracht die niet goed begrepen is, of omdat deze leerlingen in het algemeen niet heel erg vaardig zijn. In het laatste geval, zou het niet meenemen in de analyses van de leerlingen die de taak niet hebben begrepen tot gevolg hebben dat dat de vaardigheid hoger lijkt dan deze in werkelijkheid is. Om die reden zijn de verschillen ook bekeken op de gehele vaardigheidsscore in plaats van alleen op de score op de taak (zie Bijlage 10). In de meeste gevallen bleken de effecten niet te veranderen wanneer ook de resultaten op de andere taken werden meegenomen. Het lijkt er daarom op dat, over het algemeen, de leerlingen die een taak niet begrepen hadden, eerder de in het algemeen wat minder vaardige leerlingen waren dan dat ze leerlingen waren die zeer veel lager op een specifieke taak scoorden. Producten die zich kenmerken doordat de leerling de taak niet goed heeft begrepen, zijn daarom niet verwijderd uit de verdere analyses. Leerlingen die de taak verkeerd begrepen hebben konden namelijk, ondanks het niet goed begrijpen van een taak, alsnog beoordeeld worden. Reden hiervoor is dat er een diversiteit aan beoordelingsaspecten werd bevraagd, waardoor een leerling alsnog een beoordeling kon krijgen. Ter illustratie: een deel van de leerlingen heeft Taak 9 (waarin een voormalig klasgenoot om tips diende te worden gevraagd voor het houden van een interview) verkeerd geïnterpreteerd, namelijk dusdanig dat zij de klasgenoot vroegen of zij haar mochten interviewen. Hoewel een deel van de vragen in het beoordelingsformulier niet correct kon worden beoordeeld, gold dat zeker niet voor alle vragen.

5.5 Stap 2.2: Analyses per taak: van OFI-itemscores naar taakscores

Bij het geplande beoordelaarsonderzoek wilden we gebruik maken van taakscores. Daarom zijn eerst de analyses op de OFI-itemscores uitgevoerd. Per taak konden de OFI-scores op verschillende

manieren samen een score vormen. Het gemakkelijkst was de totaalscore waarbij de som van alle OFI-scores opgeteld werden (T##_som_OFI). Vervolgens kon per onderdeel een som genomen worden (T##_som_O, T##_som_F, en TK##_som_I). Van ieder van deze schalen is de betrouwbaarheid op basis van inter-item-relaties berekend (Guttman's lambda-2, of Cronbach's alpha), zoals weergegeven in Tabel 5.5⁴⁵. Daarnaast bevat Tabel 5.6 de kolom OFI*, waarin de betrouwbaarheid wordt weergegeven op het moment dat de onderling direct afhankelijke variabelen niet als afzonderlijke variabelen worden gezien. Onderlinge directe afhankelijkheid van variabelen komt in acht van de 12 taken voor. Een voorbeeld van deze direct afhankelijke variabelen is bijvoorbeeld de vraag naar de aanwezigheid van een aanhef van een brief en de geschikte toon daarvan in Taak 11, zoals in T11_v5 en T11_v6:

5.	Bevat de brief een aanhef?	<input type="radio"/> Ja	<input type="radio"/> Nee -> ga verder naar 7
6.	Is de aanhef formeel van toon, passend voor een hoger geplaatste persoon?	<input type="radio"/> Ja	<input type="radio"/> Nee

Zoals uit Tabel 5.5 kan worden opgemaakt heeft het rekening houden met de direct afhankelijke variabelen tot gevolg dat het aantal OFI en OFI* items niet gelijk is voor acht taken. Het totaal aantal OFI* items over alle taken is met 211 lager dan de OFI items over alle taken (253). In Tabel 5.5 zijn alleen de waarden gegeven van Guttman's lambda-2, aangezien die een betere benadering van de betrouwbaarheid geven dan Cronbach's alpha. De conclusies op basis van de waarden van Cronbach's alpha zijn echter vergelijkbaar. De betrouwbaarheid op subschaalniveau per taak is aan de lage kant: meestal onder de 0,60, met een gemiddelde en een mediaan van 0,50 of lager. De schalen zijn ook erg kort: in twee gevallen met slechts een item ook te kort om een betrouwbaarheid te berekenen.

De betrouwbaarheid van de te schalen opgaven samen is meestal wel groter dan 0,60. Worden de onderling direct afhankelijke variabelen niet als afzonderlijke variabelen gezien dan dalen de waarden van lambda-2 van de acht taken waarin zij voorkomen (zie kolom OFI*). Bij de schalen is ook duidelijk dat de minste opgaven gemaakt zijn voor F (formulering) en de meeste voor I (inhoud). Het aantal opgaven voor O (opbouw) ligt dicht tegen dat van I aan dan dat van F.

Er zijn meer analyses uitgevoerd. Op item-niveau zijn per taak de item-rest-correlaties (Rir) berekend, zowel op het niveau van de taak als op het niveau van de subschaal. Een vijftal opgaven had een negatieve waarde⁴⁶. Additionele analyses leverden geen aanleiding voor aanpassing in de scoring⁴⁷.

Ter controle van de codering zijn tevens analyses op de OFI-schaal en subschalen gedaan. Drie opgaven hadden een negatieve Rir-waarde op de OFI-schaal en acht opgaven hadden een negatieve Rir-waarde op subschaalniveau, maar zijn om inhoudelijke redenen niet verwijderd.

⁴⁵ Bij het bepalen van die betrouwbaarheid is geen rekening gehouden met de afhankelijke items.

⁴⁶ Het betreft: T04_i04_O, T07_i42_F, T07_i45_O, T11_i19_O, en T12_i04_O.

⁴⁷ Naar aanleiding van het onderzoek per vraag kwam opgave 12 in taak 2 naar voren als bijzonder omdat dit de enige OFI-vraag binnen alle taken was waarbij een ontbrekende waarde niet gescoord was (werd als ontbrekend gerekend), omdat de betekenis van een ontbrekende waarde inhoudelijk niet evident is. Deze was bij een leerling die binnen de selectie viel (en dus niet zeer veel ontbrekende waarden had) nog steeds ontbrekend. Om nu alle kandidaten een OFI-score te geven binnen deze taak is de waarde voor deze leerling geïmputeerd. De gemiddelde score op die opgave bij de leerlingen die een gelijke restscore hebben is 1,76. De (maximum) score 2 is de meest waarschijnlijke score op dit item voor deze leerling. Deze score wordt ook door 77% van de leerlingen met een gelijke restscore gehaald, wat inhoudelijk vanuit de vraag ook begrijpelijk is. De ontbrekende waarde is nu als 2 geïmputeerd. De gegeven waarden in Tabel 5.5 zijn onveranderd of deze imputatie wordt toegepast of niet.

Tabel 5.5: Betrouwbaarheid van de taken (totaal OFI en per subschaal O, F en I)

	Betrouwbaarheid (lambda-2)					Aantal items				
	O	F	I	OFI	OFI*	O	F	I	OFI	OFI*
T01	0,50	0,68	0,36	0,68	0,68	7	7	5	19	18
T02	0,29	0,55	0,19	0,60	@	5	4	4	13	@
T03	0,29	0,40	0,54	0,62	0,57	7	5	12	24	22
T04	0,70	0,76	0,61	0,83	0,79	14	5	6	25	19
T05	0,68	0,00	0,62	0,69	@	7	2	10	19	@
T06	0,12	#	0,67	0,62	@	5	1	11	17	@
T07	0,66	0,24	0,85	0,84	@	9	6	27	42	@
T08	0,42	0,53	0,65	0,69	0,62	5	3	15	23	21
T09	0,47	0,42	0,47	0,73	0,51	11	7	6	24	12
T10	0,54	0,28	0,33	0,66	0,60	9	3	8	20	13
T11	0,40	0,42	0,27	0,62	0,56	10	3	6	19	10
T12	0,23	#	0,38	0,53	0,18	3	1	4	8	5
	Betrouwbaarheid (lambda-2)					Aantal items				
	O	F	I	OFI	OFI*	O	F	I	OFI	OFI*
gemiddelde	0,44	0,43	0,50	0,68	0,60	7,7	3,9	9,5	21,1	17,6
minimum	0,12	0,00	0,19	0,53	0,18	3	1	4	8	5
mediaan	0,44	0,42	0,50	0,67	0,61	7	3,5	7	19,5	17,5
maximum	0,70	0,76	0,85	0,84	0,84	14	7	27	42	42

Geen inter-item-relatiebetrouwbaarheid te berekenen bij 1 item; @ Geen som items (OFI* gelijk aan OFI)

5.4 Stap 2.3: Analyses per taak: Beoordelaarsbetrouwbaarheid

De volgende stap was het beoordelaarsonderzoek, waarbij de beoordelaarsbetrouwbaarheid is nagegaan. Zoals gesteld is per taak meer dan 10% van de werken dubbel beoordeeld, waarbij de toewijzing van eerste en tweede beoordelaar willekeurig plaatsvond. De overeenstemming tussen beoordelaars wordt hier gegeven per taak en per schaal per taak. De opgaven aangaande de algemene indruk, hoeveelheid geschreven tekst en de drie correctheidsopgaven worden niet geschaald, maar de interbeoordelaarsbetrouwbaarheid op itemniveau is wel bepaald.

In totaal zijn er van 1445 producten dubbele beoordelingen. Per taak zijn er 214 (Taak 8; 107 producten) tot en met 270 (Taak 1; 135 producten) dubbele beoordelingen, met een gemiddelde van 241 dubbele beoordelingen. Van alle mogelijke 276 combinaties van 24 beoordelaars kwamen er 228 voor bij de dubbele beoordelingen. Dit betekent dat 83% van de mogelijke combinaties gebruikt is. Deze 228 combinaties van verschillende beoordelaars beoordeelden gemiddeld 6 producten. Er is geen relatie tussen het volgnummer van de beoordelaar en of ze hier als eerste of tweede beoordelaar functioneerden. Opvallend genoeg zijn er 58 producten twee keer door dezelfde beoordelaar geëvalueerd.⁴⁸ Dit is bij 18 van de 24 verschillende beoordelaars gebeurd (gemiddeld 3 producten bij deze 18 beoordelaars). Volgens het beoordelaarschema zouden beoordelaars niet twee maal dezelfde leerling moeten beoordelen. Deze groep is als aparte groep verder geëvalueerd.

Evaluatie selectievragen

Als eerste is onderzocht hoe de exclusie van producten mogelijk verschilde per beoordelaarcombinatie. Bij 1430 van de 1445 producten (99%) waren de beoordelaars het met elkaar eens en in 15 gevallen niet. In zes van deze gevallen bleek een beoordelaar twee keer te zijn begonnen aan de beoordeling van dezelfde tekst van dezelfde leerling en bleek deze beoordelaar bij één van deze beoordelingen eerder te zijn gestopt. Dit zijn eigenlijk gewoon enkel beoordeelde producten. Van de 15 afwijkende beoordelaarscombinaties bleek er bij de overige negen beoordelingen daadwerkelijk een verschil van inzicht te zijn over de selectievragen, met een uiteindelijk percentage van onenigheid van iets meer dan een half procent. Bij deze negen producten was het zes keer het geval dat één van de beoordelaars vond dat het product onleesbaar was, terwijl de andere beoordelaar het product wel leesbaar vond. Deze verschillen leveren geen probleem op: op voorhand was namelijk afgesproken dat een product dat door een beoordelaar als onleesbaar werd beschouwd, aan een andere beoordelaar werd voorgelegd voor herbeoordeling. Van de 3 overgebleven verschillen zijn er twee veroorzaakt doordat één van de twee beoordelaars dacht te moeten stoppen bij vraag 3 als hij of zij het oordeel “niet af” gegeven had. Deze twee dubbele beoordelingen zijn derhalve ook geen dubbele beoordelingen, omdat één van beide beoordelaars geen volledige beoordeling heeft gegeven. In feite was er dus maar één daadwerkelijk afwijkende dubbele beoordeling over. In dit geval is er willekeurig één van beide beoordelingen geselecteerd.

Evaluatie Opbouw, Formulering en Inhoud

In totaal zijn er 11.083 producten die niet uit de analyses verwijderd zijn met een eerste beoordelaar, waarvan er bij 1427 producten sprake was van een dubbel oordeel (met een eerste en een daadwerkelijke tweede beoordelaar).⁴⁹ Dit heeft tot gevolg dat bijna 12.9% van de producten dubbel is beoordeeld, variërend van 12% tot en met bijna 15% per taak. Bij iedere taak zijn 106 tot 136 producten dubbel beoordeeld. Als we producten die twee keer door dezelfde beoordelaar beoordeeld

⁴⁸ Zoals in de navolgende tekst zal worden beschreven bleken er zes beoordelaars te zijn, die twee keer aan de beoordeling van hetzelfde schrijfproduct waren begonnen maar in één van beide keren vroegtijdig stopte met de beoordeling. Het aantal daadwerkelijk dubbel door dezelfde beoordelaar beoordeelde producten is daarmee dus 52 (58-6, zie de tekst hierboven).

⁴⁹ Zoals beschreven bij de evaluatie van selectievragen bleek er in 8 gevallen geen sprake van een daadwerkelijke tweede beoordeling: in 6 gevallen had de eerste beoordelaar twee keer dezelfde tekst beoordeeld, waarbij hij één van beide keren eerder was gestopt, en in twee gevallen was er sprake van zowel een volledige als onvolledige beoordeling. Dit verklaart het verschil in het hier beschreven totaal aantal ‘daadwerkelijke dubbele beoordelingen’ (n=1427) en het aantal in Tabel 5.3 gepresenteerde dubbele beoordelingen (n=1435). Hierdoor verschillen de genoemde percentages ook enigszins.

zijn verwijderen (dit waren er 52), dan zijn er 1375 producten dubbel beoordeeld (12.4%). Hierbij heeft iedere taak nog steeds minstens 103 dubbele beoordelingen.

De analyses voor de beoordelaarsovereenstemming zijn vooral per taak gedaan, omdat de scoranges ook per taak verschillen. Om enige analyses over taken heen te kunnen doen is voor de totaalscore ook een aparte variabele aangemaakt: het percentage van de maximale score dat de kandidaat heeft gehaald. Als de beoordelaarsovereenstemming van deze variabele per taak geanalyseerd wordt, is deze gelijk aan die op de somscore van die taak.

Als we alle 1427 werkelijk dubbel beoordeelde producten bij elkaar nemen, en de percentages van de maximumscore die een kandidaat gescoord heeft bij de eerste en de tweede beoordelaar berekenen, dan is de correlatie tussen die twee 0.74. Als we alleen de producten bekijken die door twee verschillende beoordelaars beoordeeld zijn, dan is die 0.75. Dat is dus een hogere waarde. De correlaties tussen de 52 scores die twee maal door dezelfde beoordelaar gegeven zijn, is met 0.58 dan ook lager⁵⁰.

De rest van de analyses op de oordelen per taak heeft alleen plaatsgevonden op basis van dubbele beoordelingen van verschillende beoordelaars. De evaluatie per taak laat de volgende resultaten zien voor de gehele OFI score, en voor de scores van de drie afzonderlijke schalen. Die zijn gegeven in Tabel 5.6. Merk op dat de correlatie bij de OFI-scores per taak gelijk is aan die van de "percentage van de maximum score" correlaties per taak.

Tabel 5.6: Interbeoordelaarsbetrouwbaarheid per taak voor de OFI-score en de subschalen

	OFI	O	F	I
Taak 01	0.68	0.71	0.54	0.53
Taak 02	0.53	0.17	0.51	0.62
Taak 03	0.67	0.50	0.50	0.69
Taak 04	0.90	0.90	0.91	0.58
Taak 05	0.78	0.76	0.04	0.80
Taak 06	0.75	0.61	0.63	0.69
Taak 07	0.85	0.83	0.41	0.85
Taak 08	0.72	0.66	0.51	0.69
Taak 09	0.59	0.55	0.51	0.53
Taak 10	0.70	0.67	0.42	0.63
Taak 11	0.56	0.38	0.86	0.39
Taak 12	0.39	0.52	0.38	0.24
Gemiddelde	0.68	0.60	0.52	0.60
minimum	0.39	0.17	0.04	0.24
mediaan	0.69	0.63	0.51	0.62
maximum	0.90	0.90	0.91	0.85

⁵⁰ Mogelijk was er sprake van een administratieve fout, waardoor ogenschijnlijk dubbele beoordelingen door eenzelfde beoordelaar feitelijk de beoordeling van een andere leerling betrof. Dit is achteraf niet meer na te gaan.

Bij de evaluatie van de beoordelaarsbetrouwbaarheid per taak is te zien dat de betrouwbaarheid regelmatig onder de 0.60⁵¹ valt. Zoals te verwachten is dit vooral bij de subschalen het geval. Bij F is dat het meest het geval, hetgeen ook sterk kan samenhangen met het aantal opgaven dat bij F het laagst ligt. Ook bij de OFI-score is te zien dat de taken met relatief de meeste opgaven de hoogste betrouwbaarheid hebben. Het is belangrijk om bij de interpretatie van de resultaten in Tabel 5.6 te bedenken dat de metingen bij leerlingen niet op basis van een enkele taak plaatsvinden maar op basis van vier dan wel twee verschillende taken. Als we kijken naar de gemiddelde betrouwbaarheid en met een toetsverlenging van twee danwel vier leerlingwerken (het aantal afgenomen taken in respectievelijk sbo en bo), dan zien we dat bij de varianten met vier taken alle betrouwbaarheden boven de 0.80 uitkomen, en bij twee taken minimaal 0.68 zijn (bij F). Ook in het geval van combinaties van de minst betrouwbare taken zijn bij vier taken alle betrouwbaarheden nog steeds boven de 0.60. Geconcludeerd kan daarom worden dat de betrouwbaarheid van de in de peiling gebruikte taken redelijk tot goed is, zeker wanneer het gaat om de afgenomen combinaties van taken.

Evaluatie algemene (slot)vragen

Tot slot is de beoordelaarsbetrouwbaarheid bij de algemene en correctheidsvragen over de tekst als geheel nagegaan. Deze vragen werden als laatste door de beoordelaars ingevuld. Iedere vraag is op zich een eigen score. De vragen zijn niet onderling samengevoegd per taak. Bij deze evaluatie is wel een onderscheid gemaakt tussen de nieuwe taken en de ankertaken, aangezien de vragen per type verschilden. Bij de nieuwe taken was er een vraag naar de algemene indruk, een vraag naar de hoeveelheid zinnen in de bodytekst en drie vragen betreffende de correctheid. Bij de ankertaken zijn de eerste twee vragen ook gesteld, zij het in een ander format. In de tabellen 5.7a en 5.7b zijn de gegevens gepresenteerd als er geen onderscheid gemaakt is tussen de taken. Dat was in principe mogelijk omdat de vragen hetzelfde zijn bij alle taken, conditioneel op de uitsplitsing anker- en nieuwe taken.

Tabel 5.7a: Beoordelaarsbetrouwbaarheid bij de additionele vragen bij de Ankertaken

ANKER TAKEN	schrijffouten	N woorden	Indruk	Aantal leerlingen
ALL	0.46	0.93	0.34	352
Taak 01	0.58	0.92	0.43	129
Taak 03	0.39	0.86	0.33	120
Taak 08	0.36	0.72	0.28	103
Gemiddelde	0.44	0.83	0.35	

⁵¹ De grens van 0.60 is de grens voor een voldoende betrouwbaarheid voor metingen op groepsniveau zoals aangegeven door de COTAN.

Tabel 5.7b: Beoordelaarsbetrouwbaarheid bij de additionele vragen bij de nieuwe taken

NIEUWE TAKEN	Naam taak	Algemeen		Correctheid			N lln
		Indruk	Hoeveelheid	Interpunctie	Constructie	Spelling	
ALL		0.46	0.89	0.73	0.36	0.62	1023
Taak 02	Verboden Boek	0.28	0.90	0.63	0.30	0.51	104
Taak 04	Brief aan Annika	0.62	0.88	0.76	0.35	0.59	117
Taak 05	Nepnieuws	0.40	0.83	0.73	0.36	0.62	121
Taak 06	Vlinder	0.53	0.91	0.85	0.25	0.68	107
Taak 07	Pannenkoek	0.46	0.69	0.78	0.27	0.63	124
Taak 09	Naima	0.48	0.91	0.67	0.45	0.59	111
Taak 10	Speelveld	0.50	0.96	0.71	0.47	0.60	112
Taak 11	Korte zomervakantie	0.45	0.93	0.67	0.46	0.62	110
Taak 12	Frisdrank en snoep	0.25	0.80	0.69	0.20	0.64	117
Gemiddelde		0.44	0.87	0.72	0.34	0.61	

De beoordelaarsbetrouwbaarheid is hoog bij de vragen die gaan over de hoeveelheid tekst. Dat was ook de makkelijkste vraag van de additionele vragen. De betrouwbaarheid van de algemene indruk ('Indruk') is laag. Dit lijkt ermee samen te hangen dat beoordelaars verschillen in het gemak waarmee zij hoge scores geven bij vraag naar algemene indruk. Om dit na te gaan is de correlatie berekend tussen het algemene oordeel van een beoordelaar als deze de eerste beoordelaar is en wanneer deze de tweede beoordelaar is. Een hoge correlatie zou erop kunnen wijzen dat een beoordelaar relatief vaak vergelijkbare oordelen geeft. De berekende correlatie is met 0.74 hoog te noemen, waarmee een beoordelaarseffect niet onwaarschijnlijk is. Algemene indruk heeft waarschijnlijk meer betekenis binnen een bepaalde beoordelaar dan over beoordelaars heen. Deze indruk heeft overigens geen impact op de schaal.

Bij de vragen die betrekking hebben op correctheid (bij de nieuwe taken) zijn de beoordelingen voor interpunctie redelijk betrouwbaar en voor spelling voldoende betrouwbaar. De vraag over de hoeveelheid niet-lopende zinsconstructies (bij de nieuwe taken) is aanzienlijk minder betrouwbaar. Ook de vraag naar schrijffouten (spelling bij de ankertaken) is niet heel betrouwbaar.

5.5 Stap 3: Van taak naar schaal - OPLM-analyses

5.5.1 Voorbereiding data: taken combineren

Voor het maken van OPLM-schalen moesten eerst de resultaten van de verschillende taken per leerling bij elkaar gebracht worden. Allereerst is per leerlingnummer nagegaan welke taken de leerling gemaakt heeft en is nagegaan in hoeverre de combinatie van taken al dan niet conform design is geweest. Er waren 3459 unieke leerlingnummers in de dataset. In totaal waren er 18 combinaties die volgens het design konden voorkomen. Dat zijn zes combinaties met vier taken (BO leerlingen) en 12 combinaties met twee taken (SBO leerlingen). Alle gevonden combinaties met 4 taken waren volgens

het design bij bo-leerlingen en alle leerlingen met vier taken zijn ook bo-leerlingen, conform de verwachting. Er bleken echter voor 442 leerlingen niet-design combinaties van 1, 2, 3 of 5 taken te zijn. In totaal waren bij 354 bo-leerlingen niet-design combinaties vastgesteld en bij 84 sbo-leerlingen. Daarnaast waren er nog vier niet-design combinaties bij leerlingen van wie het schooltype niet kon worden vastgesteld. De niet-design combinaties werden relatief vaker aangetroffen bij sbo-leerlingen dan bij bo-leerlingen. Deze SBO-leerlingen hadden één taak gemaakt, hetgeen geen onverwachte uitkomst is. Het is immers best waarschijnlijk dat een deel van de sbo-leerlingen één in plaats van twee taken heeft gemaakt. Het definitieve onderscheid tussen de leerlingen die daadwerkelijk de waargenomen combinaties van taken gemaakt hebben en administratieve fouten is bij beide typen onderwijs echter slecht te maken.

Er waren drie sbo-leerlingen met drie taken. Bij deze sbo-leerlingen kan het zo zijn dat door een administratieve fout twee verschillende leerlingen als een enkele leerling worden gezien. Dat is het enige type fout dat wel schade in de schaling op kan leveren. Om die reden zijn deze drie sbo-leerlingen niet in de schaling opgenomen. Ditzelfde probleem wordt gevonden bij alle combinaties van vijf taken door bo-leerlingen. Verwijdering van één van de taken resulteerde bij deze laatste type combinaties echter altijd in een combinatie die volgens het design te verwachten was. De combinaties van vijf taken worden allen toegewezen aan een van de zes reguliere boekjes, na verwijdering van de aberrante taak. Dit leverde in totaal 61 bruikbare verschillende combinaties van taken per leerlingnummer, waarbij iedere combinatie een eigen OPLM-boekje vormde.

Om de dataset te maken zijn vervolgens per taak de scores van de eerste beoordelaar toegevoegd aan ieder leerlingnummer in de lijst. In de gevallen van een enkele beoordelaar was dat ook de enige beoordeling. Bij twee beoordelaars is de eerste beoordelaar random toegewezen. Merk hierbij op dat de leerlingnummers niet noodzakelijk unieke leerlingen waren, hoewel dit bij de ruime meerderheid het geval zal zijn. Van de 86% van de records (2963 van de 3459) die volledig volgens het design zijn ingevuld was het zeer waarschijnlijk dat dit daadwerkelijk unieke leerlingen zijn.⁵² Ook de 20 leerlingnummers met vijf taken zijn waarschijnlijk (na verwijdering van de aberrante taken) unieke leerlingen. Het is echter mogelijk dat leerlingen onder twee nummers in de lijst voorkomen. Dit was achteraf niet meer te achterhalen.

Bij de lijst van 3459 leerlingnummers met de taken die daar bij horen is vervolgens gekeken welke taken niet te beoordelen waren volgens de beoordelaar(s). Als taken hier uitvielen veranderde de kandidaat van boekje doordat er een andere selectie van taken beoordeeld is. In totaal zijn er 169 leerlingnummers van boekje veranderd, omdat er minstens een taak niet beoordeeld is. Dat betreft 95 bo-leerlingnummers en 74 sbo-leerlingnummers. Dat zijn relatief veel sbo-leerlingnummers (8,0% versus 3,7% bij bo). De aantallen leerlingnummers met het aantal taken dat daarbij niet verder beoordeeld is, is gegeven in Tabel 5.8.

⁵² Zoals uit de tekst blijkt, zijn records niet per definitie unieke leerlingen. Uiteindelijk bleken er van 3319 leerlingen van de 3358 leerlingen die aan de peiling deelnamen, beoordelingen van schrijftaken aanwezig te zijn (van 39 leerlingen waren er enkel vragenlijstdata aanwezig). Ter identificatie van de leerlingen heeft elke aangemelde leerling voorafgaand aan de afname een zevencijferige code toebedeeld gekregen (schoolXXXX, klasY, leerlingZZ). Daarnaast waren er per klas twee reservecodes. Uitgaande van de 3459 beschreven records bevatte het bestand 140 records waarbij er, door administratieve fouten of problemen bij de technische verwerking, zevencijferige codes ingevoerd zijn die, gezien de identificeringssystematiek, geen bestaand leerlingnummer konden representeren. Achteraf konden de correcte leerlingnummers voor deze beoordelingen (deze 140 records) niet meer eenduidig worden vastgesteld. Deze administratieve fouten hadden gevolgen voor het aantal beoordelingen dat voor sommige leerlingen aanwezig was. Leerlingen maakten immers meerdere taken.

Tabel 5.8: Aantal taken uit per leerlingnummer (N tkn uit)

N tkn uit	Aantal leerlingnummers			Percentage leerlingnummers		
	BO	SBO	Totaal	BO	SBO	Totaal
0	2444	846	3290	96.3%	92.0%	95.1%
1	73	69	142	2.9%	7.5%	4.1%
2	16	5	21	0.6%	0.5%	0.6%
3	6	0	6	0.2%	0.0%	0.2%
Totaal	2539	920	3459			

In Tabel 5.9 is de verschuiving aangegeven van de startsituatie (aantal taken per leerlingnummer) naar de situatie na eliminatie van de niet te beoordelen taken. De vetgedrukte aantallen betreffen de aantallen leerlingnummers waarbij alle taken te beoordelen waren. De som van die aantallen is 3290, oftewel het totaal aantal leerlingnummers waarbij geen enkele taak is geëlimineerd (zie Tabel 5.8). Het is ook duidelijk dat bij 30 leerlingnummers (in rood) taken vervallen omdat er geen enkele taak als te beoordelen is geclassificeerd. Van deze vervallen leerlingnummers zijn er 19 bo-nummers en 11 sbo-nummers. De verschuivingen hadden enerzijds de eliminatie van gevonden boekjes tot gevolg (vier boekjes) en anderzijds drie nieuwe unieke combinaties (boekjes) tot gevolg.

Tabel 5.9: Aantal taken per leerling voor en na de eliminatie van niet te beoordelen taken

Aantal taken start	Aantal taken na eliminatie niet te beoordelen taken					Totaal
	0	1	2	3	4	
1	15	174				189
2	9	69	799			877
3	6	3	18	215		242
4			9	40	2102	2151
Totaal	30	246	826	255	2102	3459

5.5.2 Voorbereiden design

Voordat de OPLM analyses gedaan konden worden, moest het design beschreven worden. Dat is gedaan door per record aan te geven welke taken daarbij horen. Per taak zijn de items gegeven. Naast de beoordeelde punten zijn ook de somscores van de "afhankelijke" vragen in de dataset opgenomen. Dat is gedaan, omdat het OPLM uitgaat van lokaal onafhankelijke opgaven. Dat is bij de afhankelijke opgaven - zoals eerder beschreven - niet het geval. Bij de losse schalen (O, F en I) zullen echter ook losse items gebruikt moeten worden, omdat de afhankelijke opgaven veelal combinaties van de drie typen betreffen. In de beschrijving van de analyses is weergegeven hoe dat is aangepakt.

In Tabel 5.10 staat een overzicht van alle soorten items. Bij het type wordt er onderscheid gemaakt tussen de OFI-items en de overige items. Een aantal items komt alleen in een bepaald type taak voor (ankertaak dan wel een nieuwe taak). Er zijn twee soorten combinatie-items (som). Enerzijds zijn er mogelijke combinaties van hetzelfde type. Die krijgen de naam van het type (O, F of I). Daarnaast zijn er mogelijke combinaties van verschillende types. In zeven van de gevallen is dat een combinatie van een O en een F item. Dit is de combinatie genaamd 'V'. In twee gevallen, aangeduid met 'Q', zijn het unieke combinaties (een maal een I, een O en 2F, en een maal drie I plus een O item).

Naast de overall schrijfvaardigheidsschaal zijn er ook subschalen (O_schaal, F_schaal en I_schaal) gemaakt. De gecombineerde somscores V en Q kunnen niet in deze subschalen gebruikt worden. Deze gecombineerde somscores V en Q worden daarom losgelaten bij de subschalen; de losse items waaruit zij zijn opgebouwd zijn in de betreffende subschalen meegenomen. Somscores die bestaan uit combinaties van hetzelfde type (bijvoorbeeld de som van 2 O-items), blijven ook bij subschalen gehandhaafd, omdat deze sommen dezelfde type items bevatten als de subschaal.

Tabel 5.10: Overzicht van items in de data

Type	Beoordeling	Aantal items
Gescoord item op formulier	Enkele items	316
O	Opbouw	92
F	Formulering	47
I	Inhoud	114
Y	Algemeen: Taak voldaan? (anker en nieuw)	12
A	Algemeen: Indruk (anker en nieuw)	12
H	Algemeen: Hoeveelheid (nieuw)	9
C1	Correctheid: interpunctie (nieuw)	9
C2	Correctheid: zinsconstructie (nieuw)	9
C3	Correctheid: spelling (nieuw)	9
S	Correctheid: Schrijffouten (Anker)	3
Som score	Combinatie-items	21
O	Combinatie Opbouw	5
F	Combinatie Formulering	1
I	Combinatie Inhoud	6
V	Combinatie O+F	7
Q⁵³	Bijzondere combinatie	2
Totaal		337

Al met al is er een lijst van 337 opgaven, met verschillende maximumscores. Deze lijst bevat zowel de losse OFI opgaven als de combinatie-items, bestaande uit diezelfde losse OFI-opgaven. Omdat er meer dan eens sprake is van testlets (gegroepeerde items) en afhankelijke vragen in de observatieformulieren, bestaat hier een vrij ruime overlap: combinatie-items bestonden soms uit wel zes losse OFI-opgaven (met een maximumscore van 8). Deels sluiten de in Tabel 5.10 genoemde opgaven elkaar dus uit, zoals de combinatie-items en de losse opgaven, die niet gelijktijdig in een schaal zullen worden opgenomen. Deels zullen de gepresenteerde items geen onderdeel van de schaal worden, zoals de Y, A, H, C en S items. Het gaat daarbij om in totaal 63 Y, A, H, C en S items. Door middel van item-selectie-files zijn de opgaven geselecteerd in de schaal. Ieder van de opgaven is toegewezen aan de boekjes. De eerste check of het design toegepast is, is op de gemiddelde scores. Dit leidde tot het verwijderen van 26 taken⁵⁴.

⁵³ In de oorspronkelijke boekjes zijn deze bijzondere combinaties opgenomen met code 'W'. In latere analyses zijn zij hergecodeerd als Q.

⁵⁴ Voor ieder van de opgaven is de gemiddelde score uitgerekend en is deze vergeleken met de gemiddelde scores die bekend zijn uit SPSS per taak, nadat de te elimineren opgaven uit de analyses gehaald zijn. Deze gemiddelden komen zeer sterk

5.6 Stap 4: OPLM analyses

5.6.1 Vooranalyses

Voordat de echte OPLM plaatsvond, zijn nog enige analyses vooraf uitgevoerd. Allereerst betrof dit factoranalyses om vooraf een indruk te hebben of er losse schalen per subschaal (O, F en I) zouden ontstaan of dat dit wellicht losse schalen per taak zouden zijn. Uit deze resultaten kwam niet een eenduidig beeld naar voren, behalve dat de algemene scores en de score voor correctheid niet echt op de schaal thuishoorden. Een tweede vooranalyse betrof analyses waarbij binnen OPLM geanalyseerd werd of bij de polytome opgaven ordening van de scorecategorieën geheel monotoon was (levert een hogere waarde ook meestal een hogere score op?). Dat bleek bij een drietal opgaven, allen uit Taak 03, niet het geval te zijn. Bij de beschrijving van de definitieve OPLM analyses staat beschreven hoe met deze items is omgegaan. Ten derde zijn de sets van losse opgaven verder onderzocht door middel van analyses waarbij gekeken is wat de resultaten zouden zijn als afhankelijkheid genegeerd zou worden en niet de somscores gebruikt zouden worden. Omdat het belang van de afhankelijke vragen wordt opgeblazen, wordt een modelschending (gebrek aan lokaal stochastische onafhankelijkheid) als modelkenmerk geïncorporeerd. Tenslotte bleek op basis van verdere analyses dat er bij Taak 04 en 07 som-items moesten worden gemaakt, zodat inhoudelijk gekoppelde opgaven de schaal niet sterk domineren. Enkele grafische weergaven van goed en slecht passende opgaven staan weergegeven in Bijlage 8.

5.6.2 OPLM schaal analyses

De OPLM schaalanalyse voor de O, F, I- items samen (OFI-schaal) zijn gedaan met de hierboven genoemde som-opgaven. Er bleken zeven items, waaronder twee combinatie-items in T03, met een dalende item-response-functie (IRF), hetgeen niet toegestaan is (zie Tabel 5.11). Inhoudelijk is nagegaan waarom deze items afwijkend functioneerden. Zo is bijvoorbeeld het vermelden van datum en plaats boven een brief sporadisch aangetroffen (item 4, taak 4), waardoor het twijfelachtig is of leerlingen überhaupt bekend zijn met deze conventie. De niet gevraagde mening (en eventuele toelichting daarop) en het opleveren van additionele informatie (resp. item 12 en 13 en item 14 en 15, Taak 3) hoeven niet per definitie een indicatie te zijn van goede/zwakke schrijfvaardigheid binnen een samenvatting. Omdat inhoudelijk geduid kon worden waarom de zeven items afwijkend functioneerden, zijn zij buiten beschouwing gebleven bij het bepalen van de vaardigheidsscores. Het betreft de volgende items uit Taak 03, 04 en 07:

overeen, maar zijn in enkele gevallen niet gelijk. Dat is het geval doordat bij de analyses per taak er geen taken zijn verwijderd bij de combinatie per 5 taken: dat is in OPLM wel het geval om de design-boekjes te verkrijgen en om zo de kans dat verschillende leerlingen onder een nummer vallen te verkleinen. In totaal scheidt dit 26 taken: 20 keer 1 taak minder (bij set van 5 taken), en twee keer 3 taken minder (bij de SBO leerlingen).

Tabel 5.11: Opgaven met dalende item-response-functies

OPLM nummer	Item Label ⁵⁵	Gam _i
70	T03S122I ⁵⁶	-0,01
71	T03S142I	-0,04
73	T04i04_O	-0,39
86	T04i17_O	-0,09
191	T07i40_F	-0,22
193	T07i42_F	-0,73
196	T07i45_O	-0,67

Na samenvoeging van verschillende items tot de combinatie-items (bestaande uit dezelfde of juist verschillende items), extra gegroepeerde items (zogenoemde testlets) en het verwijderen van de niet-goed passende items blijft er een algemene schrijfvaardigheidsschaal over van 186 opgaven⁵⁷⁵⁸. Daarvan zijn er 111 dichotoom (twee score categorieën, maximum score 1). De overige 75 opgaven zijn polytoom, waarvan 27 samengestelde som-opgaven.⁵⁹ Het totaal aantal scorepunten dat te verdelen is, is 311 als iedereen alle taken zou maken. In Tabel 5.12 wordt per taak een overzicht gegeven van het aantal OFI-opgaven in de beoordelingsformulieren en het aantal enkele opgaven en combinatie-items die in de schrijfvaardigheidsschaal zijn opgenomen. Zoals Tabel 5.12 laat zien, zijn de 253 OFI-opgaven in de beoordelingsformulieren daarbij ingedikt naar 186. In Tabel 5.5 was het aantal van 253 OFI-opgaven reeds gereduceerd tot 211 OFI*-opgaven, maar de noodzakelijke toevoeging van somscores in Taak 4 en 7 (ter voorkoming van het opblazen van losse opgaven die eigenlijk bij elkaar horen) heeft uiteindelijk geleid tot een schaal bestaande uit 186 opgaven.⁶⁰

Tabel 5.12 laat zien dat er met name bij T04, T07, T09, T10, T11 en T12 een verschil *lijkt* in de aantallen opgenomen items in de schaal, maar dit betreft dus taken met relatief veel combinatie-items als gevolg van afhankelijke en gegroepeerde items. Zo wordt bij alle brieven gevraagd naar aanhef en de gepastheid daarvan, naar de aanwezigheid van een groet en de gepastheid daarvan en naar de aanwezigheid, en naar de specifieke inhoud en herkenbaarheid van een afsluitende zin. Dit zijn afhankelijke vragen in T04, T09, T10, T11. Een ander type taak, het recept (T07), bevat bijvoorbeeld meerdere gegroepeerde items, die gericht zijn op verschillende fasen in het bereidingsproces. Tenslotte wordt bij persuasieve teksten (T11 en T12) gevraagd naar de aanwezigheid van stellingname en argumenten (afhankelijke vragen). Concluderend kan gesteld worden, dat afgezien van de eerder beschreven afgevallen zeven items met een dalende IRF, alle losse opgaven in de

⁵⁵ Item label bestaat uit drie delen: de eerste drie posities betreffen de taak (T[Taaknummer]), de volgende vier betreft het item of de som ([i]\$\$_los item, met \$\$ als item volgnummer in taak)/S\$\$%=som item, met \$\$ eerste itemnummer in taak van de som, % het aantal achtereenvolgende items gesommeerd]), de laatste positie betreft de schaal waar het item toe behoort (type schaal, O, F, I of mogelijk een combi, bij somitems)

⁵⁶ Merk op dat T03S122I en T03S142I combinatie-items zijn, bestaande uit twee afhankelijke vragen. Alle vier de losse items en het combinatie-item hadden een dalende IRF.

⁵⁷ Merk nogmaals op dat Tabel 5.10 een overlap in items bevat: items die onder enkele items vallen, vallen soms ook onder combinatie-items. Deze worden in de vaardigheidsschaal echter niet gelijktijdig opgenomen. Indien een item is opgenomen in een somscore is dit item niet ook nog als enkel item in de schaal aanwezig.

⁵⁸ Zoals in hoofdstuk 8 wordt beschreven, is voor de analyses aangaande de leerlingprestaties gebruik gemaakt van een vaardigheidsschaal bestaande uit 176 items. Daar wordt ook beschreven op welke punten deze schaal verschilt van de hier en de in hoofdstuk 7 gebruikte schaal. Overigens leidt het al dan niet meenemen van de 10 betreffende items niet tot verschillende resultaten, zoals uit analyses is gebleken.

⁵⁹ Van de polytome opgaven is bij 48 de maximumscore gelijk aan 2 en bij 15 is dat maximaal 3. Bij 7 is de maximumscore 4, bij 3 is dat 5 en bij 2 (somscores) is dat 8.

⁶⁰ Vergeleken met de lijst opgaven in Tabel 5.10 zijn er nog zes extra somscores/combinatie-items vastgesteld in T04 en T07, zie paragraaf 5.6.1.

beoordelingsformulieren in de schaal zijn opgenomen, hetzij als losse opgave, hetzij als onderdeel van een combinatie-item.

Tabel 5.12: Overzicht van de aantallen losse OFI-opgaven en OFI-combinatie-items in de vaardigheidsschaal

Taak nummer	Taak naam	Genre	Anker	Aantal OFI opgaven (in observatie-formulier) ⁶¹	Aantal losse OFI vragen	Aantal combinatie-items (OFI)
01	Wim	Verhaal	Anker	19	17	1
02	Verboden boek	Verhaal	Nieuw	13	13	0
03	Unicef	Samenvatting	Anker	24	20	0
04	Brief aan Annika	Brief	Nieuw	25	11	4
05	Nepnieuws	Tijdschriftartikel	Nieuw	19	19	0
06	Vlinder	Biologie-tekst	Nieuw	17	17	0
07	Pannenkoeken	Recept	Nieuw	42	16	7
08	Poekie	Flyer	Anker	23	19	2
09	Naima	Brief	Nieuw	24	7	5
10	Speelveld	Brief	Nieuw	20	10	3
11	Korte zomervakantie	Brief	Nieuw	19	6	4
12	Frisdrank en snoep	Stuk in de krant	Nieuw	8	4	1

De modelpassing is onderzocht door de fit-statistics te evalueren. De R1c statistics is 5262,214 bij 3295 vrijheidsgraden wat duidt op een misfit ($p=0.0000$)⁶². Bij 837 tot 1046 observaties per opgave (930 gemiddeld) is het OPLM echter een behoorlijk streng model als je kijkt naar modelpassing. Het is bij dergelijke aantallen gebruikelijk dat er een slechte passing gevonden wordt: het betekent niet automatisch dat het model niet gebruikt kan worden. Om te onderzoeken of het model kan worden gebruikt is gekeken of de voorspelling met het model enigszins deugdelijk is. Daartoe is gekeken of met het model de bekende kenmerken teruggeschat kunnen worden. Dat bleek het geval te zijn. Met behulp van MML is de populatieverdeling bepaald en samen met de itemparameters zijn zo bijvoorbeeld de p-waarden van de items geschat. Deze bleken zeer goed terug te schatten te zijn, met een zeer beperkte afwijking, namelijk een gemiddelde absolute afwijking van 0,007 verschil in p-waarde tussen schatting en geobserveerde waarde. Concluderend kan gesteld worden dat de schaling goed bruikbaar is.

Naast modelpassing is ook gekeken naar de impact per taak en per onderdeel O, F en I. Daarbij blijken met name Taak 04 en Taak 07 een relatief groot gewicht te hebben en Taak 12 een relatief laag gewicht. Dat was overigens ook al te verwachten op basis van het aantal opgaven (vergelijk Tabel 5.1). Kijken we naar de O, F en I opgaven, dan zien we daar niet een dominantie van een van de drie subschaal-items, anders dan te verwachten door de aantallen opgaven. De gemiddelde a-parameters zijn voor de O-items iets lager.

⁶¹ Hier gaat het om het aantal OFI-vragen die in het beoordelingsformulier zijn opgenomen, zie Tabel 5.1.

⁶² Daarnaast is binnen OPLM gekeken naar itemstappen. Van de 311 itemstappen hebben er 30 een misfit met $p=.01$ en 53 met een $p=.05$. Dat zijn er (aanzienlijk) meer dan je onder kans mag verwachten.

Er zijn ook analyses uitgevoerd voor itembias. Hierbij is gekeken naar het type onderwijs: bo versus sbo. Het aantal opgaven met DIF lijkt zeer beperkt. Er zijn enkele opgaven gevonden die mogelijk DIF hebben, maar de grootte van de DIF leek beperkt en het aantal opgaven lag daarbij met 7% rond het niveau dat te verwachten valt wanneer op 5% getoetst wordt. Er was geen reden om opgaven te verwijderen op basis van DIF.

Bij de opgaven was te zien dat bij 90% van de opgaven leerlingen in het bo beter presteren dan leerlingen in het sbo. Bij 10% is het verschil in p-waarden meer dan 0,20 ten faveure van het bo. Het grootste gevonden verschil waarbij sbo beter presteert dan bo is 0,11, en het grootst gevonden verschil met bo beter presterend dan sbo is 0,31. Gemiddeld is de p-waarde bij bo 0,09 hoger dan bij sbo. Dat verschil wordt gevonden bij zowel O-, F- als I-items, zij het dat als iets gedetailleerder gekeken wordt, te zien is dat bij O- en F-items het verschil 0,092 is en bij I-items 0,087. Bij de taken is er ook enig onderscheid. De grootste verschillen worden gevonden bij Taak 4 en 5, de kleinste bij Taak 10 en Taak 12.

Naast de OFI-schaal zijn ook de O-, F- en I-schaal afzonderlijk bepaald. Het deel van de samengestelde opgaven dat uit twee typen bestaat, is opgesplitst in losse opgaven, waarbij het losse item aan de juiste schaal is toegewezen. Als de opgaven afzonderlijk in de schaal zitten, zijn ze niet meer onderling afhankelijk. Dit levert een vergelijkbare passing als op de gehele schaal op. Er worden niet meer of minder items verwijderd. Ter controle zijn er ook nog analyses gedaan met de algemene en correctheidsvragen. Deze vormen samen ook een schaal, en hier is ook weinig sprake van DIF. De resultaten worden hier niet weergegeven omdat verder niet met deze schaal gewerkt wordt. Deze schaal is eerder ter controle om deze opgaven ook samen te kunnen vatten.

Hoewel de modelpassing in eerste instantie niet optimaal leek, kan op basis van de verschillende analyses die gedaan zijn om de kwaliteit van de schaling na te gaan, gesteld worden dat de schaling goed bruikbaar is. Schattingen op basis van het OPLM-model zijn namelijk goed terug te schatten, de impact per taak en subschaal is conform verwachting en het model is in staat om verschillende vaardigheden van verschillende voor de peiling relevante groepen (bo en sbo) te meten.

5.6.3 Schattingen met de OPLM schalen

Vaardigheden van leerlingen

Met de verkregen schalen zijn zowel voor de individuele leerlingen als voor de populaties schattingen van de vaardigheid gemaakt. Voor ieder van de leerlingen is een WML-schatting gemaakt van de vaardigheid op de OFI-schaal, maar ook op de drie subschalen. Met behulp van de kennis over de schaal is ook voor ieder van de afzonderlijke taken te bepalen welke prestatie het best overeenkomt met een bepaalde vaardigheid. Deze bepalingen per taak, de algemene schattingen voor de OFI-schaalwaarde per leerling en de subschaal-waarden worden gebruikt bij de selectie van leerling-werken bij de standaardbepaling (zie hoofdstuk 7).

Nu de WML-schattingen bepaald zijn voor de subschalen is ook te zien wat de correlatie is tussen de WML-schattingen. Deze kunnen gecorrigeerd worden voor de gevonden betrouwbaarheid per subschaal. Deze analyses laten een zeer hoge (latente) correlatie zien tussen O en F. De correlatie tussen I en O (0.81) en I en F (0.73) is wat lager. Het heeft de voorkeur om de eenheid van de OFI-schaal te bewaren door de gehele schaal te gebruiken.

Eigenschappen van populaties en opgaven

Naast de individuele schattingen zijn er ook schattingen voor de populaties gemaakt met behulp van OPMML. De resultaten van de MML schattingen worden gegeven in Tabel 5.13. Op basis daarvan kunnen inschattingen gemaakt worden van de effectgrootte van het verschil tussen bo- en sbo-leerlingen in de designboekjes. Deze liggen bij alle vier de schalen (OFI samen, of O, F en I afzonderlijk) rond de 1, hetgeen als groot geclassificeerd kan worden. In hoofdstuk 8 wordt nader ingegaan op de gevonden verschillen.

Tabel 5.13: Populatieverdeling zoals bepaald met OPMML

MML	Gemiddelden				Standaardafwijking			
	Totaal schaal	O schaal	F schaal	I schaal	Totaal schaal	O schaal	F schaal	I schaal
ALL	-0,07	-0,06	-0,06	-0,06	0,18	0,21	0,21	0,19
BO	-0,02	-0,02	-0,02	-0,02	0,16	0,18	0,18	0,17
SBO	-0,18	-0,22	-0,21	-0,18	0,18	0,22	0,22	0,19
Effectgrootte	1.02	1.06	0.97	0.90				

Nu de schatting van de verdeling van de populatie bekend is, kunnen de p-waarde, item-totaal en item-rest correlaties geschat worden door de MML-populatieschattingen (met gefixeerde item-parameters) te combineren met de itemparameters (zie Bijlage 9). Deze kunnen gegeven worden voor de gehele OFI-schaal, of voor de O, F en I-schaal afzonderlijk. Dit kan (zeer) kleine verschillen opleveren, maar de schattingen van de populatie lijken het meest stabiel voor de OFI-schaal, aangezien daarvoor per leerling de meeste informatie is gebruikt. De geschatte waarden zijn gegeven voor de gehele steekproef (ALL; ongewogen), en voor de groepen bo en sbo afzonderlijk. Deze laatste twee schattingen zijn het meest interessant, aangezien de aanduiding van de populatie het duidelijkst is. In de ALL-groep domineert de bo-groep namelijk. Dit zou verder alleen maar versterkt worden als de gewogen waarden voor ALL berekend zouden worden, aangezien het sbo in onze steekproef oververtegenwoordigd is. De impact van sbo op ALL is in de populatie ook kleiner dan in onze steekproef.

5.7 Vragenlijstonderzoek

Leerlingvragenlijst

De leerlingen zijn niet alleen onderzocht door middel van de hiervoor beschreven opdracht, maar er is ook een vragenlijst bij hen afgenomen. Het betreft een vragenlijst met in totaal 31 vragen. De vragenlijst bevatte vier onderdelen.

Het eerste onderdeel, bestaande uit vijf vragen, ging over de thuissituatie voor wat betreft taal. Deze vragen zijn genummerd 1 tot en met 4, waarbij de eerste vraag uit twee delen bestaat, met alleen een antwoord op het tweede deel als "Nee" het antwoord is op de eerste vraag. In de meeste

gevallen is dit antwoord ontbrekend, zoals verwacht kon worden omdat een zeer ruime meerderheid van de leerlingen het antwoord “Ja” geeft. Bij vraag 3 en 4 konden leerlingen meer hokjes aankruisen, hetgeen nominale variabelen opleverde. Het tweede gedeelte ging over de vraag hoe vaak en gedurende hoeveel tijd er geschreven wordt in de vrije tijd. Dit betreft vraag 5 tot en met 8. Het derde onderdeel, vraag 9 tot met 18, betrof vragen naar de schrijfervaring. Deze tien vragen hebben alle drie geordende antwoordcategorieën (“Nooit”, “Soms”, “Vaak”). Bij het vierde en laatste onderdeel (vraag 19) werd leerlingen gevraagd te reageren op twaalf stellingen, die gerelateerd waren aan schrijfplezier en self efficacy. Deze vragen zijn opgevat als subvragen onder vraag 19, en worden aangegeven als 19.01 tot en met 19.12. Hier moesten leerlingen reageren op een twaalfstal stellingen, waarbij ze gebruik konden maken van vijf geordende antwoordcategorieën: “Helemaal mee oneens”, “Mee oneens”, “Niet mee eens, niet mee oneens”, “Mee eens” en “Helemaal mee eens”.

Op alle scholen zijn vragenlijsten op leerlingnummer uitgezet, gebaseerd op de leerlinginformatie die op voorhand van de school was verkregen. Om voorbereid te zijn op eventuele extra leerlingen zijn tevens per klas twee vragenlijsten geprint die toegekend waren aan twee reservenummers. Ingeval van extra leerlingen in de klas zijn deze gebruikt.⁶³ Bruikbare vragenlijstinformatie is aanwezig voor 3342 leerlingen. Dit is een respons van 99.5% van de leerlingen die daadwerkelijk aan de peiling hebben deelgenomen. Van de 3341 leerlingen zaten 2438 op het reguliere basisonderwijs (bo) en 903 op het speciaal basisonderwijs (sbo). Al deze leerlingen hebben op minstens een vraag een antwoord gegeven.⁶⁴ Van deze records zijn er 3066 (91.7% van 3342) waarbij geen enkele vraag over is geslagen; hierbij is het tweede deel van vraag 1 niet meegerekend. 197 leerlingen (5.9%) hebben één vraag overgeslagen, en 35 leerlingen (1.0%) twee vragen. In totaal was er bij 99.4% van de records op meer dan 80% van de vragen een antwoord gegeven.

Schaling leerlingvragenlijst - vraag 19

De vragen in de vragenlijst hadden primair tot doel inventariserend te zijn. Behalve vraag 19 werden de vragen niet geacht schalen te vormen (naar DeMiliano, 2013). Bij vraag 19 was op voorhand de veronderstelling dat de eerste zes vragen (v19.01 tot en met v19.06) de schaal “**schrijfplezier**” vormden, en de laatste zes vragen (v19.06 tot en met v19.12) de schaal “**self efficacy**”. Hogere scores op beide veronderstelde schalen dienden indicatief te zijn voor schrijfplezier, dan wel self efficacy. Dat betekent dat een score 0 het minst indicatief is voor schrijfplezier dan wel self efficacy en score 4 het meest indicatief.⁶⁵ Om dit te bewerkstelligen is de scoring van een vijftal vragen gehercodeerd. Bij de vragen 5 en 6 waren de antwoorden “Helemaal mee oneens” juist het meest indicatief voor schrijfplezier (bij de vragen 1 tot en met 4 was dat “helemaal mee eens”). Bij self efficacy zijn de scores voor de vragen 9, 10 en 12 omgekeerd, op dezelfde wijze als bij schrijfplezier. Bij de overige drie vragen (7, 8 en 11) leverden “Helemaal mee oneens” een score 0 op en “Helemaal mee eens” een score 4.

In Tabel 5.14 zijn de kenmerken wat betreft de betrouwbaarheid gegeven van deze twee schalen. Het is duidelijk dat “schrijfplezier” voldoende betrouwbaar is, maar dat de betrouwbaarheid van self efficacy tegenvalt.

⁶³ Hetzelfde reservenummer is gebruikt voor vragenlijst en voor de schrijftaken.

⁶⁴ Overigens zijn er 7 leerlingen die slechts 1 vraag hebben ingevuld en nog eens 4 die 2 vragen hebben ingevuld. Het betreft vraag 5 en 13. De indruk bestaat dat dit geen ingevulde vragenlijsten betreft. In dat geval heeft 99.7% van de leerlingen minstens een antwoord gegeven.

⁶⁵ De tussenliggende antwoorden zijn gewoon geordend.

Tabel 5.14: Betrouwbaarheid schalen Schrijfplezier (6 items) en Self Efficacy (6 items)

	Aantal vragen	Vraagnummers	Cronbach's Alpha	Guttman's Lambda 2
SP6: Schrijfplezier 6	6	.01 - .06	0.797	0.804
SE6: Self efficacy 6	6	.07 - .12	0.556	0.589

Bij self efficacy (SE) wordt de betrouwbaarheid van de schaal negatief beïnvloed door vraag 7, zoals Tabel 5.15 laat zien. De betrouwbaarheid van de SE-schaal stijgt als deze vraag uit de schaal verwijderd wordt.

Tabel 5.15: Itemeigenschappen per schaal.

Schrijfplezier				Alpha		Self Efficacy				Alpha	
	Gem	SD	Rit	als	impact		Gem	SD	Rit	als	impact
S19.01	1,58	1,21	0.57	0.76	0.04	S19.07	1,81	1,29	0.19	0..57	-0.01
S19.02	1,34	1,19	0.48	0.78	0.02	S19.08	2,18	1,12	0.36	0..48	0.07
S19.03	1,51	1,22	0.67	0.74	0.06	S19.09	3,16	1,02	0.30	0..51	0.05
S19.04	1,83	1,30	0.45	0.79	0.01	S19.10	2,87	1,21	0.22	0..55	0.01
S19.05	2,27	1,31	0.57	0.76	0.04	S19.11	1,82	1,22	0.42	0..45	0.11
S19.06	2,12	1,34	0.58	0.76	0.04	S19.12	2,84	1,07	0.32	0..50	0.06

Op deze analyses zijn ook factoranalyses uitgevoerd die suggereren dat vraag v19.07 beter bij de schaal "schrijfplezier" past. Bij het verplaatsen van vraag 7 naar de schaal schrijfplezier stijgt namelijk de betrouwbaarheid van zowel schrijfplezier als die van self efficacy. Ook inhoudelijk is de verplaatsing te onderbouwen omdat vraag 7, net als vraag 1 en 2, over schrijven in de vrije tijd gaat. Hoewel de betrouwbaarheid van de schaal self efficacy met een betrouwbaarheid van 0.59 net niet voldoende betrouwbaar is (voor analyses op stelselniveau, zie Tabel 5.16), denken wij dat een opdeling in SP7 en S5 goed te verdedigen is.

Tabel 5.16: Betrouwbaarheid schalen Schrijfplezier (7 items) en Self Efficacy (5 items)

Aantal vragen	aantal vragen	Vraagnummers	Cronbach's Alpha	Guttman's Lambda 2
SP7: Schrijfplezier 7	7	.01 - .07	0.810	0.815
SE 5: Self efficacy 5	5	.08 - .12	0.569	0.591

Tot slot is gekeken naar de vragen V09 tot en met V18. Deze hebben op zich weinig met elkaar te maken; de correlaties tussen deze opgaven zijn ook vrij laag (gemiddeld 0.14). De som van de opgaven kan echter gebruikt worden als een score voor "ervaring met schrijven in de vrije tijd". Deze schaal heeft een betrouwbaarheid van 0.62 (Lambda-2). Uitgaande van een minimale voldoende

betrouwbaarheid is deze schaal dus (net) voldoende betrouwbaar voor analyses op stelselniveau en daarom wordt ook over deze schaal score gerapporteerd in hoofdstuk 6.

Schoolvragenlijst en aanbodvragenlijst: schalen en variabelen in meerniveau-analyses

Er is gebruikt gemaakt van twee vragenlijsten om informatie te verzamelen over kenmerken van leerkrachten, het onderwijsleerproces en de school. Informatie over het onderwijsleerproces is daarbij verzameld middels een digitale schoolvragenlijst en een leerkrachtvragenlijst. Deze laatste bevatte ook vragen naar achtergrondkenmerken van de leerkracht.

Op alle 133 deelnemende scholen zijn digitale vragenlijsten uitgezet. In totaal zijn er 113 reacties geregistreerd, ondanks reminders om de vragenlijst in te vullen (85% van 133 scholen). Nadere inspectie liet echter zien dat 17 van deze reacties geen antwoorden bevatten (15% van de 113 reacties). Dit waren gevallen waarin men de vragenlijst weliswaar had aangeklikt, maar niet ingevuld. Op twee scholen is de vragenlijst zelfs twee keer geopend, maar is niets ingevuld. Uiteindelijk hebben 96 personen de schoolvragenlijst geopend en (deels) ingevuld (72% van de 133 scholen), waarvan de vragen op vijf scholen door twee medewerkers (gedeeltelijk) zijn beantwoord. Dit betekent dat in totaal 91 unieke scholen de vragenlijst hebben ingevuld. Voor 68% (91) van de 133 deelnemende scholen is daarmee informatie op schoolniveau beschikbaar. Hiervan zijn er 67 scholen voor het basisonderwijs (74%) en 24 scholen voor het speciaal basisonderwijs (26%). Nagegaan is het percentage missings per vraag. Alle respondenten hebben de vraag naar aanwezig taalbeleid op school ingevuld. Bij vraag 2 (naar details over het taalbeleid) hadden 21 respondenten vier items niet ingevuld (23%) en 20 respondenten (22%) de overige acht items. Het aantal missings bij items van vraag 3 (schoolbrede aandacht in schrijfonderwijs) en vraag 4 (visie op schrijfonderwijs) lag beduidend lager: 3%.

Van de 204 leerkrachten in groep 8 of de schoolverlaters aan wie de vragenlijst is voorgelegd, hebben 193 leerkrachten de lijst ingevuld (95% van de 204 leerkrachten). Specifiek voor het bo gaat het om 126 leerkrachten (65% van 193 leerkrachten), voor het sbo betreft het 67 (35% van de 193 leerkrachten). Van de 193 leerkrachten heeft 65% (n=125) alle vragen en subitems ingevuld. 32 leerkrachten (17%) hadden twee ontbrekende waarden en nog eens 15 (8%) vier. 3.4% van de leerkrachten had 10 ontbrekende waarden of meer. Het maximum aantal ontbrekende waarden was 38 van de 85 items (45%, bij 1 leerkracht).

Gebruikte schalen, indexen en variabelen schoolvragenlijst en leerkrachtvragenlijst

Zowel de schoolvragenlijst als de leerkrachtvragenlijst hadden primair tot doel te inventariseren op welke manier het schrijfonderwijs schoolbreed of op klasniveau wordt vormgegeven. De vragen waren niet primair opgesteld om schalen te vormen. Om de verzamelde informatie toch zinvol in te kunnen dikken voor de meerniveau-analyses, is met behulp van exploratieve factoranalyses gezocht naar onderliggende factoren. Voor de selectie van vragen waren de bewezen effectieve schrijfbepalingen die door Janssen en Van Weijen (2017) zijn beschreven leidend. Deze exploratieve factoranalyses zijn uitgevoerd ten behoeve van:

- Schoolvragenlijst
 - Vraag 2 (nadruk in taalbeleid)
 - Vraag 3 (nadruk in schrijfonderwijs)
 - Vraag 4 (visie op schrijfonderwijs) – geen schaling mogelijk
- Leerkrachtvragenlijst:

- Vraag 14 (feedback)
- Vraag 13 (activiteiten leerlingen)
- Vraag 17 (kennis van schrijfprocessen)

Voorts zijn indexen gemaakt voor vraag 5, 6, 7, 8 en 12. Tabel 5.17 toont een overzicht van de gemaakte schalen en indexen, en hun betrouwbaarheid (Cronbach's α). Alleen factoren en indexen met een voldoende of goede betrouwbaarheid worden in de analyses meegenomen, waarbij een minimale betrouwbaarheid van .60 is gehanteerd.

Tabel 5.17: Overzicht schalen en indexen onderwijsleerproces

Schoolvragenlijst	Schaal/index	Items	α
Taalbeleid gericht op het wegwerken van taalachterstanden bij (anderstalige)	Schaal	2_1, 2_2, 2_5, 2_9 en 2_10	0.77
Taalbeleid gericht op rijk/breed taalaanbod voor allen	Schaal	2_3, 2_4, 2_6, 2_7, 2_8 en 2_13	0.75
Focus brede facilitering van het schrijfonderwijs	Schaal	3_3, 3_4, 3_5, 3_6, 3_7, 3_8 en 3_9	0.88
Leerkrachtvragenlijst	Schaal/index	Items	α
Soorten teksten* ^a	Index	5_1 – 5_13	0.77
Soorten doelen* ^a	Index	6_1 – 6_5	0.67
Soorten publiek* ^a	Index	7_1 – 7_5	0.75
Feedback op inhoud	Schaal	14 a-c	0.79
Feedback op technische aspecten	Schaal	14 d-f (allen gehercodeerd)	0.64
Prewriting activiteiten	Schaal	13_1, 13_2, 13_3, 13_10	0.71
Revisie activiteiten	Schaal	13_4, 13_6, 13_7, 13_11,	0.88
Leren van anderen	Schaal	13_5, 13_8, 13_9	0.67
Kennis van schrijfprocessen	Schaal	17_1, 17_2, 17_3, 17_4, 17_5, 17_6	0.81
Samenwerking	Index	12_1, 12_2, 12_3	0.81
Schrijven om te leren	Index	8_1, 8_2, 8_3	0.69

* indexen gebaseerd op gedichotomiseerde waarden van de items (ja/nee)

^a: bij vraag 5, 6 en 7 is de optie niet gespecificeerd niet meegenomen.

Gebruikte variabelen in meerniveau-analyses

Tabel 5.18 laat verder een overzicht van de schalen, indexen en losse variabelen zien, die in de meerniveau-analyses zijn meegenomen. Een deel daarvan wordt tevens gebruikt in de beschrijvende analyses van de kenmerken van het onderwijsleerproces, de scholen, de leerkrachten en de leerlingen (hoofdstuk 6). In Tabel 5.18 staan, naast de in Tabel 5.16 gegeven schalen en indexen, tevens losse variabelen. Omdat er vaak sprake was van ordinale variabelen met 4 of 5 antwoordcategorieën zijn deze antwoorden meestal ingedikt tot dichotome variabelen. Dichotomisering vond grotendeels plaats op basis van frequentieverdelingen. Alleen bij de inzet van soorten teksten, doelen en publiek is per

type tekst, doel en publiek op voorhand onderscheid gemaakt tussen het wel of niet gebruiken van die typen, ongeacht de mate waarin dit gebeurde.

Tabel 5.18: Overzicht schalen, indexen en ingedikte variabelen t.b.v. meerniveau-analyses

Thema/onderwerp	Onderwerp	bron	Naam variabelen	Type variabele/antwoordcategorie
Leerlingkenmerken				
Achtergrond				
Achtergrond	land van herkomst	ll-vrl	V01_MC	NL; anders
	gebruikte thuistaal leerling met ouders	ll-vrl	V04_MC	NL (in combinatie met andere taal); dialect of Fries; andere taal
	leerlinggewicht	admin	lln_gw_num	0.0; 0.3; 1.2
	voorlopig schooladvies	admin	Advies	PRO, vmbo b/k, vmbo g/t, havo/vwo
	geslacht	admin	Geslacht	Jongen/meisje
	leeftijd	admin	XX	leeftijd per 1-4-2019
Domeinspecifiek				
Frequentie en duur schrijven	frequentie schrijven op mobiel (vorige week)	ll-vrl	V05_MC	5 categorieën
	frequentie schrijven vorige week	ll-vrl	V07_MC	5 categorieën
Attitude	Schrijfplezier	ll-vrl	SP7	Somscore
	Self efficacy	ll-vrl	SE5	somscore
Klasniveau				
Domeinspecifieke kenmerken				
Algemene aandacht voor schrijfonderwijs	bestede tijd aan schrijven (3_2)	Lkr-vrl	Schrijftijd_pw	aantal minuten per week
	frequentie schrijfpodrachten	Lkr-vrl	Sc_4_1_dich	minder dan 1 keer per maand vs 1 keer per maand of vaker
	frequentie schrijfpodrachten binnen andere vakken	Lkr-vrl	Sc_4_2_dich	minder dan 1 keer per maand vs 1 keer per maand of vaker

Tabel 5.18: Overzicht schalen, indexen en ingedikte variabelen t.b.v. meerniveau-analyses (vervolg)

Klasniveau				
Domeinspecifieke kenmerken				
Communicatief taalonderwijs	nadruk doel en publiek	Lkr-vrl	Sc_9_1_dich	(helemaal) niet van toepassing <i>versus</i> een beetje en (helemaal) van toepassing
	feedback op inhoud (schaal)	Lkr-vrl	schaal_fb_inhoud_mean	gemiddelde schaalscore
	kennis over genre	Lkr-vrl	Sc_16_1_dich	een keer per kwartaal of minder <i>versus</i> een keer per maand of meer
Technisch taalonderwijs	feedback op technische aspecten (schaal)	Lkr-vrl	schaal_fb_correctheid_mean	gemiddelde schaalscore
	kennis over gebruik leestekens en hoofdletters	Lkr-vrl	Sc_16_4_REC_dich	minder dan 1 keer pm <i>versus</i> 1 keer pm of meer
Strategie-onderwijs	prewriting (schaal)	Lkr-vrl	schaal_aanpak_mean	gemiddelde schaalscore
	revisie (schaal)	Lkr-vrl	Schaal_revisie_mean	gemiddelde schaalscore
	kennis schrijfprocessen (schaal)	Lkr-vrl	Schaal_proceskennis_mean	gemiddelde schaalscore
focus op leerdoel	feedback op leerdoelen	Lkr-vrl	Sc_14_b_dich	weinig(nooit en soms) <i>versus</i> veel (meestal en vaak)
Samenwerking tijdens het schrijfproces	samenwerking (som 12_1/3)	Lkr-vrl	Samenwerking_mean	gemiddelde schaalscore
leren van anderen	leren van anderen (schaal)	Lkr-vrl	Schaal_leren_van_anderen_mean	gemiddelde schaalscore
integratie met andere vakken	schrijven_om_te_leren	Lkr-vrl	schrijven_om_te_leren_mean	gemiddelde schaalscore
aandachtsgebieden binnen schrijfonderwijs	soorten_teksten_NEW	Lkr-vrl	Sc_5_1/13_dich	gemiddelde schaalscore
	soorten_doelen_NEW	Lkr-vrl	Sc_6:1/5_dich	gemiddelde schaalscore
	soorten_publiek_NEW	Lkr-vrl	Sc_7:a-e_dich	gemiddelde schaalscore

Tabel 5.18: Overzicht schalen, indexen en ingedikte variabelen t.b.v. meerniveau-analyses (vervolg)

Klasniveau				
Domeinspecifieke kenmerken				
Didactiek	methode	Lkr-vrl	2_1	6 methoden + anders
	andere leermiddelen	Lkr-vrl	2_2	ja/nee
	zelf ontwikkeld materiaal	Lkr-vrl	2_3	ja/nee
Beoordeling	oordeel obv aspecten tekst (Sc_15_2)	Lkr-vrl	Sc_15_2_dich	1=nooit, soms <i>versus</i> 2=meestal, vaak
	oordeel obv rubrics (Sc_15_3)	Lkr-vrl	Sc_15_3_dich	1=nooit, soms <i>versus</i> 2=meestal, vaak
Differentiatie	differentiatie_dich	Lkr-vrl	18	0=nee, 1=ja
Leerkrachtniveau				
Achtergrond				
	geslacht	Lkr-vrl	1_1	man/vrouw
	leeftijd	Lkr-vrl	1_2	20-29;30-39;40-49;50-59;>60
	vooropleiding	Lkr-vrl	1_3	pabo, verkort, additioneel
werkervaring	werkervaring als lkr	Lkr-vrl	1_4	0-5;6-10;11-15;16-20;>21
Domeinspecifiek				
Professionalisering	professionalisering	Lkr-vrl	1_6	ja/nee
	benodigde kennis en vaardigheden	Lkr-vrl	1_7	nvt en een beetje van toepassing <i>versus</i> (helemaal) van toepassing)
Schoolniveau				
Domeinspecifiek				
Doorlopende leerlijn	doorlopende leerlijn	Sch-vrl	Sc_3_1	niet en beetje van toepassing <i>versus</i> (helemaal) van toepassing (n.b.: helemaal niet van toepassing komt niet voor)

Tabel 5.18: Overzicht schalen, indexen en ingedikte variabelen t.b.v. meerniveau-analyses (vervolg)

Schoolniveau				
Domeinspecifiek				
Schoolbrede focus op schrijfonderwijs	monitoren schrijfontwikkeling	Sch-vrl	Sc_3_10	(helemaal) niet van toepassing <i>versus</i> een beetje en (helemaal) van toepassing=1
	brede focus	Sch-vrl	Focusschool_brede_facilitering ⁶⁶	somscore (range: 1-35)
Aanwezigheid taalbeleid	expliciet taalbeleidsplan	Sch-vrl	1	niet, bezig met en niet op papier <i>versus</i> expliciet taalbeleidsplan
	aanwezigheid taalcoördinator	Sch-vrl	0	deels afgeleid uit vraag 0, ontbrekende informatie via internet: Ja/nee
Algemene kenmerken				
%leerlinggewicht op schoolniveau		DUO	gewicht_cat	S1-S4
aantal leerlingen op school		DUO	Grootte	aantal leerlingen
stedelijkheid		DUO	Stedelijkheid	5 categorieën
denominatie		DUO	den_def	4 categorieën

⁶⁶ Zie ook Tabel 5.17 voor opbouw schaal (3_3, 3_4, 3_5, 3_6, 3_7, 3_8 en 3_9)

Hoofdstuk 6. Scholen, leerkrachten, leerlingen en onderwijs- leerproces

Naast het vaststellen van de schrijfvaardigheden van leerlingen in groep 8 en schoolverlaters in het sbo heeft de peiling tot doel het onderwijsleerproces op de scholen in kaart te brengen. Zowel voor de directie als voor leerkrachten zijn vragenlijsten opgesteld die informatie hieromtrent op respectievelijk school- en klasniveau verzamelen. Met de leerkrachtvragenlijst zijn tevens achtergrondkenmerken van de leerkrachten vastgelegd. Daarnaast is een vragenlijst bij leerlingen afgenomen. In principe is deze vragenlijst, als eerste onderdeel van de afname, door alle aanwezige leerlingen ingevuld. Na afloop van de hoofdafname is het merendeel van de scholen bezocht voor een verdiepend onderzoek, waarin observaties en interviews (met leerkrachten en leerlingen) zijn uitgevoerd.

Met behulp van de resultaten van de school-, leerkracht en leerlingen worden de kenmerken van de deelnemende scholen, leerkrachten/groepen en leerlingen beschreven. Gestart wordt met een paragraaf waarin wordt beschreven hoe we te werk zijn gegaan bij de verzameling van informatie over deze kenmerken en bij hoeveel scholen, groepen en leerlingen data zijn verzameld. In paragraaf 6.2 tot en met 6.5 staan kenmerken van de scholen, de leerkrachten en het onderwijsleerproces centraal, terwijl dit in paragraaf 6.6 voor de leerlingen het geval is. Bij de presentatie van de onderzoeksbevindingen worden naast resultaten van de totale onderzoekspopulatie veelvuldig resultaten op basis van groepen bo-scholen weergegeven. Daarbij komen eventuele verschillen aan bod tussen scholen met 0% gewichtenleerlingen (de S1-scholen), scholen met 0-10% gewichtenleerlingen (de S2-scholen), scholen met 10-25% gewichtenleerlingen (de S3-scholen) en scholen met 25% of meer gewichtenleerlingen (de S4 scholen). Op deze manier wordt recht gedaan aan mogelijke verschillen als gevolg van het percentage gewichtenleerlingen op een school. In dit hoofdstuk worden resultaten op school-/klas- en leerkrachtniveau op basis van ongewogen analyses gepresenteerd. Verder worden bo en sbo weliswaar beschrijvend met elkaar vergeleken, maar de verschillen tussen beide onderwijssoorten worden niet getoetst op significantie.

6.1 Gebruikte informatiebronnen: dataverzameling en respons

Met behulp van verschillende instrumenten is informatie verzameld over de deelnemende scholen, leerkrachten en leerlingen en over het onderwijsleerproces op de scholen.

Schoolkenmerken

Voor het bevragen van kenmerken op schoolniveau is een digitale vragenlijst met behulp van Qualtrics ontworpen. De uitnodiging voor deze vragenlijst is aan de directie van de school verstuurd, met het verzoek de vragen zelf te beantwoorden of de beantwoording over te laten aan de taalcoördinator. In totaal hebben 91 unieke scholen de vragenlijst ingevuld. Daarmee is er voor 68% (n=91) van de 133 deelnemende scholen informatie op schoolniveau beschikbaar. Hiervan zijn 67 scholen voor het basisonderwijs (74%) en 24 scholen voor het speciaal basisonderwijs (26%).⁶⁷ De consequentie is dat de tabellen over het schoolproces in dit hoofdstuk niet optellen tot het totaal aantal van 133

⁶⁷ Zie hoofdstuk 5, paragraaf 5.5, voor meer informatie over de respons op de schoolvragenlijst.

deelnemende bo- en sbo-scholen.⁶⁸ In totaal hebben 53 (adjunct-)directeuren/locatieleiders de schoolvragenlijst ingevuld (55%), 11 taalcoördinatoren/specialisten (11%), 13 leerkrachten en 19 ib'ers/bouw-coördinatoren.

De schoolvragenlijst bevatte drie delen: de aanwezigheid en invulling van taalbeleid op de school, het schrijfonderwijs op de school en de visie op schrijfonderwijs.

Leerkrachtkenmerken en informatie over het onderwijsleerproces

Bij leerkrachten is op twee manieren informatie verzameld over hun onderwijsaanbod. Allereerst is er een vragenlijst op papier afgenomen. Deze vragenlijst bevatte vier onderdelen met vragen over 1) achtergrondkenmerken, 2) onderwijstijd en methoden voor schrijven, 3) de schrijfproducten en 4) het schrijfonderwijs op de school.

De vragenlijst is afgenomen bij leerkrachten van groep 6, 7 en 8 in het reguliere basisonderwijs en bij leerkrachten van de schoolverlaters in het sbo. Voor de verspreiding van deze vragenlijst zijn wij als volgt te werk gegaan. Voorafgaand aan het onderzoek is bij iedere school informatie over het aantal leerlingen en klassen (groepen 8 en groepen schoolverlaters) opgevraagd. In het reguliere basisonderwijs (bo) was het aantal groepen 8 op de school indicatief voor de inschatting van het aantal groepen 6 en 7 op dezelfde school en dus voor de verspreiding van de vragenlijsten. Bij iedere deelnemende groep 8 of schoolverlaters is de leerkracht gevraagd de vragenlijst op de dag van afname van het onderzoek in te vullen. Van de 204 benaderde leerkrachten in groep 8 of de schoolverlaters aan wie de vragenlijst is voorgelegd, hebben 193 leerkrachten de lijst ingevuld (95%). Specifiek voor het bo gaat het om 126 leerkrachten (65% van 193 leerkrachten), voor het sbo betreft het 67 leerkrachten (35% van de 193 leerkrachten). Per bo-school is ook van iedere aanwezige groep 6 en 7 één leerkracht gevraagd de vragenlijst op eigen gelegenheid in te vullen en aan de onderzoekers op te sturen in een retourenvelop. In totaal zijn er 458 vragenlijsten verspreid en zijn er 379 ingevuld ontvangen (83%). Omdat dit rapport zich richt op het onderwijsleerproces in groep 8 en haar resultaten zijn voor de beschrijvingen in dit hoofdstuk de antwoorden van de leerkrachten van groep 8 en de schoolverlaters (n=193) meegenomen. In dit hoofdstuk van het technisch rapport laten we de resultaten voor de leerkrachten van de groepen 6 en 7 buiten beschouwing. In bijlage 7 worden de resultaten van deze groepen in tabelvorm gepresenteerd en waar sprake is van verschillen met de leerkrachten van groep 8 worden deze nader geduid. Op die manier ontstaat een vollediger beeld van het onderwijsleerproces in de gehele bovenbouw van de deelnemende bo- scholen.

Daarnaast heeft een verdiepend onderzoek plaatsgevonden in de groepen 8 en in de groepen schoolverlaters, bestaande uit interviews en observaties van de schrijflessen. Gedurende de interviews zijn leerkrachten (nader) bevraagd over hun visie over en aanbod op het gebied van schrijfonderwijs. Deze interviews zijn bij in totaal 135 leerkrachten afgenomen. Hiervan werkten 99 leerkrachten in het bo (op 85 van de 95 scholen (89%)) en 36 leerkrachten in het sbo (op 22 scholen van de 38 participerende scholen (58%)). De interviews werden afgenomen aan de hand van pre-gecodeerde interviewleidraden. Deze bevatten open vragen, waarvan de antwoorden door de interviewer zijn gecodeerd in vooraf gedefinieerde antwoordcategorieën. Opgemerkt moet worden dat de interviewers per vraag *niet* elke antwoordcategorie apart hebben bevraagd. Per vraag was er ruimte om een nadere toelichting in steekwoorden te vermelden, bijvoorbeeld om de categorie 'anders' te beschrijven of om één van de reeds geformuleerde categorieën nader te illustreren. De extra informatie is dus gebaseerd op de ingevulde steekwoorden die genoteerd zijn door de interviewer.

⁶⁸ De aantallen tellen steeds op tot maximaal 67 (bo) en 24 (sbo) scholen.

Indien bij een vraag geen van de antwoordcategorieën is aangekruist, is een negatief antwoord verondersteld (niet van toepassing).

De observaties zijn bij een deel van dezelfde groep leerkrachten uitgevoerd. Hiermee is getracht aan de hand van pre-dominant activity sampling een overzicht te bieden van de aanwezige fasen in de lessen (en hun relatieve duur) en de uitgevoerde leerkrachtactiviteiten daarbinnen (zie paragraaf 1.6.2 voor een nadere beschrijving).

Met behulp van de interviews en observaties is aanvullende en verdiepende informatie ten opzichte van de vragenlijst verzameld.

Leerlingkenmerken

Om zicht te krijgen op de leerlingkenmerken is ook bij de leerlingen op twee manieren informatie verzameld. Bij alle leerlingen is een vragenlijst afgenomen en per deelnemende school in het verdiepende onderzoek zijn twee leerlingen per klas geïnterviewd.

Met behulp van de leerlingvragenlijst is informatie over achtergrond, schrijfervaring en attitude van de leerlingen bevraagd. Bruikbare vragenlijstinformatie is aanwezig voor 3341 leerlingen. Dit is een respons van 99.5% van de leerlingen die daadwerkelijk aan de peiling hebben deelgenomen. Van de 3341 leerlingen zaten 2438 op het reguliere basisonderwijs (bo) en 903 op het speciaal basisonderwijs (sbo).

Naast de afname van een vragenlijst bij alle deelnemende leerlingen uit groep 8 en schoolverlaters, zijn er tijdens de verdiepende studie op iedere bezochte school met twee leerlingen per klas interviews uitgevoerd. De leerlingen zijn door de bezoekende inspecteur (bo) of onderzoeksmedewerker (sbo) willekeurig geselecteerd en individueel geïnterviewd. In totaal zijn er interviewgegevens van 271 leerlingen, namelijk 199 bo-leerlingen (uit 99 klassen/groepen op 85 scholen) en 72 sbo-leerlingen (uit 35 klassen/groepen op 22 scholen).⁶⁹ Er is geen informatie beschikbaar over de identiteit van de geïnterviewde leerlingen. Net als bij de leerkrachtinterviews is bij de leerlinginterviews gebruik gemaakt van pre-gecodeerde interviewleidraden. Per vraag was er ruimte om een nadere toelichting in steekwoorden te vermelden, bijvoorbeeld om de categorie 'anders' te beschrijven of ter illustratie van één van de reeds geformuleerde categorieën.

Informatie verzameld uit de interviews met leerlingen wordt beschreven in de paragraaf over de leerlingkenmerken (paragraaf 6.6) en in de paragraaf over het onderwijsleerproces (6.5), waar het informatie over de door leerlingen ervaren uitgevoerde aspecten van en activiteiten binnen het schrijfonderwijs weergeeft. Ook hier geldt dat ontbrekende informatie is gecodeerd als 'geen antwoord', en dat deze worden geïnterpreteerd als niet van toepassing.

Weergave resultaten in dit hoofdstuk

Vanwege het aanvullende karakter van de interviews en de lesobservaties bij de vragenlijsten worden de resultaten van beide instrumenten in paragrafen 6.3 tot en met 6.6 waar mogelijk in onderlinge relatie met elkaar beschreven. Daarbij is gekozen voor een thematische aanpak. In de in deze paragrafen gepresenteerde tabellen en teksten wordt duidelijk weergegeven van welke bron de informatie afkomstig is. Wanneer de informatie afkomstig is uit de interviews met leerlingen/leerkrachten of van de lesobservaties, wordt dit standaard genoemd. De tabellen die geen

⁶⁹ Voor één klas geldt dat er maar voor 1 leerling bruikbare data is. Verder zijn in zowel één bo klas als in één sbo klas 4 leerlingen in plaats van 2 leerlingen gesproken.

toelichtingen bevatten zijn gebaseerd op het bestand met de resultaten van de leerkrachtvragenlijst.⁷⁰ Tabel 6.0 geeft een overzicht van de beschikbare informatie per bron en per schooltype.

Tabel 6.0: Overzicht beschikbare informatie per bron en schooltype

	Bo (N=95 scholen)			Sbo (N=38 scholen)		
	n scholen	% scholen	Totaal Aantal respondenten	n scholen	% scholen	Totaal Aantal respondenten
Schoolvragenlijst	67	74%	67	24	63%	24
Leerkracht vragenlijst*	95	100%	126	36	95%	67
Leerkracht interview*	85	89%	99	22	58%	36
Lesobservaties ⁷¹	58	61%	70	21	55%	30
Leerling vragenlijst	95	100%	2438	38	100%	903
Leerling interview**	85	91%	199	22	58%	72

* max. één leerkracht per klas

** twee leerlingen per klas

De lesobservaties en de interpretatie van deze resultaten

In het bo zijn 70 schrijflessen (70 leerkrachten op 58 scholen) en in het sbo 30 schrijflessen (30 leerkrachten op 21 scholen) geobserveerd met behulp van Predominant activity sampling (Van Gelderen & Blok, 1991). Observatoren, in het bo inspecteurs in de rol van onderzoeker en in het sbo studenten orthopedagogiek, observeerden gedurende 20 seconden de leerkracht en de les en scoorden gedurende de 20 seconden daarna de meest voorkomende leerkrachtactiviteit gedurende de observatieperiode, waarna weer een volgende observatieperiode van 20 seconden observatie en 20 seconden scoring volgde, enzovoorts. De kern van het observatie-instrument richtte zich op de geobserveerde fasen ('globale interactie episodens tussen leerkracht en leerlingen') in de schrijfles en op de activiteiten die de leerkracht daarin uitvoerde. De observator stelde per observatie vast in welke fase de les zich bevond en bepaalde welke activiteit dominant was gedurende de 20 seconden waarin is geobserveerd. Observatoren hebben per tijdsblok tevens informatie genoteerd over algemene aspecten van de les en het type opdracht waar leerlingen aan werkten. In totaal zijn er 4849 geobserveerde tijdsblokken van 20 seconden in het bo en 1753 van deze tijdsblokken in het sbo gescoord.

⁷⁰ Ook aan de N in de meeste tabellen is zichtbaar van welke bron de informatie afkomstig is.

⁷¹ Over de lesobservaties dient het volgende te worden opgemerkt. Oorspronkelijk waren er voor 79 bo-lessen (op 67 bo-scholen, 71%) observatiedata aanwezig. Helaas bleek het percentage ontbrekende waarden in sommige lessen relatief groot. Uitgegaan is van het inclusie criterium dat minimaal 85% van de te observeren lesmomenten aanwezig moest zijn. Dit heeft geresulteerd in het verwijderen van 9 lessen, die niet meegenomen zijn in de analyses.

In de paragrafen 6.3 tot en met 6.6 worden de resultaten van de vragenlijsten en interviews waar relevant en mogelijk aangevuld met resultaten afkomstig uit de lesobservaties. Wij willen benadrukken dat de lesobservaties slechts één keer hebben plaatsgevonden. De deelnemende leerkrachten is gevraagd een reguliere les te geven, waardoor bovendien de mogelijkheid bestaat dat er een les binnen een lessenserie is geobserveerd, hetgeen de representativiteit van de observatieresultaten mogelijk heeft beïnvloed. In het bo zijn de observaties door onderwijsinspecteurs uitgevoerd. Ook al was deze inspecteur niet in zijn/haar toezichthoudende rol aanwezig en waren scholen hierover ook expliciet geïnformeerd, beide aspecten (eenmaligheid van de observatie en de observator zelf) kunnen van invloed zijn op de representativiteit en generaliseerbaarheid van de gevonden resultaten. De uitkomsten van de lesobservaties dienen daarom met grote voorzichtigheid te worden geïnterpreteerd. Dit geldt zeker voor situaties waarin onderdelen uit de eenmalig geobserveerde les worden gerelateerd aan de door leerlingen en leerkrachten gegeven antwoorden over het reguliere handelen van de leerkrachten en de uitgevoerde activiteiten door leerlingen.

6.2 Kenmerken van de deelnemende scholen

6.2.1 Organisatie van taal/schrijven op school

De manier waarop scholen het schrijfonderwijs organiseren en inbedden in het taalonderwijs verschilt. Dit geldt ook voor de mate waarin taal-/schrijfbeleid op een school is geformaliseerd en waarin schrijven aandacht krijgt op een school. Een taal-/leescoördinator kan daarbij een trekkersrol op zich nemen. Daarbij kan een taalcoördinator werken als een katalysator die het taaldomein, en daarbinnen de subdomeinen, blijvend onder de aandacht van het team kan brengen. De vragenlijst bevatte geen expliciete vraag naar de aanwezigheid van een taalcoördinator. Wel kon de respondent invullen taalcoördinator te zijn. Dit was het geval bij 15 scholen het geval.⁷² Daarna zijn de websites van de overige 118 scholen doorzocht op teamsamenstelling en zijn de schoolgidsen doorgenomen op aanwezigheid van taalcoördinator of taalspecialist én leescoördinator of leesspecialist (ook dyslexiespecialist). In het bo blijkt 31% van de deelnemende scholen over een taal-/leescoördinator te beschikken en in het sbo geldt dit voor 21% van de scholen. Dit betekent dus dat de meerderheid van de scholen (respectievelijk 70% en 78%) niet over een dergelijke coördinator beschikt (zie Tabel 6.1). De expertise van een taalcoördinator kan tevens ingezet worden om te komen tot (meer) weloverwogen (en geformaliseerd) taalbeleid, ook al is een taalcoördinator hiervoor niet noodzakelijk. Een veel gehanteerde definitie van taalbeleid is dat het een 'structurele en strategische poging van een schoolteam [is] om de onderwijspraktijk aan te passen aan de taalleerbehoeften van de leerlingen, met het oog op het bevorderen van hun algemene ontwikkeling en het verbeteren van hun onderwijsresultaten' (<https://www.slo.nl/thema/vakspecifiekethema/nederlands/taalbeleid/>). Deze gestructureerde benadering is idealiter vastgelegd in een taalbeleidsplan. Op 34% van de deelnemende bo-scholen blijkt een expliciet taalbeleidsplan te zijn opgesteld, terwijl 22% van de bo-scholen geen taalbeleidsplan heeft (en deze ten tijde van de afname ook niet ontwikkeling was). Op

⁷² In vier van de 15 gevallen is de functie als taal-/leescoördinator afgeleid van de additionele, doordat de respondent in de toelichtingsruimte een omschrijving van zijn/haar taak gaf.

de sbo-scholen is zowel het percentage scholen dat helemaal geen taalbeleidsplan heeft (8%) als het percentage met een expliciet taalbeleidsplan (21%) lager dan in het bo.

Tabel 6.1: Aantallen scholen met een expliciet taalbeleidsplan en waar taalcoördinator aanwezig is

		bo (N=67)		sbo (N=24)	
		N	%	n	%
Helemaal niet		15	22%	2	8%
Wij zijn bezig met ontwikkeling van taalbeleid		16	24%	7	29%
Er is taalbeleid, maar het staat nog niet op papier		13	19%	10	42%
We hebben een expliciet taalbeleidsplan		23	34%	5	21%
		bo (N=95)		sbo (n=38)	
Aanwezigheid taalcoördinator	Ja	29	31%	8	21%
	Nee	66	70%	30	78%

6.2.2 Aandachtspunten in het taalbeleid

SLO stelt dat taalbeleid smal, met een focus op schoolse taalvaardigheid (zoals spelling, grammatica), of breed kan zijn. Breed taalbeleid veronderstelt dat er binnen het totale curriculum aandacht is voor taalonderwijs, en dat taalonderwijs geïntegreerd is met andere vakken, hetgeen tevens een hogere mate van samenwerking tussen leerkrachten veronderstelt. Directeuren is gevraagd naar schoolbeleid gericht op taal en meer specifiek gericht op schrijfonderwijs. Opmerkelijk is dat de vraag naar de invulling van het taalbeleid relatief veel ontbrekende waarden bevatte. Vierentwintig respondenten (26%), die wel de vragenlijst invulden, hebben geen enkel item binnen de vraag naar (algemeen) taalbeleid ingevuld. Dit betreft met name de bo-respondenten. Op basis van een factoranalyse zijn twee schalen vastgesteld, namelijk *Taalbeleid gericht op het wegwerken van taalachterstanden bij (anderstalige) leerlingen* en *Taalbeleid gericht op rijk/breed taalaanbod voor allen*⁷³. Uitgaande van een maximumscore van respectievelijk 25 en 30, kan uit Tabel 6.2 worden afgeleid dat de scholen die deze vragen beantwoorden, relatief hoog scoren op beide schalen. Ze geven daarmee aan dat de meeste items voor hen van toepassing zijn. Hierbij verschillen de bo-scholen en sbo-scholen niet van elkaar. De vraag naar de schoolbrede focus in het schrijfonderwijs kan worden samengevat aan de hand van twee schalen. De schaal *Doorgaande leerlijnen en monitoring* heeft een maximale score van 15⁷⁴. De gemiddelde scores binnen bo en sbo laten zien dat scholen deze aspecten weliswaar schoolbreed benadrukken, maar dat zij ze niet volledig op hen van toepassing achten. Een schoolbrede nadruk op *Facilitering van het schrijfonderwijs in brede zin*⁷⁵ is weliswaar van toepassing, maar – uitgaande van een maximale score van 35 – geeft een gemiddelde score van 21.02 (bo) en 18.52 (sbo) weer dat scholen zich niet herkennen in een sterke toepassing van aspecten als het professionaliseren van leerkrachten op het gebied van schrijfonderwijs, het ontwikkelen van een eigen schoolvisie op

⁷³ Deze schalen hebben respectievelijk een betrouwbaarheid van Cronbach's α 0.77 en 0.75

⁷⁴ De betrouwbaarheid van Doorgaande leerlijnen en monitoring is α =0.66

⁷⁵ De betrouwbaarheid is Facilitering van het schrijfonderwijs in brede zin is α =0.88

schrijfvaardigheid en –didactiek of het faciliteren van collegiaal overleg over de aanpak van het schrijfonderwijs.

Tabel 6.2: Focus in het taalbeleid en schrijfonderwijs op schoolniveau

		N	M	SD	Min	Max
Focus op wegwerken achterstanden (anderstalige) leerlingen	bo	48	17.54	4.07	8	24
	sbo	21	17.95	3.51	11	25
Focus op rijk taalaanbod voor alle leerlingen	bo	49	25.49	3.18	15	30
	sbo	21	25.00	2.70	20	29
Focus schrijfonderwijs: doorgaande leerlijn en monitoring	bo	65	10.34	2.21	6	15
	sbo	23	9.17	1.72	6	13
Focus schrijfonderwijs: faciliteren schrijfonderwijs in brede zin	bo	65	21.02	5.57	10	35
	sbo	23	18.52	4.18	14	27

6.2.3 Visie op het schrijfonderwijs

De laatste vraag in de schoolvragenlijst bestond uit zeven stellingen waarmee de visie van de school op schrijfonderwijs is gemeten. In Tabel 6.3 worden de resultaten per stelling gepresenteerd, waarbij weer een onderscheid wordt gemaakt tussen bo en sbo. De overgrote meerderheid van schoolleiders/taalcoördinatoren in beide schooltypen vindt dat schrijfonderwijs in direct verband moet staan met de andere onderdelen binnen het taalonderwijs, dat het gericht moet zijn op levensechte taken die dichtbij de belevingswereld van de leerlingen liggen en dat het leerlingen bekend moet maken met verschillende soorten teksten. In het bo lijkt men een nadruk op schrijfprocessen relevanter te vinden dan in het sbo (bo: 75%, sbo: 48%), terwijl in het sbo meer scholen dan in het bo het eens zijn met de stelling dat schrijven vooral gericht moet zijn op correct taalgebruik (70% in sbo en 55% in bo).

Tabel 6.3: Visie t.a.v. schrijfonderwijs op schoolniveau

		bo (N=65)		sbo (N=23)	
Schrijfonderwijs is vooral gericht op correct schrijven van woorden en zinnen	Ja	36	55%	16	70%
	Nee	29	45%	7	30%
Schrijfonderwijs is vooral gericht op het bereiken van doelen bij de lezer	Ja	43	66%	13	57%
	Nee	22	34%	10	43%
Schrijfonderwijs moet leerlingen vooral bekend maken met veel verschillende soorten teksten	Ja	53	82%	19	83%
	Nee	12	18%	4	17%
Schrijfonderwijs moet zich vooral richten op de schrijfprocessen van leerlingen (plannen en reviseren)	Ja	49	75%	11	48%
	Nee	16	25%	12	52%
Schrijfonderwijs moet uitgaan van levensechte situaties waarin leerlingen zich kunnen herkennen	Ja	55	85%	20	87%
	Nee	10	15%	3	13%
Schrijfonderwijs moet in direct verband staan met andere onderdelen van het taalonderwijs (lezen, grammatica, woordenschat, spreken etc.)	Ja	63	97%	2	96%
	Nee	2	3%	1	4%
Bij schrijfonderwijs moet veel aandacht zijn voor het schrijven met tablet en/of computer	Ja	30	46%	10	43%
	Nee	35	54%	13	57%

6.3 Leerkrachtkenmerken

Als achtergrondkenmerken zijn met behulp van de vragenlijst, naast individuele kenmerken, tevens de genoten opleiding en werkervaring, genoten professionaliseringsactiviteiten en self-efficacy gemeten. Daarnaast is met behulp van de interviews geïnventariseerd welke aspecten binnen het schrijfonderwijs de leerkrachten lastig vinden en waar zij positief over zijn.

6.3.1 Achtergrondkenmerken van de leerkrachten

De eerste vragen in de leerkrachtvragenlijst betroffen achtergrondinformatie, waarin eveneens naar ervaring is gevraagd. In Tabel 6.4 staan de leeftijd en het geslacht van de deelnemende leerkrachten weergegeven, uitgesplitst naar schooltype (bo of sbo). Zoals uit Tabel 6.4 kan worden afgeleid is in beide schooltypen ongeveer twee derde van de leerkrachten vrouw en een derde man. De bo-leerkrachten zijn over het algemeen wat jonger met 60% in de leeftijd tot 40 jaar ten opzichte van 38% bij de sbo-leerkrachten.

Tabel 6.4: Achtergrondkenmerken deelnemende leerkrachten

		bo			sbo		
		N	n	%	N	n	%
Geslacht	Man	125	42	34%	67	20	30%
	Vrouw		83	66%		47	70%
Leeftijd	20-30 jaar	125	28	22%	66	9	14%
	30-40 jaar		47	38%		16	24%
	40-50 jaar		28	22%		23	35%
	50-60 jaar		17	14%		12	18%
	Ouder dan 60		5	4%		6	9%

Gevraagd naar het aantal jaren werkervaring is in de vragenlijst een onderscheid gemaakt tussen werkervaring in het primair onderwijs en werkervaring in de bovenbouw (groep 6-8). Zoals uit Tabel 6.5 blijkt is het percentage bo-leerkrachten dat maximaal 10 jaar ervaring in het po heeft met 45% hoger dan het percentage sbo-leerkrachten met maximaal 10 jaar ervaring (23%). Met name de groep sbo-leerkrachten met meer dan 21 jaar ervaring is relatief hoog (36%). Als het gaat om werkervaring in de bovenbouw blijkt dat circa een derde in beide groepen leerkrachten hier tot 5 jaar werkervaring heeft (38% in bo versus 30% in sbo). Verreweg het merendeel van de leerkrachten in beide schooltypen heeft de reguliere pabo afgerond (89% in bo en 90% in sbo).

Tabel 6.5: Ervaring en opleiding respondenten

		bo			sbo		
		N	n	%	N	n	%
Werkervaring in het PO	0-5 jaar	122	22	18%	66	8	12%
	6-10 jaar		33	27%		7	11%
	11-15 jaar		25	21%		15	23%
	16-20 jaar		19	16%		12	18%
	Meer dan 21 jaar		23	19%		24	36%
Werkervaring in bovenbouw (groep 6-8)	0-5 jaar	124	47	38%	66	20	30%
	6-10 jaar		30	24%		14	21%
	11-15 jaar		25	20%		11	17%
	16-20 jaar		13	11%		15	23%
	Meer dan 21 jaar		9	7%		6	9%
Behaald diploma	PABO	124	112	89%	65	60	90%
	Verkorte PABO		10	8%		4	6%
	Academische PABO		2	2%		1	2%
	Diploma nog niet behaald		0	0%		0	0%

6.3.2 Professionalisering en self efficacy

Om kennis en vaardigheid in het schrijfonderwijs te vergroten kunnen leerkrachten deelnemen aan professionaliseringsactiviteiten. De vragenlijst bevatte een vraag naar eventueel gevolgde professionaliseringsactiviteiten die gericht zijn op verbetering van het schrijfonderwijs. Deze vraag is door 186 leerkrachten ingevuld. De overgrote meerderheid van beide groepen leerkrachten heeft zich niet nader geprofessionaliseerd op het gebied van schrijfonderwijs (zie Tabel 6.6).⁷⁶ Van de sbo-leerkrachten heeft één leerkracht aangegeven op welke manier hij/zij nader geschoold is. Deze leerkracht heeft scholing gehad ten behoeve van de omgang met de methode. Dit geldt ook voor twee bo-leerkrachten. In het bo worden verder de opleiding tot taalcoördinator (waarbinnen ook aandacht voor schrijven) genoemd en scholing op het gebied van doorgaande leerlijnen. Twee andere bo-leerkrachten zeggen *andere* professionaliseringsactiviteiten op schrijfgebied te hebben doorlopen, zonder specificering.

Tabel 6.6: Professionaliseringactiviteit(en) van de leerkrachten op het gebied van schrijven

	N		n	%
bo	120	Nee	111	92%
		Ja	9	8%
sbo	66	Nee	63	95%
		Ja	3	5%

6.3.3 Eigen inschatting aanwezige kennis en vaardigheden

Zoals uit Tabel 6.6 blijkt is het aantal leerkrachten dat zich na de initiële opleiding nader heeft geprofessionaliseerd op het gebied van schrijfonderwijs zeer beperkt (bo: 8% en sbo: 5%). In de vragenlijst konden leerkrachten ook reageren op de stelling of ze over voldoende kennis en vaardigheden met betrekking tot schrijfonderwijs beschikken. Van de 190 leerkrachten die deze vraag invulden, is, afgezien van één sbo-leerkracht, elke leerkracht van mening dat deze stelling in meer of mindere mate voor hem/haar opgaat (zie Tabel 6.7).

⁷⁶ Daarbij dient te worden vermeld dat het aanbod op het terrein van schrijfonderwijs beperkt is, zeker in vergelijking tot het aanbod voor bijvoorbeeld lezen.

Tabel 6.7: Professionalisering van de leerkrachten op het gebied van schrijven

Self efficacy/zelfwaardering/eigen inschatting		bo			sbo		
		N	n	%	N	n	%
Beschikking hebben over voldoende kennis en aardigheden m.b.t. schrijfonderwijs	Helemaal niet van toepassing	123	0	0%	67	0	0%
	Niet van toepassing		0	0%		1	2%
	Een beetje van toepassing		31	25%		14	21%
	Van toepassing		81	66%		46	69%
	Helemaal van toepassing		11	9%		6	9%

6.3.4 Sterke punten en lastige aspecten in het schrijfonderwijs

Sterke punten

Aan het eind van de *interviews* is leerkrachten gevraagd te vertellen over wat zij sterke kanten van hun schrijfonderwijs vonden, wat zij als lastige aspecten van schrijfonderwijs zagen en in te gaan op redenen waarom die aspecten lastig zijn. De reacties van leerkrachten zijn in steekwoorden door de interviewer genoteerd. Hiervoor waren op voorhand geen categorieën vastgesteld. De genoteerde steekwoorden zijn door de onderzoekers thematisch gecategoriseerd. Bij de lastige aspecten zijn tevens de genoemde redenen gecategoriseerd. De meest genoemde aspecten staan weergegeven in Tabellen 6.8 en 6.9. De 135 leerkrachten die zijn geïnterviewd en die zijn ingegaan op sterke kanten van hun schrijfonderwijs zeggen met name er goed in te zijn de lessen en opdrachten aantrekkelijk te maken voor leerlingen. Een derde van de sbo- en bo-leerkrachten noemt dat zij de leerlingen kunnen motiveren en enthousiasmeren voor schrijven/de opdracht. Ook vinden zij dat het hen lukt opdrachten te laten aansluiten bij de belevingswereld van leerlingen en/of leerlingen mee te laten denken over de opdracht. Met name sbo-leerkrachten (25%) noemen dat zij oog hebben voor diversiteit in de klas en hun schrijfonderwijs aan laten sluiten bij verschillende behoeften van leerlingen, en dat zij een duidelijke, vooral gestructureerde instructie kunnen geven. Een sbo-leerkracht noemt in dit kader dat hij/zij *'dichtbij het kind (en zijn/haar belevingswereld) wil blijven en kijkt naar wat leerlingen nodig hebben'*. Daarbij wordt expliciet benoemd dat de leerlingen vertrouwen moeten krijgen. Deze focus op vertrouwen, complimenten, hulp en redzaamheid komt bij sommige sbo-leerkrachten terug. Ook enkele bo-leerkrachten *'proberen de drempel [voor leerlingen] laag te houden'*. Bo-leerkrachten noemen ook hun affiniteit met schrijven en hun eigen vaardigheid in schrijven of taalverzorging⁷⁷ als sterk punt van hun schrijfonderwijs.

⁷⁷ Onder taalverzorging, in de Tabel grammatica/interpunctie, vallen opmerkingen over het eigen vaardigheidsniveau, maar ook over affiniteit met deze technische taalaspecten.

Tabel 6.8: Genoemde sterke aspecten in de schrijflessen (interviews leerkrachten)

	bo		sbo	
	n	%	n	%
Motiveren/enthousiasmeren	37	37%	11	31%
Schrijfonderwijs sluit aan bij belevingswereld en/of leerlingen hebben bijdrage in invulling schrijfles/-opdrachten	21	21%	5	14%
Creativiteit stimuleren en creatieve lessen	13	13%	2	6%
Duidelijke doelen/eisen/afspraken	16	16%	6	17%
Differentiatie/aansluiting bij verschillende behoeften leerlingen	10	10%	9	25%
Duidelijke/goede instructie	12	12%	9	25%
Begeleiding schrijfproces	14	14%	1	3%
Feedback/terugkoppeling aan leerlingen	7	7%	3	8%
Schrijfopdrachten (b.v. variatie in opdrachten)	11	11%	3	8%
Affiniteit met schrijven zelf/geven van schrijflessen	9	9%	1	3%
Opbouw van de tekst	7	7%	0	0%
Grammatica/interpunctie	5	5%	0	0%
Eigen schrijf-beheersing	3	3%	0	0%
Eigen pedagogische vaardigheid	3	3%	1	3%

NB: Dit betrof een open vraag. Leerkrachten hebben regelmatig verschillende aspecten naar voren gebracht. Meerdere antwoorden waren dus mogelijk.

Lastige aspecten

Niet alle leerkrachten hebben gesproken over aspecten die het schrijfonderwijs voor hen bemoeilijken: 117 leerkrachten (87% van de geïnterviewde leerkrachten) zijn op dit onderdeel ingegaan. Hiervan zijn 81 bo-leerkrachten (82% van de geïnterviewde bo-leerkrachten) en 36 sbo-leerkrachten (100%). In Tabel 6.9 staan de uitkomsten gepresenteerd. Opmerkelijk is dat bij de sterke en lastige aspecten relatief veel dezelfde aspecten worden genoemd: wat de ene leerkracht een sterk aspect van zijn/haar schrijfonderwijs vindt, blijkt voor de ander juist lastig uit te voeren. Voor het meest genoemde lastige aspect, het beoordelen/evalueren en monitoren van de schrijfontwikkeling van leerlingen, geldt dit niet. Ruim een derde van de bo-leerkrachten benoemt dit onderdeel van het schrijfonderwijs (maar niemand heeft dit als positief aspect genoemd), terwijl dit in het sbo door 19% van de leerkrachten naar voren is gebracht. Daarbij noemen sommige leerkrachten (bo: n=9, sbo: n=3) dat het voor hen niet duidelijk is naar welke aspecten van een leerlingtekst zij moeten kijken. Meermaals wordt daarbij genoemd dat *'een instrumentarium ontbreekt'* en dat de beoordeling daarom subjectief wordt. Een ander aspect dat onduidelijk is, is dat sommige leerkrachten geen helder beeld hebben van het gewenste schrijfniveau (zoals *'[het is] onduidelijk waar ze horen te zitten'*, *'haal ik er voldoende uit? is de opbrengst hoog genoeg?'* en *'de ontwikkeling van de leerlingen nagaan, ik weet niet echt wat ervan verwacht wordt'*). Zeven leerkrachten in het bo en drie in het sbo refereren aan dit punt. Met name sommige sbo-leerkrachten (n=6) noemen het lastig om in te schatten wat hun leerlingen kunnen. Hun verwachtingen van leerlingteksten en wensen voor het niveau zijn soms te hoog, maar te laag komt ook voor (*'Soms onderschat je de leerlingen, terwijl ze dan wel iets moois neerzetten'*). De groep leerkrachten die refereert aan verwachtingen problematiseert niet de onduidelijkheid van beoogd

niveau of doelen, maar refereert met name aan de eigen verwachtingen in relatie tot wat ze willen bereiken of de juistheid daarvan.

Een vijfde van de bo-leerkrachten vindt het lastig om te differentiëren en in te spelen op de niveauverschillen in de klas. Leerkrachten benadrukken daarbij verschillende problemen, zoals de grootte van de klas, de beperkte handreikingen binnen de methode, het niet weten hoe te differentiëren bij het beoordelen of het niet goed weten wat de sterke leerlingen nodig hebben. In beide schooltypen wordt verder regelmatig genoemd dat leerkrachten het lastig vinden dat de doorgaande leerlijn of opbouw in schrijflessen (over het leerjaar) onvoldoende is uitgewerkt en/of dat ze hier onvoldoende kennis over hebben. Ook het motiveren van *alle* leerlingen voor schrijven en het uitlokken van creativiteit wordt regelmatig genoemd als lastig onderdeel.

De leerkrachten is tevens gevraagd toe te lichten waarom zij de door hun genoemde aspecten lastig vinden. Tien sbo-leerkrachten refereren hierbij aan de leerlingpopulatie en drie aan de motoriek van sbo-leerlingen, die het schrijven *an sich* in de weg lijkt te staan. Zeven bo- en vier sbo-leerkrachten noemen de methode als (een van de) reden(en) waarom het de leerkracht niet goed lukt om de creativiteit van leerlingen te activeren (*'saaie opdrachten'*). Het gebrek aan tijd wordt regelmatig genoemd als reden waarom beoordeling/evaluatie of de doorgaande leerlijn lastige aspecten zijn (bo: 14, sbo: 9): er is geen tijd om zich hierin te verdiepen. Ook wordt tijdgebrek in combinatie met het niet geven van feedback of lezen van de opdrachten genoemd (*'schrijven kost veel tijd'*). Het gebrek aan prioritering wordt gekoppeld aan dezelfde onderwerpen (evaluatie, doorgaande leerlijn), maar de moeilijkheid om goede instructie te geven wordt ook aan prioritering gekoppeld. Zoals een interviewer heeft genoteerd: *'We doen het te weinig, de kernvakken gaan voor'*. Gebrek aan prioritering wordt door elf bo-leerkrachten en 3 sbo-leerkrachten benoemd.

Tabel 6.9: Genoemde lastige aspecten van schrijflessen (interviews leerkrachten)

	bo (N=81)		sbo (N=36)	
	n	%	n	%
Beoordeling/monitoring/evaluatie schrijfontwikkeling leerlingen	29	36%	7	19%
Differentiatie/aansluiting bij niveau leerlingen	16	20%	2	6%
Doorgaande leerlijn/opbouw schrijflessen	10	12%	4	11%
Verwachtingen schrijfvaardigheid (te hoog of te laag)	1	1%	5	14%
Motiveren van <i>alle</i> leerlingen en creativiteit stimuleren	10	12%	2	6%
Begeleiden van product en/of proces	14	17%	4	11%
Schrijfopdrachten (geschikte/goede opdrachten en/of variatie in opdrachten, aandacht voor specifieke teksten)	3	4%	3	8%
Gebruik feedback	5	6%	1	3%
Spelling/grammatica/interpunctie	3	4%	0	0%
Eigen schrijfbeheersing	6	7%	1	3%
Gebrek aan visie (eigen/van school)	3	4%	0	0%

Tabel 6.9: Genoemde lastige aspecten van schrijflessen (interviews leerkrachten, vervolg)

	n	%	n	%
<i>Aangevoerde redenen:</i>				
Tijdgebrek/schrijven is tijdrovend vak	14	17%	9	25%
Geen prioriteit/andere prioriteiten	11	14%	3	8%
Methode	7	9%	4	11%
Motoriek leerlingen	0	0%	3	8%
Leerlingpopulatie	2	2%	10	28%

NB: Dit betrof een open vraag. Leerkrachten hebben regelmatig verschillende aspecten naar voren gebracht. Meerdere antwoorden waren dus mogelijk.

6.4 Visie en afstemming

6.4.1 Visie op schrijfonderwijs

Hetgeen leerkrachten belangrijk vinden binnen schrijfonderwijs zal de voorbereiding en uitvoering van hun schrijflessen beïnvloeden. Datzelfde geldt voor de doelen die zij door middel van hun schrijflessen willen bereiken. De visie op schrijfonderwijs is daarom in het interview bevestigd door leerkrachten de volgende vraag voor te leggen: *Wat is volgens u goed schrijfonderwijs?* De antwoorden heeft de interviewer vertaald naar één of meerdere van de volgende vier vooraf gedefinieerde hoofdvisies: 1) de communicatieve, 2) de conceptuele, 3) de expressieve functie van het schrijfonderwijs en/of 4) het schrijfonderwijs als middel om goed te leren spellen en correcte zinnen te leren maken (zie Tabel 6.10). Daarnaast heeft de interviewer genoteerd welke centrale begrippen (of beschrijvingen daarvan) voorkwamen in de antwoorden van de leerkrachten (zie Tabel 6.11). Merk op dat de interviewer de spontane antwoorden heeft omgezet naar van tevoren vastgestelde categorieën en begrippen en de aanwezigheid daarvan heeft genoteerd. Specifieke bevestiging van alle verschillende hoofdvisies en begrippen heeft dus *niet* plaatsgevonden. Door deze werkwijze is getracht sociaal wenselijke antwoorden te verminderen. De spontaan gegeven antwoorden worden verondersteld de door de leerkracht meest relevant geachte aspecten binnen schrijfonderwijs weer te geven.

In totaal hebben de interviewers voor 28 leerkrachten, 25 bo-leerkrachten en drie sbo-leerkrachten, geen hoofdcategorie ingevuld, terwijl er wel antwoorden waren aangevinkt bij de centrale begrippen. Wanneer we kijken naar de vier onderscheiden hoofdvisies, zie Tabel 6.10, dan blijken leerkrachten in bo en sbo met name de communicatieve en de expressieve functie van het schrijven belangrijk te vinden. Dit wil zeggen dat leerkrachten het belangrijk vinden dat leerlingen leren hun boodschap goed over te brengen. Daarnaast vinden ze het relevant dat leerlingen met behulp van een tekst uiting leren geven aan hun eigen gevoelens en ideeën en dat leerlingen in de schrijflessen ruimte krijgen om creatief te schrijven. Opmerkelijk is dat bijna een derde van de bo-leerkrachten schrijven (ook) ziet als een middel om correcte teksten te produceren, terwijl dat in het sbo veel minder vaak wordt genoemd. Overigens noemt de helft van de leerkrachten meer dan één functie. Met name combinaties van de communicatieve, expressieve en/of correctheidsfuncties worden gevonden.

Tabel 6.10: Genoemde visies in termen van vier hoofdcategorieën (interviews leerkrachten)

	bo N=74		sbo N=33	
	n	%	n	%
Communicatieve functie	40	54%	17	52%
Conceptuele functie	8	11%	6	18%
Expressieve functie	38	51%	13	39%
Middel om goed te leren spellen/correcte zinnen te bouwen (correctheid)	29	39%	4	12%

Op voorhand waren diverse centrale begrippen bepaald waar leerkrachten aan zouden kunnen refereren om hun ideeën over goed schrijfonderwijs te beschrijven of toe te lichten. In Tabel 6.11 staat

het percentage leerkrachten weergegeven dat inging op deze centrale begrippen, uitgesplitst naar schooltype. Voor vier leerkrachten, twee bo-leerkrachten en twee sbo-leerkrachten, zijn geen centrale begrippen genoteerd. Voor deze leerkrachten was wel een hoofdvisie ingevuld. Respectievelijk 16% en 6% van de bo- en sbo-leerkrachten refereerde in hun antwoord aan doelen die door de methode (dus extern) worden aangereikt. Aan de overige twee extern aangereikte doelen die vooraf waren gedefinieerd, te weten de kerndoelen en het referentiekader, werd in het sbo niet gerefereerd en in het bo door 8% van de leerkrachten. Over het algemeen bleken de bo-leerkrachten vaker te refereren aan domeinspecifieke kennis of vaardigheden (nr. 3-6) dan de sbo-leerkrachten. Dit geldt ook voor de bevordering van motivatie voor het schrijven en van schrijfplezier (nr. 7). Gekeken naar de aanwezigheid van begrippen die ingaan op het daadwerkelijke schrijven van de leerlingen (nr. 8-10) blijkt dat met name correctheid vaak wordt genoemd (bo: 42% en sbo: 24%). Relatief veel leerkrachten wijzen op het belang van een goede schrijfvaardigheid voor de toekomstige schoolloopbaan (met name bo) of het toekomstig communicatief functioneren in onze maatschappij (met name sbo). Ongeveer een vijfde van de leerkrachten in bo en sbo spreekt over onderwerpen die vooraf niet gedefinieerd staan (*anders*). De steekwoorden laten zien dat sommige leerkrachten het belangrijk vinden dat leerlingen netjes schrijven, dat ze hun fantasie gebruiken of dat ze hun gedachten begrijpelijk op papier kunnen zetten. Een enkeling refereert aan de structuur of opbouw van teksten.

Tabel 6.11: Genoemde centrale begrippen (interviews leerkrachten)

	bo N=97		sbo N=34	
	n	%	N	%
1. Referentiedoelen/kerndoelen (extern aangereikte doelen)	8	8%	0	0%
2. Methodedoelen (extern aangereikte doelen)	16	16%	2	6%
3. Doelgericht schrijven	47	48%	9	26%
4. Publiekgericht schrijven	22	23%	6	18%
5. Kennis over verschillende tekstsoorten	39	40%	7	21%
6. Authentieke schrijftaken/echt publiek	11	11%	3	9%
7. Motivatie/schrijfplezier	35	36%	5	15%
8. Schrijfgemak (<i>fluency</i>)	11	11%	1	3%
9. Correct taalgebruik	41	42%	8	24%
10. Schrijfstrategieën/schrijfaanpak	18	19%	4	12%
11. Schrijven met tablet/computer	5	5%	1	3%
12. Samenhang met leesvaardigheid	5	5%	4	12%
13. Transfer naar andere leergebieden (b.v. wereldoriëntatie)	8	8%	3	9%
14. Belang schrijfvaardigheid voor vervolgonderwijs	20	21%	4	12%
15. Belang schrijfvaardigheid voor toekomst leerlingen (communicatieve vaardigheid)	5	5%	12	35%
Anders	20	21%	8	22%

NB: meerdere antwoorden mogelijk.

Visie in de praktijk

Leerkrachten is vervolgens gevraagd hoe zij hun genoemde visie op schrijfonderwijs in de praktijk brengen tijdens hun schrijflessen. Voor één bo-leerkracht en voor twee sbo-leerkrachten is voor deze vraag niets ingevuld. De resultaten staan weergegeven in Tabel 6.12. Ook bij deze vraag refereerden bo-leerkrachten in hun spontane antwoorden op de open vraag over het algemeen vaker aan de op voorhand gedefinieerde categorieën dan de sbo-leerkrachten. Bo-leerkrachten zeggen met name dat

hun visie tot uitdrukking komt in de keuze voor de leerdoelen, in de aard van de schrijftaken en in de inhoud van de schrijftaken. Bij de keuze voor leerdoelen wordt daarbij meermaals gerefereerd aan het gebruik van methode- en landelijke doelen, het vaststellen en/of benoemen van lesdoelen en de doorgaande leerlijn. Bij de vertaling van de visie in de *inhoud van de schrijftaken* gaat het vooral om het type opdrachten dat leerkrachten aan leerlingen voorleggen. Zo refereren leerkrachten (bo en sbo) aan opdrachten die aansluiten bij de belevingswereld van leerlingen, op betekenisvolle en actuele onderwerpen en op opdrachten die de creativiteit van leerlingen stimuleren. Wanneer de leerkracht ingaat op de soorten teksten die leerlingen moeten schrijven (*de aard van de schrijftaken*) wordt het bewust inzetten van diverse tekstsoorten genoemd, zoals het schrijven van gedichten, handleidingen, recepten, klachten, verhalen. In het sbo wordt gerefereerd aan teksten die voor het dagelijks leven van belang zijn, zoals kaartjes of boodschappenbriefjes. In het bo wordt door enkele leerkrachten ook genoemd dat ze met name bij de keuze van hun taken willen aansluiten op het VO, door leerlingen ter leren samenvatten.

Opmerkelijk is dat leerkrachten bij de categorie *anders* regelmatig refereren aan de methode. Het gaat om 11 leerkrachten in het bo (11% van de geïnterviewde leerkrachten) en drie leerkrachten in het sbo (8% van de geïnterviewde leerkrachten). In het bo noemen vier leerkrachten dat ze hun visie in de praktijk brengen door te werken met voorbeelden of dat ze dat doen door leerlingen te motiveren voor schrijven en/of het respecteren van eigenaarschap van de leerling. De sbo-leerkrachten spreken hier niet over. Uitgaande van de steekwoorden komt bij *anders* daarnaast een diversiteit aan onderwerpen aan de orde. Te denken valt aan interpunctie, differentiatie, handschrift, criteria noemen, 'vaak doen'.

Tabel 6.12: Manieren waarop leerkrachten hun visie in de praktijk brengen (interviews leerkrachten)

	bo N=98		sbo N=34	
	n	%	n	%
Keuzes leerdoelen	30	31%	5	15%
Keuzes in inhoud schrijftaken	32	33%	14	41%
Keuzes in aard van de schrijftaken (tekstsoorten)	32	33%	7	21%
Gebruik authentieke schrijftaken	23	23%	5	15%
Keuzes voor bepaalde werkvormen	11	11%	7	21%
Aandacht voor schrijfstrategieën	17	17%	2	6%
Aandacht voor zelfstandig werken	8	8%	0	0%
Aandacht voor samen schrijven	13	13%	1	3%
Aandacht voor peer feedback	19	19%	1	3%
Aandacht voor transfer naar andere vakken	15	15%	4	12%
Werkwijze feedback op schrijfproducten	15	15%	4	12%
Aard van de beoordeling van schrijfproducten	7	7%	1	3%
Nadruk op schrijfproces	16	16%	7	21%
Nadruk op correct taalgebruik	18	18%	7	21%
Nadruk op creatief schrijven	10	10%	4	12%
<i>Anders/toelichting</i>	31	32%	14	41%

NB: meerdere antwoorden mogelijk.

Nagegaan is of er veelvoorkomende combinaties tussen hoofdvisie en de genoemde uitwerking in de praktijk (de geprecodeerde begrippen) zijn. In Tabel 6.13 staan de meest voorkomende combinaties tussen de hoofvisies en genoemde uitwerking in de praktijk weergegeven. Zoals uit Tabel 6.13 kan worden afgeleid zijn de aantallen leerkrachten die de centrale begrippen noemen vrij vergelijkbaar voor de communicatieve en expressieve en correctheidsvisie. Alleen het rekening houden met de inhoud van schrijftaken wordt beduidend minder vaak genoemd door leerkrachten die schrijven met name zien als middel om correct te schrijven. De conceptuele functie is niet weergegeven vanwege de kleine aantallen.

Tabel 6.13: Meest genoemde centrale begrippen bij drie hoofdvisies op schrijven (interviews leerkrachten)

Visie op schrijven	Communicatief	Expressief	Correctheid
Centrale begrippen	n	n	n
Keuze leerdoelen	11	13	7
Inhoud schrijftaken	20	21	10
Tekstsoorten	16	13	13
Authentieke taken	16	12	13
Werkvormen	8	11	5
Schrijfproces	10	10	5
Transfer andere vakken	10	7	7
Schrijffouten	10	8	10

6.4.2 Leerplan (uitvoering visie/doelen in planning en voorbereiding)

Een leerplan is een concrete uitwerking van het beoogde curriculum voor een specifiek domein, in dit geval schrijven. De manier waarop leerkrachten streven naar een samenhangend schrijfaanbod zou tot uitdrukking moeten komen in de doelen die de leerkrachten over de verschillende jaargroepen heen vaststellen (de doorgaande leerlijn), maar ook in de doelen die een leerkracht voor een bepaalde jaargroep specificeert. Allereerst is daarom in het interview gevraagd of de leerkrachten voor het huidige schooljaar *zelf* expliciete doelen voor het schrijfonderwijs hebben bepaald. Uit Tabel 6.14 kan worden afgeleid dat het merendeel van de bo- en sbo-leerkrachten afgaan op de taalmethode voor het bereiken van de doelen voor schrijven en zelf geen doelen vaststellen. Bij de leerkrachten die zelf hun doelen vaststellen is gevraagd naar de basis van deze doelen. Deze open vraag is door 49 bo-leerkrachten beantwoord en door zes sbo-leerkrachten⁷⁸. De meeste leerkrachten zeggen de leerdoelen vast te stellen op basis van selecties binnen de methode. Gebruikmaking van het referentiekader, tussendoelen en kerndoelen komt veel minder vaak voor. Uit de genoteerde steekwoorden kan worden afgeleid dat er vaak sprake is van een combinatie van doelen. Genoemd wordt bijvoorbeeld een *'combinatie van methodedoelen, referentiekader en kerndoelen'* of *'aanvullende doelen [bij methode] passend bij thema, andere lessen of extra opdrachten (zoals b.v. een opdracht voor drama of een boek schrijven)'*. Incidenteel is een concreet (additioneel) leerdoel beschreven: *'eerst in het klad schrijven, daarna in het net'* en *'leestekens n.a.v. analyse citotoets'*.

⁷⁸ De vraag naar de basis van de zelf vastgestelde doelen was enkel bedoeld voor leerkrachten die aangaven *zelf* hun doelen te bepalen. Uiteindelijk blijkt de vraag ook gesteld te zijn aan een deel van de leerkrachten dat zei doelen uit de methode te gebruiken. Deze resultaten zijn in de beschrijving meegenomen.

Enkele leerkrachten hebben benadrukt dat zij hun doelen bepalen op basis van wat ze (aan het begin van het jaar) zien bij de leerlingen.

Tabel 6.14: Gebruik van tevoren vastgestelde doelen voor het schooljaar (interviews leerkrachten)

	bo (N=99)		sbo (N=36)	
	N	%	N	%
Ik stel zelf doelen vast	22	22%	6	17%
Ik stel zelf geen doelen vast	19	19%	16	44%
Ik stel zelf geen doelen vast, maar gebruik de taalmethode (en haar doelen)	58	59%	14	39%
<i>Ik stel de leerdoelen vast op basis van...</i>	N		N	
... selecties binnen de taalmethode	36	36%	0	0
... het referentiekader taal	5	5%	0	0
... tussendoelen	3	3%	1	3%
... kerndoelen	9	9%	2	6%
In overleg met collega's	7	7%	1	3%
Ik heb ze zelf opgesteld/geformuleerd	7	7%	4	11%

Ook in de *vragenlijst* is leerkrachten gevraagd naar doelen die zij hebben voor het schrijfonderwijs. Concreet is bevraagd op welke doelen leerkrachten de meeste aandacht leggen. Leerkrachten konden voor vier doelen (*doel en publiek, schrijfaanpak, tekstsoorten* en *correct taalgebruik*) aangeven hoezeer zij deze van toepassing vonden op hun schrijfonderwijs. Zoals uit Tabel 6.15 blijkt, heeft van de 126 bo-leerkrachten die deze vraag hebben ingevuld 55% van de leerkrachten de vraag naar nadruk op correct taalgebruik opengelaten; ook in het sbo is dit item bij 50% van de leerkrachten (=34 van de 67 sbo-leerkrachten die de vraag hebben ingevuld) leeggelaten. Opmerkelijk is dat in het bo en sbo nadruk op alle vier de doelen bij het overgrote deel van de leerkrachten minimaal een beetje van toepassing is. Leerkrachten in het bo benadrukken allemaal het correct leren spellen en zinnen schrijven tijdens de schrijflessen. In het sbo lijkt die focus op correctheid wat minder sterk aanwezig.

Tabel 6.15: Doelen waarop de leerkracht de meeste nadruk legt (leerkrachtvragenlijst)

		N	Helemaal niet van toepassing	Niet van toepassing	Een beetje van toepassing	Van toepassing	Helemaal van toepassing
Doel en publiek van de teksten	Bo	126	1%	6%	29%	55%	10%
	Sbo	66	3%	8%	44%	36%	9%
Schrijfaanpak	Bo	126	0%	5%	27%	56%	12%
	Sbo	66	3%	9%	35%	47%	6%
Tekstsoorten	Bo	126	0%	0%	18%	58%	24%
	Sbo	67	2%	3%	13%	58%	24%
Correct taalgebruik	Bo	57	0%	0%	2%	26%	72%
	Sbo	33	0%	6%	9%	55%	30%

6.4.3 Aandacht voor product en procesaspecten

In de vragenlijst is gevraagd naar de doelen die leerkrachten binnen hun schrijfonderwijs benadrukken (zie Tabel 6.15). In de interviews is dieper ingegaan op deze doelen door te vragen naar aspecten waaraan leerkrachten aandacht besteden in hun schrijflessen. De vraag was aan welke aspecten van schrijfvaardigheid de leerkracht dit schooljaar de meeste aandacht besteedde. Wederom was er dus sprake van een open vraag, waarbij de interviewer de spontane antwoorden vertaalde naar voor gedefinieerde antwoordcategorieën. Bij de vertaling van de antwoorden van de leerkracht heeft de interviewer a) genoteerd of de nadruk binnen het antwoord vooral lag op het schrijfproduct of het schrijfproces en b) welke aspecten hij/zij spontaan benoemt. Uit Tabel 6.16 kan worden afgeleid dat de interviewer heeft ingevuld dat de nadruk in de gegeven antwoorden bij bo- en sbo-leerkrachten vaker ligt op het schrijfproduct (44% en 53%) dan op het schrijfproces (34% en 17%).

Wanneer we in tabel 6.16 kijken naar het aantal genoemde aspecten, zien we dat leerkrachten vaker aspecten noemen die op het product gericht zijn. Ongeveer een derde van bo- en sbo-leerkrachten zegt vooral aandacht te besteden aan tekstsoorten, -opbouw en correct taalgebruik. Deze uitkomst sluit aan bij het eerdergenoemde in de praktijk brengen van de visie. Het aspect in het schrijfproces dat het vaakst is genoemd, is de nadruk op het verbeteren van spelling en grammatica (circa een derde van de bo-leerkrachten en 17% van de sbo-leerkrachten). Het geven of krijgen van feedback wordt in het bo ook regelmatig genoemd. Bij de categorie *anders* en de toelichting wordt een diversiteit aan steekwoorden gevonden. De meest genoemde zijn de 'creativiteit' in het schrijfonderwijs, en correctheidsaspecten (vaak interpunctie) (negen leerkrachten). Ook noemen leerkrachten verschillende tekstsoorten (zoals vrij schrijven, informatieve teksten, gedicht, betoog) of spreken ze over benodigde vaardigheden (beargumenteren van een standpunt). Meerdere leerkrachten gaan in op leesbaarheid, handschrift en schrijfgerei: genoemd worden aspecten als *verbonden schrift*, *duidelijkheid van de tekst – netheid van het handschrift*, *netjes schrijven* en *het juiste schrijfgerei*.

Tabel 6.16: Aspecten waaraan de leerkracht het huidige schooljaar de meeste *aandacht* geeft (interviews leerkrachten)

	bo (N=99)		sbo (N=36)	
	n	%	N	%
Nadruk in antwoord ligt op aspecten van het schrijfproduct (<i>door interviewer aangekruist</i>)	44	44%	19	53%
Leerkracht noemt aspecten waarmee nadruk wordt gelegd op product	85	88%	29	83%
<i>Dit schooljaar besteed ik de meeste aandacht aan...</i>				
Doelen van de tekst	24	24%	4	11%
Publiek van de tekst	14	14%	4	11%
Tekstsoorten	31	31%	11	31%
Tekstopbouw	37	37%	11	31%
Correct taalgebruik	37	37%	9	25%
Omvang tekst	1	1%	2	6%
Lay-out	13	13%	3	8%
Publiceren tekst	5	5%	1	3%
Geen specifiek aspect, ik volg de methode	9	9%	2	6%
Anders	6	6%	11	31%

Tabel 6.16: Aspecten waaraan de leerkracht het huidige schooljaar de meeste *aandacht* geeft (interviews leerkrachten) (vervolg)

	bo (N=99)		sbo (N=36)	
	n	%	N	%
Nadruk in antwoord ligt op aspecten van het schrijfproces (door interviewer aangekruist)	34	34%	6	17%
Leerkracht noemt aspecten waarmee nadruk wordt gelegd op proces	71	72%	15	42%
<i>Dit schooljaar besteed ik de meeste aandacht aan...</i>				
Ideeën genereren	10	10%	4	11%
Tekstplan maken	15	15%	2	6%
Formuleren	6	6%	3	8%
Feedback geven/krijgen	23	23%	1	3%
Feedback verwerken	11	11%	1	3%
Reviseren op inhoud	8	8%	4	11%
Spelling en grammatica verbeteren	34	34%	6	17%
Schrijfstrategieën	3	3%	4	11%
Samen schrijven	3	3%	1	3%
Geen specifiek aspect, ik volg de methode	6	6%	0	0%
Anders	8	8%	4	11%

Mening leerlingen: hoe hebben zij geleerd een goede tekst te schrijven?

Tijdens de interviews die met de leerlingen zijn gehouden is de leerlingen de vraag voorgelegd op welke manier zij hebben geleerd een goede tekst te schrijven. Dit geeft een beeld van het ervaren curriculum: op welke manier ervaren de leerlingen dat ze het door de leerkracht beoogde doel aanleren? De resultaten staan weergegeven in Tabel 6.17. Voor twee leerlingen, één in het bo en één in het sbo, is bij deze vraag niets ingevuld.

De helft van beide groepen leerlingen (bo en sbo) noemt dat zij vooral hebben geleerd teksten te schrijven omdat de leerkracht het heeft uitgelegd, maar niet omdat hij/zij het heeft voorgedaan. Ruim de helft van de bo-leerlingen noemt ook dat ze teksten hebben leren schrijven door de methode; dit geldt voor een kwart van de sbo-leerlingen. Ook de frequentie van het schrijven van teksten wordt door de leerlingen meermaals genoemd: 31% van de bo-leerlingen en 25% van de sbo-leerlingen.

Relatief weinig tot zeer weinig leerlingen zegt een goede tekst te hebben leren schrijven door voorbeelden (modelteksten), feedback van andere leerlingen of het herschrijven van eigen teksten.

Tabel 6.17: Manieren waarop leerlingen aangeven te hebben geleerd een goede tekst te schrijven (interviews leerlingen)

	bo (N=198)		sbo (N=71)	
	n	%	n	%
Via de taalmethode/ het lesboek	112	57%	17	24%
De leerkracht heeft het uitgelegd	102	52%	36	51%
De leerkracht heeft het voorgedaan	21	11%	3	4%
Met een stappenplan	34	17%	10	14%
Met voorbeeldteksten (bijv. een modelbrief)	22	11%	0	0%
Door vaak teksten te schrijven	62	31%	18	25%
Door echte teksten te schrijven, bijvoorbeeld een brief (authentieke taken)	19	10%	2	3%
Door samen met andere kinderen te schrijven	11	6%	1	1%
Door mijn eigen teksten te verbeteren	8	4%	0	0%
Door commentaar van leerkrachten op mijn teksten	19	10%	7	10%
Door commentaar van klasgenoten op mijn teksten	11	6%	0	0%
Anders	29	15%	15	21%

NB: meerdere antwoorden mogelijk

De opmerkingen in steekwoorden bij de categorie anders zijn zeer divers. Er zijn leerlingen die aangeven met name thuis te leren van ouders of via Squala. Ook zeggen leerlingen goede teksten te leren schrijven via andere taalonderdelen (zoals begrijpend lezen en met name spelling). Concrete soorten teksten worden ook genoemd: teksten schrijven aan de hand van plaatjes, Elfjes (=het schrijven van *lieve teksten* met 11 woorden), gedichten (bijvoorbeeld over 4 mei), werkstukken, fantasieverhalen, een brief, de schoolkrant en opdrachten over *echte onderwerpen* (zoals een bezoek aan Madurodam en de Tweede Kamer of verbeterpunten voor het volgende jaar). Sommige leerlingen gaan in op onderdelen in het schrijfproces waar ze veel van leren. Genoemd worden het maken van een woordweb (drie keer), het kijken van filmpjes (waarschijnlijk om ideeën over het onderwerp op te doen) en één leerling noemt het belang van de nabespreking van teksten. De inzet van een

wedstrijdelement en het meenemen van geschreven verhalen naar huis lijken enkele leerlingen te motiveren.

6.4.4 Afstemming met collega's en doorgaande leerlijn

In het interview is de leerkrachten gevraagd of zij afstemming hebben met hun collega's over een doorgaande leerlijn schrijfvaardigheid. Dit blijkt vaker het geval in het bo dan in het sbo: 52% van de 99 bo-leerkrachten rapporteert dit te doen, terwijl dit slechts voor een kwart van de 36 sbo-leerkrachten geldt. De genoemde redenen staan weergegeven in Tabel 6.18. Van de leerkrachten die de doorgaande leerlijn **niet** afstemmen met collega's hebben 42 bo-leerkrachten hiervoor één of meerdere redenen genoemd; voor 6 bo-leerkrachten die de doorgaande lijn niet afstemmen is geen reden genoteerd. De 27 sbo-leerkrachten geven allen minimaal één reden. Genoemde percentages in Tabel 6.18 zijn allemaal percentages van de totale groep bo-leerkrachten (n=99) en sbo-leerkrachten (n=36).

Met name het gebrek aan ervaren noodzaak door het volgen van de methode wordt in het bo, maar vooral in het sbo als reden aangedragen. Ook het gebrek aan prioriteit wordt meermaals genoemd in het bo en sbo. Additionele opmerkingen van leerkrachten over het gebrek aan afstemming laten met name zien dat afspraken slechts incidenteel zijn (bijvoorbeeld voor losse opdrachten), niet vastgelegd worden of dat extra afspraken niet nodig zijn, omdat men de methode hanteert. Van de leerkrachten die wel afstemming met hun collega's over de doorgaande leerlijn hebben (52% in bo en 25% in sbo) is vooral voor de bo-leerkrachten genoteerd hoe die afstemming in zijn werk gaat. Bo-leerkrachten overleggen vooral met collega's van de midden- en bovenbouw en daarbij stemmen ze vooral een gezamenlijke keuze in het aanbod af. Ook het taalbeleidsplan wordt wel gevolgd, waarin de doorgaande leerlijn is weergegeven. In de steekwoorden wordt de taalcoördinator of een taalcommissie op school wel genoemd, de afstemming tussen parallelklassen, de afstemming tijdens de overdracht (bijvoorbeeld: welke opdrachten zijn wel/niet gedaan) of de afstemming bij specifieke thema's als kerst. Een enkeling noemt het samen nadenken over criteria bij de beoordeling van schrijven (bijvoorbeeld het verschil tussen datgene wat verwacht wordt in groep 7 en groep 8).

Tabel 6.18: Redenen voor aan-/afwezigheid van afstemming schrijfonderwijs (interviews leerkrachten)

Afstemming met collega's andere jaargroepen over doorgaande leerlijnen	Bo (N=99) ⁷⁹		Sbo (N=36)	
	n	%	n	%
Aantal leerkrachten dat <i>geen</i> afstemming zegt te hebben	48	49	27	75%
<i>Genoemde reden(en)</i>				
Geen tijd	4	4%	0	0%
Geen prioriteit	11	11%	7	19%
Geen reden/noodzaak	7	7%	1	3%
Geen facilitering	0	0%	2	6%
Geen <i>knowhow</i>	0	0%	0	0%
Geen sturing	1	1%	0	0%
Niet nodig, we volgen de methode	19	19%	14	39%
Anders	7	7%	6	17%
Aantal leerkrachten dat <i>wel</i> afstemming zegt te hebben	51	52%	9	25%
<i>Manier van afstemmen</i>				
Aan de hand van jaarplan	1	1%	1	3%
Vastgelegd in taalbeleidsplan	12	12%	1	3%
Gezamenlijke keuzes aanbod (methode)	19	19%	2	6%
Individuele collega's	4	4%	3	8%
Met hele team	9	9%	0	0%
Met collega's midden-/bovenbouw	22	22%	1	3%
We maken samen selecties uit de methode	8	8%	2	6%
Anders	8	8%	1	3%

NB: meerdere genoemde redenen mogelijk

⁷⁹ Bij twee bo-leerkrachten was door de interviewer geen waarde ingevuld, maar op basis van de antwoordpatronen in de subcategorieën konden zij toch worden ingedeeld.

6.5 De onderwijspraktijk

Bij het in kaart brengen van het onderwijsaanbod en de praktijk van het schrijfonderwijs is gekeken naar een aantal verschillende aspecten. Hieronder wordt in paragraaf 6.5.1 allereerst ingegaan op het formele aanbod in de klas: de onderwijstijd die aan schrijven wordt besteed; de algemene doelstellingen waaraan gewerkt wordt; het publiek waar schrijfp opdrachten aan worden gericht; de gebruikte methoden, methodieken en schrijfp opdrachten waarmee wordt gewerkt; en de integratie van schrijven in ander (taal)onderwijs. Paragraaf 6.5.2 gaat dieper in op het onderwijsleerproces: gestelde lesdoelen; de instructiesetting; de activiteiten in de klas (incl. samenwerken); het monitoren, feedback geven en evalueren in (fases van) de les; het gebruik van samenwerken in de klas; en de mogelijkheden tot differentiatie. In paragraaf 6.5.3 wordt aandacht besteed aan en het volgen/monitoren van het schrijfproces; het geven van feedback en het (summatief) beoordelen en registreren van prestaties van de schrijfvaardigheden van leerlingen.

Al deze aspecten zijn net als in de vorige paragrafen gebaseerd op twee elkaar completerende bronnen: de leerkrachtvragenlijst en de interviews met leerkrachten in de verdiepende studie. Enkele van de bovengenoemde aspecten zijn ook aan bod gekomen in de interviews met leerlingen. In de gepresenteerde tabellen wordt duidelijk uit welke bron de betreffende informatie afkomstig is.⁸⁰

6.5.1 Het onderwijsaanbod

Onderwijstijd

Aan de leerkrachten is gevraagd naar de hoeveelheid onderwijstijd die aan schrijfonderwijs wordt besteed. Dit is opgevraagd en weergegeven in termen van het aantal lesminuten per week (zie Tabel 6.19a). Gemiddeld wordt er op de 133 scholen 65 minuten per week besteed aan schrijfonderwijs. De bestede tijd aan schrijfonderwijs is in het bo en sbo vergelijkbaar (62 versus 69 minuten). Wanneer we kijken naar de verschillende groepen scholen in termen van gewichtenleerlingen zien we een aflopend patroon. Hoe meer gewichtenleerlingen op de school des te lager het aantal minuten dat wordt besteed aan schrijven. Leerkrachten van scholen met geen enkele leerling met een formatiegewicht (S1, 0%) besteden gemiddeld gezien 26 minuten meer aan schrijfonderwijs dan de groep leerkrachten op scholen met relatief veel gewichtenleerlingen (S4, >25%). Dit verschil is echter niet significant ($p=0,14$).

Er dient echter te worden opgemerkt dat er sprake is van uitschieters naar boven die het gemiddelde sterk beïnvloeden⁸¹. In totaal zijn er 26 van de 184 leerkrachten (16 in het bo en 10 in het sbo) die aangeven meer dan 95 minuten per week aan schrijfonderwijs te besteden. Naast de gemiddelde en mediane waarde is daarom ook het eerste deciel en negende deciel berekend ($p_{10bo} = 30$ minuten en $p_{10sbo} = 20$ minuten en - in zowel bo als sbo - $p_{90} = 120$ minuten). Dit betekent dat de 10% leerkrachten die de minste tijd aan schrijfonderwijs besteden, zeggen dat zij dit wekelijks maximaal 20 minuten (sbo) en 30 minuten (bo) doen. De 10% leerkrachten die de meeste tijd zeggen te besteden aan schrijfonderwijs, rapporteren dit minimaal 120 minuten per week te doen. Dit geldt zowel voor bo

⁸⁰ Hier moet eveneens opgemerkt worden dat het bij de interviews steeds gaat om de door de interviewer genoteerde antwoorden.

⁸¹ Zo is er een leerkracht die aangeeft 480 minuten per week (8 uren) aan schrijfonderwijs te besteden (en 900 minuten aan taalonderwijs).

als sbo. Tabel 6.19b laat de gemiddelde bestede tijd aan schrijfonderwijs zien zonder dat deze leerkrachten worden meegenomen, dus zonder deze uitschieters. In dat geval wordt er gemiddeld in het bo 48 minuten en in het sbo 43 minuten aan schrijfonderwijs besteed. Dit beeld zou realistischer kunnen zijn. Ook het gebruik van de mediaan kan een realistischer beeld van de onderwijstijd voor schrijfonderwijs geven. Voor de volledige groep ligt de mediaan op 45 minuten per week (zowel in bo als sbo). De helft van de leerkrachten besteedt 45 minuten of minder en de andere helft 45 minuten of meer per week aan schrijfonderwijs. Wanneer we ook hier weer dezelfde uitschieters met een zeer hoog aantal minuten per week verwijderen, zien we dat de mediaan in het bo bij 45 minuten ligt, terwijl dat bij 30 minuten in het sbo is. Overigens is niet duidelijk of leerkrachten altijd alleen maar dachten aan onderwijstijd ten behoeve van de doelen voor het schrijven, of dat ze onder tijd voor schrijfonderwijs (soms) ook schrijven als middel om grip te krijgen op leerstof hebben verstaan (de conceptualiserende functie, denk b.v. aan een werkstuk).

Een vergelijking met de vorige peiling (Kuhlemeijer e.a., 2013) is voor de gemiddelde tijdsbesteding niet mogelijk, omdat deze vraag niet is gesteld. Waar het gaat om het aantal minuten voor schrijfonderwijs binnen het totaal van het taalonderwijs kan wel een beschrijvende vergelijking met de vorige peiling wel worden gemaakt. Uit deze peiling blijkt dat – wanneer we de 26 uitschieters ook hier weglaten - het aantal minuten voor schrijfonderwijs gemiddeld 10.9% (SD=6.6) van de tijd voor taalonderwijs uitmaakt, waarbij dit gemiddeld in het bo 10.8% (SD=5.7) en in het sbo 11.1% (SD=8.1) is.⁸² Hoe dit percentage zich precies verhoudt ten opzichte van de vorige peiling is niet goed na te gaan. Bekend is dat het aandeel van schrijfonderwijs binnen het totale taalonderwijs aan het eind van het (speciaal) basisonderwijs in 2009 17% in het bo en 18% in het sbo was. Echter, het gemiddelde aantal uren voor taalonderwijs, de standaarddeviatie en het aantal hoge uitschieters van de meting uit 2009 zijn niet bekend.

Tabel 6.19a: Gemiddeld aantal minuten voor schrijfonderwijs, per week (N=184) (leerkrachtvragenlijst)

Stratum	n	M	SD	Mediaan	Min	Max
S1 (0%)	5	76	34.6	60	30	120
S2 (max 10%)	82	65	53.2	45	5	300
S3 (10-25%)	22	56	28.6	45	30	120
S4 (>25%)	12	50	29.7	45	15	120
Schooltype						
bo	121	62	47.0	45	5	300
sbo	63	69	80.5	45	10	480

⁸² De vraag is of buiten de genoemde 26 leerkrachten met proportioneel veel tijdsaandacht voor schrijfonderwijs ook rekening zou moeten worden gehouden met de uitschieters bij het aantal minuten voor het gehele taalonderwijs. Zo zijn er vijf leerkracht die aangeven tussen 1000 en 1500 minuten per week (tot 5 uur per dag) aan taal te besteden. Het is onduidelijk of in de peiling van 2009 dergelijke uitschieters (naar boven) zijn verwijderd uit de data.

Tabel 6.19b: Gemiddeld aantal minuten voor schrijfonderwijs (N=158, zonder uitschieters naar boven) - leerkrachtvragenlijst

Schooltype	n	M	SD	Mediaan	Min	Max
bo	105	48	20.4	45	5	90
sbo	53	43	21.6	30	10	95

De geobserveerde lessen, waarvan de resultaten zijn gepresenteerd in Tabel 6.19c, geven een vergelijkbaar beeld van de les weer als Tabel 6.19b. De lessen in het bo duren gemiddeld significant langer dan die in het sbo ($t(df=98)=2.4, p=.02$). In beide schooltypen waren er grote verschillen in de lesduur; de lessen duurden tussen de 15 en ruim 100 minuten⁸³.

Tabel 6.19c: Gemiddelde lesduur geobserveerde schrijflessen, in aantal minuten

Schooltype	n	M	SD	Mediaan	Min	Max
bo	70	46	14.17	45	23	103
sbo	30	39	13.61	38	15	73

Gebruik van methodes

In de onderwijstijd die beschikbaar is voor schrijfonderwijs wordt een uiteenlopend palet aan methoden gebruikt. 190 van de 193 leerkrachten (98%) gaven informatie over de door hen gebruikte methode(n) (Tabel 6.20)⁸⁴. Zowel in het bo als het sbo worden de methodes van uitgever Malmberg (Taal actief en Staal) het meest gebruikt. Ongeveer een op de tien leerkrachten geeft aan dat er meerdere methodes worden gebruikt voor het schrijfonderwijs. Opvallend is dat alleen in het bo er leerkrachten zijn (3%) die aangeven geen enkele methode te gebruiken, maar louter gebruik maken van (online) gevonden en/of zelfontwikkeld materiaal, dat soms door de leerkracht een 'zelfontwikkelde' methode wordt genoemd.

In de categorie 'anders' worden (over de gehele groep van 190) onder andere Taalverhaal, Taaltrapeze, Taaljournaal, Taalfontein en de Montessori Taalset genoemd. Zeven leerkrachten noemen het gebruik van Snappet Connected learning. Dat is een online community waar schrijfopdrachten en instructiematerialen beschikbaar zijn (zie <https://nl.snappet.org> voor meer informatie). Snappet wordt net als de Montessori Taalset meestal als aanvulling op een meer traditionele methode ingezet, maar er zijn ook leerkrachten die deze bronmaterialen voor al hun taalonderwijs inzetten.

⁸³ Uitgaande van de range in geobserveerde lesduur doen de resultaten op basis van de data waaruit de uitschieters zijn verwijderd (zoals in Tabel 6.19b) mogelijk geen recht aan de werkelijkheid.

⁸⁴ In Tabel 6.20 worden de data op leerkrachtniveau weergegeven. Nagegaan is of er binnen dezelfde school in verschillende groepen 8 dezelfde methode wordt gebruikt. Dit is het geval.

Tabel 6.20: Gebruikte methodes (leerkrachtvragenlijst)

Methode	bo (N=124)		sbo (N=66)	
	n	%	n	%
Taal Actief (Malmberg)	35	28%	12	18%
Taalverhaal.nu (ThiemeMeulenhoff)	10	8%	3	5%
Staal (Malmberg)	25	20%	17	26%
Taal op maat (Noordhoff)	18	15%	11	17%
Taal in Beeld (Zwijssen)	9	7%	9	14%
Een combinatie* van (delen uit) methodes	12	10%	6	9%
Andere methode/ander methodemateriaal	11	9%	8	12%
Geheel geen methode	4	3%	0	-

* De vijf bovengenoemde en/of 'andere methode'

De geobserveerde lessen bleken in merendeel geen methodelessen te zijn. Dit gold zowel voor het bo als voor het sbo. In het bo is van 75 geobserveerde schrijflessen duidelijk of het een methodeles betrof: in 50 lessen was dit niet het geval (67%). In het sbo was voor 29 lessen ingevuld of de geobserveerde les een methodeles betrof. Hiervan waren 19 lessen geen les uit de methode (66%).

Aanvullende leermaterialen en hulpmiddelen

De gehanteerde methode staat meestal centraal in het schrijfonderwijs. De helft van de leerkrachten gebruikt louter de door hen genoemde methode(n). De andere helft van de leerkrachten (54% bo, 47% sbo) van de leerkrachten gebruikt ook aanvullende (officiële) leermiddelen. Meest genoemd, in de zeer uiteenlopende voorbeelden, zijn daarbij schrijfopdrachten n.a.v. *Nieuwsbegrip*, *Pennenstreken* (en andere schrijf oefeningen voor handschrift) en Snappet. Ook online (stel)opdrachten & schrijfkaarten van diverse (online) uitgevers worden genoemd.

Naast methode en aanvullend materiaal geven 64 bo-leerkrachten (53%) en 19 (30%) sbo-leerkrachten aan eigen zelfontwikkelde materialen bij het schrijfonderwijs inzetten. Er is in de vragenlijst niet gevraagd naar wat dit inhoudt.

In de interviews is eveneens gevraagd naar de inzet van hulpmiddelen voor leerlingen bij het maken van schrijfopdrachten. Van de geïnterviewde leerkrachten – in zowel bo als sbo - geeft het overgrote deel (93% bo, 89% sbo) in het gesprek aan dat leerlingen bij het maken van een schrijfopdracht één of meerdere hulpmiddelen kunnen gebruiken.⁸⁵ Tabel 6.21a bevat een overzicht van de genoemde hulpmiddelen. Leerlingen in bo-klassen kunnen, daar waar hulpmiddelen voorhanden zijn, vooral gebruik maken van de (standaard gehanteerde) taalmethode⁸⁶, een stappenplan, woordenboeken of een tekst/schrijfplan (door respectievelijk 43, 43, 29, 26 procent van de leerkrachten genoemd). Een tekstverwerker met spellingscontrole is in het bo minder gebruikelijk dan in het sbo (door 29% van de bo- en 52% van de sbo-leerkrachten genoemd). In het sbo is daarentegen een stappenplan, een schrijfplan of het gebruik van de taalmethode als hulpmiddel veel minder gebruikelijk dan in het reguliere basisonderwijs.

⁸⁵ Ongeveer 1 op de 10 leerkrachten (totale groep) geeft aan geen enkel hulpmiddel in te zetten, terwijl van 15 procent niet duidelijk is of deze in de schrijflessen (kunnen) worden gebruikt.

⁸⁶ Taalmethoden bieden diverse hulpmiddelen, bijvoorbeeld stappenplannen voor leerlingen

De meeste ‘andere’ hulpmiddelen die worden genoemd zijn de laptop/pc met internet, en online applicaties (Wikipedia, Woordweb, online kranten). Enkele leerkrachten noemen hulpmiddelen voor dyslectici: vergrote teksten/opdrachten, spraak gestuurde digitale opdrachten.

Tabel 6.21a: Percentage leerkrachten dat aangeeft dat hulpmiddel beschikbaar is (interviews leerkrachten)

	bo (N=89)	sbo (N=31)
	%	%
Stappenplannen	43%	23%
Schrijfplan/tekstplan	26%	7%
Taalmethode	43%	19%
Hulpkaarten voor taalverzorging	14%	16%
Voorbeeldtekst (modeltekst)	24%	13%
Bij tekstverwerker: spelling/grammatica controle	24%	52%
Woordenboek	29%	29%
Anders	26%	19%

NB: meerdere antwoorden mogelijk

De observatoren is ook gevraagd te noteren of en zo ja, welk (hulp)materiaal gebruikt werd tijdens de geobserveerde schrijfles. Met name in het bo werd geregeld gebruik gemaakt van schrijf- of stappenplannen (n=15, in 19% van de lessen), terwijl dit in het sbo in twee lessen is gerapporteerd (7%). Het gebruik van voorbeeldteksten is in drie bo-lessen (4%) en in één sbo-les genoteerd (3%). Benadrukt wordt dat dit slechts één schrijfles per leerkracht betreft.

Ook de leerlingen is gevraagd naar het gebruik van hulpmiddelen tijdens de schrijflessen. Respectievelijk 66% en 24% van de bo- en sbo-leerlingen geeft aan dat ze gebruik kunnen maken van hulpmiddelen. Tabel 6.21b bevat de door de leerlingen in het gesprek genoemde voorbeelden. Naast het lage aantal sbo-leerlingen dat antwoord geeft op deze vraag (17) laat de tabel zien dat de verschillende hulpmiddelen steeds door een gering aantal sbo-leerlingen wordt genoemd (twee tot negen leerlingen). In de categorie ‘anders’ worden, net als bij de leerkrachten, computerapparatuur en online bronnen genoemd (‘Google’, ‘Woordweb’, ‘Chromebook’, ‘Youtube’) maar ook spellingskaarten en (relatief vaak genoemd) de schrijfmateriaal zelf: potlood, gum, kladpapier.

Vergeleken met Tabel 6.17, waarin leerlingen noemen van wie/wat ze vooral leren om goede teksten te schrijven, zien we bij de bo-leerlingen dat zij bij de expliciete vraag naar hulpmiddelen relatief vaker het stappenplan noemen (in T6.21b 30% en in T6.17 17%) en minder vaak de algemene taalmethode (in T6.21b 28% en in T6.17 57%). Voorbeeldteksten werden bij beide vragen relatief weinig genoemd (tussen 11% en 15%).

Tabel 6.21b: Hulpmiddelen die worden genoemd door leerlingen (interviews leerlingen)

	bo (N=132)		sbo (N=17)	
	n	%	n	%
Stappenplannen	39	30%	2	12%
Schrijfplan/tekstplan	21	16%	2	12%
Taalmethode	37	28%	4	24%
Hulpkaarten voor taalverzorging	16	12%	2	12%
Voorbeeldtekst (modeltekst)	20	15%	0	-
Bij tekstverwerker: spel/grammatica controle	20	15%	2	12%
Woordenboek	26	20%	9	53%
Anders	43	33%	5	29%

NB: meerdere antwoorden mogelijk

Frequentie schrijfp opdrachten

In de leerkrachtenvragenlijst is gevraagd hoe vaak per maand de leerkracht schrijfp opdrachten aan zijn/haar leerlingen geeft. Daarbij is een onderscheid gemaakt naar de frequentie van taken binnen het taalonderwijs en daarbuiten. Uit Tabel 6.22 kan worden opgemaakt dat ongeveer twee derde van de leerlingen twee keer of vaker per maand een schrijfp opdracht binnen het taalonderwijs moet maken. Dit geldt voor 72% van de leerlingen in groep 8 en voor 62% van de schoolverlaters. Bij de vorige peiling (Kuhlemeijer, 2013) lagen deze percentages op 74% bij groep 8-leerlingen en 72% bij sbo-leerlingen. 22% (bo) en 24% (sbo) schrijft gemiddeld één keer per maand een tekst als opdracht binnen het taalonderwijs. Het percentage leerlingen dat minder dan één keer per maand een schrijfp opdracht binnen het taalonderwijs moet maken is in het bo lager (7%) dan dat in het sbo (13%). In 2009 lagen deze percentages nog op 4% voor het bo en 2% voor het sbo.

Tabel 6.22: Frequentie van schrijfp opdrachten, binnen het taalonderwijs (leerkrachtvragenlijst)

Frequentie	bo (N=124)		sbo (N=67)	
	n	%	n	%
Minder dan 1 keer per maand	8	7%	9	13%
1 keer per maand	27	22%	16	24%
2 keer per maand	40	32%	19	28%
Vaker dan 2 keer per maand	49	40%	23	34%

Wanneer wordt gekeken naar verschillen tussen scholen met een verschillende leerlingpopulatie qua gewichtenleerlingen dan zien we – geaggregeerd naar twee categorieën – een opvallend oplopend patroon. Hoe meer gewichtenleerlingen hoe relatief vaker de leerlingen twee keer of meer per maand

een schrijfpdracht moeten maken.⁸⁷ Leerlingen op scholen met relatief veel gewichtenleerlingen lijken dus vaker teksten te schrijven. Wanneer we deze resultaten echter combineren met de resultaten in Tabel 6.19a, dan zien we dat leerlingen op scholen met relatief veel gewichtenleerlingen vaker teksten schrijven, terwijl de door leerkrachten ingeschatte totale tijd besteed aan schrijfonderwijs op deze scholen lager ligt. Echter, de verschillen in tijd bleken niet significant. Hierdoor is het de vraag of we kunnen stellen dat leerlingen in S3 en S4 weliswaar meer schrijven maar in beperkter tijd voor schrijfonderwijs.

Tabel 6.23: Frequentie van schrijfpdrachten, binnen het taalonderwijs, naar schoolgewicht (leerkrachtvragenlijst)

Frequentie	S1 (n=7)	S2 (n=83)	S3 (n=22)	S4 (n=12)
1 keer per maand of minder	43%	33%	17%	8%
2 keer per maand of meer	57%	67%	82%	92%

Schrijfpdrachten - zowel methodegebonden als zelfontwikkelde opdrachten – worden veelal opgezet om in fases te worden uitgevoerd. Bijvoorbeeld: plannen eerst, vervolgens schrijven, daarna feedback en revisie, en in de laatste les: lay-out en publicatie. Aan de leerkrachten is daarom gevraagd in hoeverre de schrijfpdrachten zijn verspreid over meerdere lessen (Tabel 6.24). Uit de Tabel blijkt dat in het bo relatief vaker schrijfpdrachten over meerdere lessen worden verspreid dan in het sbo.

Tabel 6.24: Schrijfpdrachten over meerdere lessen (leerkrachtvragenlijst)

Frequentie	bo (N=126)		sbo (N=66)	
	n	%	n	%
Nooit	8	6	10	15
Soms	56	44	40	61
Meestal	45	36	11	17
Altijd	17	14	5	8

Tablet of PC als transcriptiemedium

Voor het schrijven wordt in de meeste gevallen pen en papier gebruikt. Toch geeft twee derde (66.4%) van de bo-leerkrachten en ruim twee vijfde (44.8%) van de sbo-leerkrachten aan dat leerlingen minstens één keer per maand schrijfpdrachten maken op de pc, een laptop of een tablet. Voorbeelden die uit de interviews naar voren komen zijn: het schrijven van een e-mail, een stuk voor

⁸⁷ Gezien het lage aantal leerkrachten in de groepen 0% en >25% gewichtenleerlingen moet dit echter voorzichtig worden geïnterpreteerd.

de schoolkrant of de schoolwebsite, maar vooral het werken aan een werkstuk (dat ingeleverd en beoordeeld dient te worden). Respectievelijk 2 en 13 procent van de bo- en sbo-leerkrachten geeft aan dat hij/zij nooit digitale middelen inzet als transscriptiemedium bij de schrijflessen.

Positie schrijfvaardigheid binnen andere subdomeinen van het taalonderwijs

Binnen een brede visie op taalbeleid vervlechten leerkrachten aandacht voor meerdere taaldomeinen in hun taalonderwijs (SLO). Zo kan een leerkracht in een les, die primair gericht is op begrijpend lezen, ook bewust aandacht besteden aan de manier waarop de tekst is geschreven. Leerkrachten is daarom in de interviews gevraagd of en op welke manier zij binnen andere taaldomeinen bewust aandacht besteden aan schrijfvaardigheid. Interviewers konden de op voorhand gedefinieerde categorieën leesonderwijs, mondelinge taalvaardigheid en taalverzorging aanvinken, met daaronder de specifieke focus die de leerkracht dan kan aanbrengen (zoals 'tekstsoorten' of 'doelgericht communiceren' communiceren bij leesonderwijs, 'doelgericht communiceren' bij mondelinge taalvaardigheid of 'spelling' bij taalverzorging).⁸⁸ Van de bo-leerkrachten zegt 82% dit te doen en van de sbo-leerkrachten 83%. Tabel 6.25 laat zien dat (afgerond) 58% van de geïnterviewde bo-leerkrachten benoemt dat zij het schrijfonderwijs koppelen aan het leesonderwijs en 56% van de sbo-leerkrachten. Schrijven wordt door meer dan de helft van de bo-leerkrachten gekoppeld aan taalverzorging, terwijl dit door ruim een derde van de sbo-leerkrachten wordt gedaan. Het schrijfonderwijs is bij eveneens een derde van de sbo-leerkrachten verbonden aan de mondelinge taalvaardigheid. In het reguliere basisonderwijs lijkt dit wat meer los te staan van het schrijven: één op de vijf leerkrachten benoemt een directe relatie/koppeling met het schrijfonderwijs. Wanneer vervolgens naar de aspecten binnen de taaldomeinen wordt gekeken, valt allereerst op dat de meeste leerkrachten slechts één aspect benoemen. Voor leesonderwijs geldt dat steeds minder dan de helft van de leerkrachten het sub domein benoemd in het gesprek. Tekstsoorten is het sub domein dat het meeste wordt genoemd door bo-leerkrachten, terwijl de sbo-leerkrachten in het leesonderwijs met name voor 'tekstopbouw' en 'taalgebruik' de koppeling te maken met het schrijfonderwijs. Het schrijfonderwijs wordt door de bo-leerkrachten binnen het domein taalverzorging vooral gekoppeld aan spelling en interpunctie, en minder (40%) aan grammatica. Van de sbo-leerkrachten noemt 50-75% de drie sub domeinen van Taalverzorging, het aantal leerkrachten is echter steeds gering (9,7, en 9 leerkrachten).

⁸⁸ Bij deze vraag geldt dat de domeinen (lezen, mondelinge taalvaardigheid en taalverzorging) soms zijn aangevinkt maar geen subdomeinen (aspecten). Om een indruk te geven van de mate waarin leerkrachten aangeven het schrijfonderwijs te koppelen aan deze subdomeinen, zijn symbolen gebruikt i.p.v. exacte percentages. NB: het symbool weerspiegelt steeds een percentage van het aantal leerkrachten waarvoor ook het hoofdomein is aangekruist.

Tabel 6.25: Koppeling met andere domeinen in taalonderwijs (interviews leerkrachten)

Koppeling met		bo (N=99)		sbo (N=36)	
		n	%	n	%
Leesonderwijs		57	57.6%	20	55.6%
	Tekstsoorten	++		+	
	tekstopbouw	+		++	
	doelgericht communiceren	+		+	
	publiekgericht communiceren	-		+	
	taalgebruik (zinsbouw en woordkeuze)	+		++	
	publiceren	-		-	
Mondelinge taalvaardigheid		20	20.2%	12	33.3%
	doelgericht communiceren	++		++	
	publiekgericht communiceren	+++		++	
	feedback geven en ontvangen	+		-	
Taalverzorging		52	52.5%	13	36.1%
	Spelling	*		+++	
	Grammatica	++		+++	
	Interpunctie	+++		+++	

NB: frequenties; - = genoemd door 0-10% , + = 11-25% of minder leerkrachten, ++ = 25-50%, +++ = 50-75% * = >75%

Het samenvatten van een tekst, o.a. bij *Nieuwsbegrip* of een (boek)verslag, wordt door relatief veel leerkrachten genoemd bij 'anders' of in de toelichting als het gaat om een koppeling maken met het leesonderwijs: "begrijpend lezen en dan samenvatten op papier". Ook het koppelen aan de woordenschat (als onderdeel van het leesonderwijs) wordt enkele keren genoemd. Als het gaat om de mondelinge taalvaardigheid wordt een omschrijving bij 'anders' maar beperkt gegeven. Door enkele leerkrachten wordt de koppeling gelegd met zaken als argumenteren, spreken, (geschreven) mening verwoorden, spreekbeurten en het presenteren van een boekverslag. Bij de koppeling met taalverzorging wordt in de categorie 'anders' door ongeveer tien leerkrachten een interpretabel antwoord geven, meestal gaat het om 'verzorgd' of 'correct' taalgebruik en de link met het begrip van woorden.

Ongeveer vijftien leerkrachten geven een toelichting hoe de verschillende disciplines (kunnen) worden gecombineerd. Enkele voorbeelden:

- Bij Aardrijkskunde, Geschiedenis en Biologie verwerken de leerlingen al het geleerde op verschillende manieren die raakvlak hebben met taalonderwijs (w.o. schrijven);
- Woordenschat: leerkracht noemt (nieuwe) woorden en leerlingen schrijven daarmee passende zinnen;
- Link met Nieuwsbegrip of andere begrijpend lezen opdrachten: leerlingen vatten teksten samen, maken een 'woordspin', schrijven voorkennis op;
- Activiteiten als "vertel of schrijf het verhaal verder" (op basis van aanzet van leerkracht);
- Leerlingen schrijven en maken vervolgens een radio-uitzending.

Integratie in/met andere vakken

Binnen breed taalbeleid is er naast aandacht voor schrijfvaardigheid binnen andere *taaldomeinen* ook aandacht voor schrijfvaardigheid binnen andere domeinen en vakken. Zo kunnen leerlingen werken aan schrijfoopdrachten binnen andere vakken. Te denken valt aan het maken van werkstukken bij andere vakken (zoals wereldoriëntatie) of schrijfoopdrachten binnen projectonderwijs. De meeste leerkrachten geven aan leerlingen dit ook regelmatig te laten doen: 55% van de groep 8-leerlingen schrijft twee keer per maand of vaker buiten het taalonderwijs aan een opdracht en in het sbo ligt dit percentage op 40.3%. In de vorige peiling (Kuhlemeijer, 2013) lag dit percentage voor beide groepen leerlingen op 40%.

Tabel 6.26 laat zien dat er in het bo gemiddeld gezien per jaar iets vaker dan in het sbo buiten het schrijfonderwijs in andere vakken door leerlingen wordt gewerkt aan schrijfoopdrachten. In het sbo werkt ruim twee derde (68.7%) één keer per maand of meer aan een schrijfoopdracht in andere vakken, terwijl dat in het bo 83.3% is. Buiten het taalonderwijs werken leerlingen in het bo in 56% van de klassen en leerlingen in het sbo in 40% van de klassen, twee keer of vaker per maand aan een geschreven tekst. Het gaat dan bijvoorbeeld om het schrijven van een tekstje binnen projectonderwijs of om werkstukken bij andere vakken.

Tabel 6.26: Frequentie van integratie schrijfonderwijs in andere vakken (leerkrachtvragenlijst)

	bo (N=126)		sbo (N=67)	
	n	%	n	%
Nooit	0	-	1	1.5%
Minder dan één keer per maand	21	16.7%	20	29.9%
Eén keer per maand	35	27.8%	19	28.4%
Twee keer per maand	28	22.2%	11	16.4%
Vaker dan twee keer per maand	42	33.3%	16	23.9%

In de interviews is ook gevraagd of binnen andere vakken, zoals Wereldoriëntatie en Wetenschap en Techniek, aan schrijfonderwijs wordt gewerkt. Ruim vier vijfde van de bo- leerkrachten geeft aan dat dit inderdaad gebeurt (80.8%). Bij de sbo-leerkrachten is dat precies twee derde (66.7%). Voor 79.8% (79), respectievelijk 63.9% (23, sbo) van de leerkrachten is aangekruist dat het om wereldoriëntatie en/of wetenschap/techniek gaat. Gevraagd naar de manier waarop schrijflessen zijn gekoppeld aan deze vakken worden ruim 50 (combinaties van) antwoorden gegeven, maar ontbreekt er voor één bo-leerkracht en twee sbo-leerkrachten informatie over deze manieren. Tabel 6.27 bevat de frequenties van de manieren waarop schrijven bij vakken zoals wereldoriëntatie en wetenschap/techniek aan bod komt. De meest voorkomende aspecten zijn het gebruik van authentieke schrijftaken (met name in het bo), het toepassen in de les van kennis van tekstsoorten, en aandacht voor de verzorging van teksten. In het sbo komt schrijven bij een derde van de leerkrachten ook naar voren via het (aanleren van) het maken van aantekeningen (en deze te verwerken).

Tabel 6.27: Integratie schrijfonderwijs in wereldoriëntatie en wetenschap/techniek (interviews leerkrachten)

	bo (N=78)		sbo (N=21)	
	n	%	N	%
Gebruik van authentieke schrijftaken, gerelateerd aan de besproken thema's	37	47%	4	19%
Toepassen van schrijfstrategieën	15	19%	2	10%
Toepassen van kennis van tekstsoorten	26	33%	5	24%
Toepassen van kennis over doelgericht schrijven	19	24%	4	19%
Toepassen van kennis over publiekgericht schrijven	19	24%	2	10%
Aantekeningen maken (en verwerken)	19	24%	7	33%
Teksten herschrijven	8	10%	1	5%
Aandacht voor verzorging van teksten	22	28%	5	24%
Anders	16	21%	7	33%

Ook hier wordt in de toelichting vooral gesproken over 'samenvatten van teksten' voor de vakken wereldoriëntatie en techniek, hetzij als opdracht of als onderdeel van het maken van een presentatie, een muurkrant of een verslag. Ook het schriftelijk weergeven van de mening van de leerling - bijvoorbeeld in de les burgerschap of naar aanleiding van een (geponeerde) stelling in Nieuwsbegrip - worden genoemd. Een aansprekend voorbeeld van een integratie van het schrijfonderwijs in een ander vak is het inleven en beschrijven van een [historisch] figuur bij de geschiedenisles.

Schrijven ten dienste van leren

In de vragenlijst is leerkrachten gevraagd hoe vaak zij een schrijfofdracht inzetten als middel om kennis te vergaren of inzicht in verbanden te krijgen, de conceptualiserende functie van schrijven (Tabel 6.28). Hoewel de leerkrachten in de interviews amper refereren aan de conceptualiserende functie als onderdeel van goed schrijfonderwijs (zie Tabel 6.10), laten de vragenlijstresultaten een tamelijk gunstig beeld zien. Leerkrachten rapporteren regelmatig gebruik te maken van schrijfofdrachten die leerlingen helpen grip op de leerstof te krijgen. Het schrijven om kennis op te doen over een onderwerp (bv. een werkstuk) of een samenvatting maken zijn relatief vaker onderdeel van het onderwijs in het bo. Schrijven door het laten maken van aantekeningen wordt ongeveer in gelijke mate ingezet in bo en sbo. Ten behoeve van de meerniveau-analyses is tevens een schaalscore (gemiddelde) berekend, bestaande uit de drie items. Gemiddeld rapporteren de bo- en sbo-leerkrachten schrijfofdrachten t.b.v. de conceptualiserende functie iets vaker dan één keer per kwartaal in te zetten, waarbij de bo-leerkrachten (M=3.33) dat iets vaker lijken te doen dan de sbo-leerkrachten (M=3.12).

Tabel 6.28: Frequentie leerkrachten dat schrijven ten dienste van leren inzet (leerkrachtvragenlijst)

		N	Nooit	Een keer per jaar	Een keer per kwartaal	Minimaal een keer per maand
Een tekst schrijven om kennis op te doen over een onderwerp	bo	125	1%	13%	52%	34%
	sbo	66	6%	17%	50%	18%
Een samenvatting maken van andere teksten in een eigen tekst	bo	125	2%	4%	30%	65%
	sbo	67	3%	9%	43%	45%
Aantekeningen maken over leerstof om er later een tekst over te schrijven	bo	126	7%	11%	37%	45%
	sbo	66	9%	17%	33%	41%
		N	Min	Max	M	SD
Schaalscore: 'schrijven om te leren'	bo	126	1.33	4.0	3.33	.57
	sbo	67	1.66	4.0	3.12	.70

Alhoewel er in de interviews niet specifiek is gevraagd naar (voorbeelden) hoe leerkrachten schrijfopdrachten inzetten in dienst van het leerproces blijkt uit de toelichtingen bij andere vragen dat teksten schrijven om iets te leren (inclusief het zoeken naar informatie/bronnen), het samenvatten van teksten en het maken van aantekeningen vooral bij de zaakvakken een rol speelt. Dit is met name het geval bij wereldoriëntatie en burgerschap en komt zowel in afzonderlijke lessen als bij het maken van grotere werkstukken voor.

6.5.2 De lespraktijk

Ten aanzien van het de lespraktijk zijn in de vragenlijst een achttal vragen gesteld over de volgende thema's: het soort schrijfopdrachten (soort teksten) dat een leerkracht aanbiedt; de doelstellingen die schrijfproducten dienen te bereiken; het publiek waarvoor de schrijfproducten worden gemaakt; de leeractiviteiten die in de les rondom schrijven worden toegepast; de mate waarin leerlingen samenwerken en elkaar van feedback voorzien in de schrijfles; de mate waarin in de lessen aandacht wordt besteed aan soorten teksten en schrijfprocessen; en de manier(en) waarop leerkrachten in hun schrijfonderwijs differentiëren tussen leerlingen. Ook in de interviews zijn vragen over enkele van deze aspecten gesteld. Extra thema's die in de interviewleidraad waren opgenomen zijn: de gehanteerde instructiesetting en de nadruk die wordt gelegd in de instructie (schrijfproces versus schrijfproduct); het gebruik van lesdoelen per schrijfles; en de manieren waarop leerkrachten trachten de motivatie van leerlingen te bevorderen in/voor de schrijfles. In deze paragraaf worden de resultaten van bovengenoemde onderwerpen beschreven. Daar waar van toepassing worden deze gecompleteerd (dan wel geïllustreerd) met gegevens uit de interviews met leerlingen en uit de lesobservaties.

Geobserveerde schrijflessen: interactie-episodes

Zoals paragraaf 1.6.2 is beschreven, is gebruik gemaakt van het observatie-instrument van Van Gelderen en Blok (1991). Daarin zijn de volgende lesfasen onderscheiden:

1. De **voorbereidende fase**, waarin de leerkracht kan ingaan op het overzicht van de les (lesoverzicht), het lesdoel kan formuleren, het onderwerp kan introduceren, de opdracht kan definiëren en

leerstof kan geven (over de schrijfpdracht). De leerkracht kan daarbij strategieën aanreiken en activiteiten initiëren om gezamenlijk het onderwerp van de schrijfpdracht te verkennen of dit onderwerp door leerlingen zelf laten verkennen. Dit kan door middel van een brainstorm, discussie, verzamelen van materiaal en het maken van een planning of op een andere manier (bijvoorbeeld door het aanreiken van een voorbeeldtekst).

2. De **schrijffase**, waarin de leerlingen aan het werk zijn. De leerkracht kan hierbij hulp bieden of motiveren (aanmoedigen of ter beschikking staan).
3. De **revisiefase**, waarin het schrijfproduct wordt aangepast. De leerkracht kan in deze fase hulp aanbieden en aanmoedigen, maar ook leerstof aanbieden of met leerlingen discussiëren.
4. De **evaluatiefase**, waarin de leerkracht de schrijfproducten en het schrijfproces met de klas bespreekt.
5. De **publicatiefase**, waarin de tekst wordt bewerkt of vormgegeven, zodat hij effectief de doelgroep kan bereiken. De leerkracht kan hierbij wederom hulp bieden of feedback geven of vragen.

Per fase kan het voorkomen dat de leerkracht gericht is op iets anders dan de schrijfles (non-instructie) of op de overgang naar een andere taak, werkvorm of fase (transitie). Conform de werkwijze van Van Gelderen en Blok wordt bij de indeling in lesfasen uitgegaan van het perspectief van de leerling. Uitgaande van dit perspectief hebben Van Gelderen en Blok ervoor gekozen om non-instructie als onderdeel van de tijd besteed aan een lesfase mee te nemen. Omdat het hoogste percentage non-instructie wordt gevonden in de schrijffase als de leerlingen zelfstandig aan het werk zijn, is de non-instructie geen onderbreking, juist eerder een onderdeel van die fase. Transitie is als aparte categorie meegenomen. Op basis van de data kan een beeld worden geschetst van de tijdsverdeling over de verschillende lesfasen, uitgesplitst naar bo en sbo (zie Tabel 6.29).

Tabel 6.29: Tijdsverdeling over de lesfasen (inclusief non-instructie)

		Peiling 2019					B&vG, 1991
Schooltype	N	Lesfase	Gem. % lestijd	SD	Min	Max	Gem
bo ⁸⁹	70	Voorbereiding	34.2%	0.18	0%	81%	27.8%
		Schrijven	38.6%	0.18	0%	73%	36.5%
		Revisie	8.2%	0.12	0%	55%	2.3%
		Evaluatie	7.2%	0.09	0%	55%	10.5%
		Publicatie	1.3%	0.04	0%	19%	15.9%
		Transitie	6.1%	0.05	0%	21%	7%
sbo	30	Voorbereiding	34.2%	0.20	3%	90%	
		Schrijven	45.2%	0.18	2%	82%	
		Revisie	5.0%	0.11	0%	45%	
		Evaluatie	14.4%	0.12	0%	49%	
		Publicatie	0.3%	0.01	0%	4%	
		Transitie	7.5%	0.06	0%	26%	

⁸⁹ Doordat we in het bo een inclusiecriteria van minimaal 85% gescoorde lesmomenten hanteren, zijn niet alle lesmomenten gescoord. De percentages lopen daarom niet op tot 100%.

Zoals uit Tabel 6.29 kan worden afgeleid, bestaat het grootste deel van de schrijflessen in zowel bo als sbo uit de voorbereidende fase (in zowel bo als sbo 34%) en uit de schrijffase (bo: 39%, sbo: 45%), ook al verschillen de lessen onderling wel. De spreiding in het bo is voor beide fasen .18 en in het sbo .20 en .18. De overige fasen nemen aanzienlijk minder tijd in beslag, met name aan publicatie wordt amper tijd besteed (1% in bo en iets meer dan 0% in sbo). Opmerkelijk is wel dat in het sbo gemiddeld 14% van de geobserveerde tijd is besteed aan de evaluatiefase.

Wanneer we deze uitkomsten vergelijken met Van Gelderen en Blok (1991), die hun onderzoek in het bo hebben uitgevoerd, dan zien we in het bo vrij vergelijkbare uitkomsten voor schrijffase en transitie (zie Tabel 6.29). De gemiddelde lestijd die aan voorbereiding wordt besteed is echter wel hoger in 2019. De in de les bestede aandacht aan revisie was begin jaren '90 lager dan in 2019, terwijl dit toen voor het publiceren/presenteren van de teksten aanzienlijk meer was dan in 2019. Ook aan evaluatie van teksten leek toentertijd meer tijd besteed te worden.

Per lesfase is een nadere onderverdeling gemaakt naar de specifieke onderdelen waaraan tijd wordt besteed door de leerkrachten. In Tabel 6.30 worden de resultaten per onderdeel gepresenteerd, wederom uitgesplitst naar schooltype, waarbij de onderdelen die gemiddeld meer dan 5% van de totale lestijd in beslag namen grijs gearceerd zijn. Binnen de voorbereidende fase wordt gemiddeld de meeste tijd aan de definiëring van de opdracht en het geven van leerstof besteed in bo en sbo. In het sbo besteden leerkrachten daarnaast ook ruim 5% van de tijd aan het introduceren van het onderwerp waarover de tekst wordt geschreven. In de schrijffase zijn leerkrachten in bo en sbo veelal actief in het geven van hulp en/of aanmoediging van leerlingen (tussen 16% en 20% van de totale lestijd). In het sbo werd gemiddeld 8% van de lestijd aan de algemene/klassikale feedback op de gemaakte teksten besteed, maar de range in minimum en maximum bestede tijd is groot. Gezien de gegeven ranges tussen minimum en maximum bestede tijd blijken de lessen soms flink te verschillen waar het gaat om de tijd besteed aan de onderdelen. Dit geldt voor beide schooltypen.

Tabel 6.30: Gemiddelde percentages tijd besteed aan onderdelen per lesfase, naar schooltype⁹⁰

Fase	% tijd voor..	bo				sbo			
		Gem	S.D.	Min	Max	Gem	S.D.	Min	Max
Voor- bereidende fase	Lesoverzicht	0.6%	0.01	0%	8%	1.3%	0.02	0%	6%
	Lesdoel	1.4%	0.02	0%	8%	0.2%	0.01	0%	3%
	Introductie onderwerp	4.4%	0.05	0%	24%	5.6%	0.05	0%	19%
	Non-instructie	1.2%	0.03	0%	15%	0.6%	0.01	0%	3%
	Transitie	3.7%	0.04	0%	20%	4.6%	0.04	0%	18%
	Definiëring opdracht	8.6%	0.07	0%	34%	6.1%	0.06	0%	24%
	Leerstof geven	7.6%	0.07	0%	25%	5.5%	0.09	0%	40%
	Brainstorm	3.2%	0.07	0%	42%	4.8%	0.06	0%	19%
	Discussie met/tussen leerlingen	1.3%	0.03	0%	13%	0.4%	0.01	0%	6%
	Speciale voorbereiding	1.2%	0.03	0%	29%	4.7%	0.10	0%	41%
	Verzamelen	1.2%	0.03	0%	13%	0.0%	0.00	0%	0%
	Schrijfplan maken	1.0%	0.03	0%	13%	0.5%	0.03	0%	16%
	Combi doel/definiëring opdracht	0.2%	0.01	0%	3%	-	-	-	-
	Combi onderwerp/definiëring opdracht	0.2%	0.01	0%	6%	-	-	-	-
	Combi onderwerp/discussie	0.4%	0.02	0%	15%	-	-	-	-
	Combi definiëring/leerstof	0.5%	0.01	0%	8%	-	-	-	-
	Combi definiëring/discussie	0.1%	0.01	0%	7%	-	-	-	-
	Combi leerstof/brainstorm	0.2%	0.01	0%	8%	-	-	-	-
Combi leerstof/discussie	0.3%	0.02	0%	12%	-	-	-	-	
Schrijffase	Hulp geven	14.6%	0.13	0%	53%	20.1%	0.15	0%	56%
	Aanmoedigen	17.6%	0.14	0%	59%	15.9%	0.13	0%	44%
	Non-instructie	4.0%	0.07	0%	30%	6.9%	0.08	0%	29%
	Transitie	1.6%	0.03	0%	12%	2.3%	0.03	0%	8%
	Combi hulp-aanmoedigen	2.4%	0.07	0%	32%	-	-	-	-
	Feedback mbt bewerkte producten	1.0%	0.03	0%	19%	0.1%	0.00	0%	2%
	Non-instructie	0.0%	0.00	0%	3%	0.0%	0.00	0%	0%
	Transitie	0.0%	0.00	0%	1%	0.0%	0.00	0%	0%

⁹⁰ Observatoren is gevraagd per scoremoment, binnen één lesfase, één score te noteren. De dataset bevatte voor de bo-lessen echter ook scoremomenten waarop twee of drie aspecten binnen hetzelfde of (in enkele gevallen) over verschillende lesfasen heen gescoord waren. Nagegaan is hoe vaak alle combinaties voorkwamen. Voor combinaties die 10 keer of vaker voorkwamen zijn combinatievariabelen aangemaakt. Dit heeft negen nieuwe combinatievariabelen opgeleverd, zoals Tabel 6.30 laat zien. Bij minder frequent voorkomende combinaties, die dus maximaal 9 keer voorkwamen, is willekeurig één van de twee (en in enkele gevallen drie) aspecten toegekend aan het scoremoment.

Tabel 6.30: Gemiddelde percentages tijd besteed aan onderdelen per lesfase, naar schooltype (vervolg)

Fase	% tijd voor..	bo				sbo			
		Gem	S.D.	Min	Max	Gem	S.D.	Min	Max
Revisiefase	Hulp geven	2.2%	0.05	0%	34%	2.8%	0.07	0%	27%
	Aanmoedigen	2.5%	0.05	0%	30%	1.3%	0.03	0%	15%
	Leerstof geven	0.7%	0.02	0%	8%	0.2%	0.01	0%	3%
	Discussie over tekstrevisie	1.0%	0.03	0%	14%	0.0%	0.00	0%	0.00
	Non-instructie	1.4%	0.07	0%	52%	0.7%	0.04	0%	19%
	Transitie	0.3%	0.01	0%	8%	0.0%	0.00	0%	0%
	Combi hulp-aanmoedigen	0.4%	0.02	0%	17%	-	-	-	-
Evaluatie-fase	Feedback/evaluatie in het algemeen	3.4%	0.06	0%	42%	7.6%	0.07	0%	33%
	Individuele feedback/evaluatie	0.0%	0.00	0%	0%	4.3%	0.08	0%	38%
	Feedback op specifieke kenmerken	0.7%	0.01	0%	6%	1.5%	0.02	0%	6%
	Peer-feedback	2.6%	0.05	0%	18%	0.3%	0.01	0%	5%
	Non-instructie	0.2%	0.01	0%	6%	0.1%	0.01	0%	3%
	Transitie	0.5%	0.01	0%	6%	0.7%	0.02	0%	7%
	Combi alg/specifieke kenmerken	0.3%	0.01	0%	9%	-	-	-	-
Publicatie	Hulp bij bewerking	0.3%	0.02	0%	15%	0.3%	0.01	0%	4%
	Feedback mbt bewerkte producten	1.0%	0.03	0%	19%	0.1%	0.00	0%	2%
	Non-instructie	0.0%	0.00	0%	3%	0.0%	0.00	0%	0%
	Transitie	0.0%	0.00	0%	1%	0.0%	0.00	0%	0%

Instructiesetting

Tabel 6.31, gebaseerd op de interviews, bevat gegevens over hoe de leerkrachten normaliter hun instructie geven in de schrijflessen. Leerkrachten konden meerdere antwoorden geven. Het merendeel geeft aan klassikaal instructie te geven met de taalmethode als uitgangspunt (bijna driekwart van de bo-leerkrachten en drie vijfde van de sbo-leerkrachten). Klassikale instructie zonder de taalmethode wordt door twee op de vijf leerkrachten toegepast (beide schooltypen). Zelfstandig werken met de instructie in de taalmethode komt slechts beperkt voor.

Een combinatie van instructie met en zonder methode komt het meeste voor. Daarnaast geven acht leerkrachten, in de tabel onder de noemer 'anders', aan dat leerlingen na de klassikale instructie vaak zelfstandig verder werken met de instructie die in de methode staat. Additionele instructie die wordt genoemd onder 'anders' betreft ook: 'verlengde instructie voor wie dat nodig heeft', 'pre-teaching', 'verdere instructie in subgroepen naar niveau (na algemeen deel)', geschreven instructie via weektaken. Een leerkracht noemt het elkaar helpen met instructie (modellering) als iets wat wordt ingezet in de instructiefase.

Tabel 6.31: Instructiesetting (interviews leerkrachten)

Setting	bo (N=99)		sbo (N=36)	
	n	%	n	%
Klassikale instructie uit de taalmethode	73	73.7%	22	61.6%
De leerlingen werken zelfstandig met de instructie in de taalmethode	7	7.1%	2	5.6%
Klassikale instructie los van de taalmethode	43	43.4%	15	41.7%
Anders	21	21.2%	11	30.6%

NB: combinaties van meerdere antwoorden mogelijk

Als aanvulling op deze vraag wordt meerdere keren genoemd dat na klassikale instructie (met methode) de leerlingen zelf verder moeten met de instructie in de methode of de opdracht. Vier leerkrachten noemen dat leerlingen na klassikale instructie (uit de methode) samen in tweetallen of in subgroepen verder werken. Soms wordt de instructie in deelgroepen gegeven (bijvoorbeeld groep 7 en 8 in een combinatieklas). Circa vijf leerkrachten vullen ook aan dat zij naast de groepsinstructie ook – vaak later in de les – nog eens individuele instructie geven.

Tijdens de observaties is ook in kaart gebracht tot welke (groep) leerling(en) de leerkracht zich tijdens de schrijfles richtte (setting/gerichtheid van de leerkracht). Het overgrote deel van de les is er in beide schooltypen sprake van een klassikale focus (Tabel 6.32), zoals de leerkrachten ook in de interviews al aangaven. In bo en sbo wordt er gemiddeld 4% en 3% van de tijd besteed aan een subgroep van leerlingen. Wanneer we de data echter bekijken, dan zien we dat het werken met een subgroep in het sbo in 28 lessen niet voorkwam (93%) en in het bo in 62 lessen (90% van de 69 lessen waarvoor informatie beschikbaar is). Daarnaast bleek dat er in drie lessen (een in het bo en twee in het sbo) nul procent van de tijd aan de hele klas werd besteed. In het bo en sbo betrof het hier ieder één klas waar 100% van de tijd aan een subgroep werd besteed. Het zou hier om combinatiegroepen kunnen gaan, waarbij de leerkracht les gaf aan groep 8 of de schoolverlaters. In de derde klas (sbo) waarin nul procent tijd aan de hele klas werd besteed, werd 78% van de tijd besteed aan individuele instructie (en over de overige 22% is niets ingevuld).

Tabel 6.32: Instructiesetting (observaties)

Setting/ gerichtheid leerkracht	bo (n=69)*				sbo (n=30)			
	Gem	SD	Min	Max	Gem	SD	Min	Max
Percentage lestijd besteed aan hele klas	84%	25.0	0%	100%	79%	28.5	0	100
Percentage lestijd besteed aan een subgroep	4%	14.0	0%	100%	3%	18.2	0	100
Percentage lestijd besteed aan individuele leerlingen	9%	20.4	0%	74%	16%	22.9	0	78

*: bij één bo-klas is geen informatie genoteerd over gerichtheid. Deze klas is buiten beschouwing gelaten. De percentages lopen niet op tot 100%, omdat setting niet voor ieder lesmoment is gescoord.

Naast de setting is de leerkracht verder gevraagd te benoemen welke aspecten hij/zij in de instructiefase van de schrijflessen belangrijk vindt (Tabel 6.33). Deze aspecten werden niet aan de leerkracht voorgelegd, maar konden door de interviewer worden gecodeerd in de interviewleidraad (zie bijlage 5A). In meer algemene zin heeft de interviewer de genoemde antwoorden (soms achteraf) gecodeerd als 'leerkracht legt nadruk op het schrijfproces' in de instructie of als 'leerkracht legt nadruk op het schrijfproduct' in zijn/haar instructie. Indien er (bijna) geen antwoorden werden gegeven of als procesmatige of productmatige aspecten in vergelijkbare mate werden genoemd is geen codering toegepast.

Tabel 6.33: Nadruk in de instructie (interviews leerkrachten)

	bo (N=99)		sbo (N=36)	
	n	%	n	%
Nadruk op het schrijfproduct	35	35.4%	17	47.2%
Nadruk op het schrijfproces	30	30.3%	8	22.2%
Beide aangekruist	3	3.0%	0	-
Geen van beide aangekruist	31	31.3%	11	30.6%

Uit Tabel 6.33 valt op te maken dat in er in het sbo in de instructiefase relatief vaker nadruk wordt gelegd op het schrijfproduct en – ten opzichte van het bo – minder op het schrijfproces. Wanneer naar de specifieke aspecten van product en proces in de instructie wordt gekeken, blijkt dat voor beide groepen leerkrachten productaspecten relatief vaker worden genoemd dan procesaspecten (Tabel 6.34). Met name aandacht voor tekstopbouw, inhoudelijke eisen van het schrijfproduct, en correct taalgebruik vinden leerkrachten belangrijk in de instructie.

Tabel 6.34: Aspecten van de instructie, die belangrijk zijn volgens leerkrachten (interviews leerkrachten)

Aspecten van het schrijfproduct	bo (N=95)	sbo (N=35)	Aspecten van het schrijfproces	bo (N=95)	sbo (N=35)
Expliciteren publiek van de tekst	11%	9%	Aandacht voor schrijfstrategieën	20%	9%
Expliciteren inhoudelijke eisen	30%	51%	Uitleg: plannen van de tekst	23%	6%
Ingaan op kenmerken tekstsoort	14%	14%	Uitleg: formuleren van zinnen	6%	6%
Aandacht voor tekstopbouw	38%	23%	Uitleg: reviseren van een tekst	6%	-
Aandacht voor correct taalgebruik	23%	17%	Uitleg: verbeteren van fouten in een tekst	6%	-
Omvang van de tekst	4%	6%	Instructie voor samen schrijven	5%	3%
Lay-out van de tekst	7%	3%	<i>modeling</i> : voordoen van bovenstaande	15%	14%
Laten zien van modelteksten	13%	9%	Instructie op basis van tekstbespreking	3%	-
Anders	23%	17%	Reflecteren op (eerdere) taakaanpak	18%	11%
			Anders	17%	14%

Gevraagd naar een toelichting op de manier waarop in de instructiefase aandacht wordt besteed aan product- en procesmatige zaken van het schrijfproduct wordt door 17 leerkrachten genoemd dat het doel van de schrijfles vooral wordt benadrukt: dat kan inhoudelijk zijn (onderwerp, eindproduct), gaan over het publiek waarvoor de leerlingen moeten schrijven, of gaan over ‘het perspectief waaruit leerlingen schrijven’. Wat betreft het schrijfproduct worden ook instrumentele zaken genoemd als voorbeeldteksten, filmpjes uit de methode, omvang tekst en spelling. Bij het schrijfproces worden voorbeelden genoemd over hoe leerlingen het beste de te nemen stappen kunnen plannen, hoe leerlingen het beste samen kunnen werken of dat leerlingen juist zo zelfstandig mogelijk de opdracht maken.

Gerichte motivatie en stimulerende woorden In de instructiefase worden door acht leerkrachten genoemd: zowel als toelichting bij het sturen op het schrijfproces als het op het schrijfproduct. Twee leerkrachten benoemen expliciet dat het belangrijk is aan te sluiten bij de belevingswereld van leerlingen in de instructiefase.

Leerdoelen

De leerkrachten is tijdens de interviews gevraagd of en op welke manier zij de doelstellingen van een schrijfles aan de leerlingen kenbaar maken. Leerlingen in het bo werken bij ruim vier vijfde (82.8%) van de leerkrachten (in minimaal twee derde van alle schrijflessen aan specifieke vastgestelde leerdoelen,

zie Tabel 6.35). In het sbo is dit 66.7%. Het deel van de leerkrachten dat aangeeft nooit aan leerdoelen per schrijfles te werken is in het sbo hoger dan in het bo (17 versus 2 procent).

Tabel 6.35: Gebruik van lesdoelen (interviews leerkrachten)

Doel	bo (N=99)		sbo (N=36)	
	N	%	n	%
Nooit	2	2.0%	6	16.7%
Weinig, in ongeveer een derde van de lessen	4	4.0%	1	2.8%
Soms, in ongeveer de helft van de lessen	11	11.0%	5	13.9%
Vaak, in ongeveer twee derde van de lessen	23	23.2%	5	13.9%
Altijd, in elke les	59	59.6%	19	52.8%

Met betrekking tot het gebruik van specifieke lesdoelen is voor ongeveer de helft van de leerkrachten een toelichting genoteerd. In verreweg de meeste gevallen gaat het om de leerdoelen die reeds in de methode (voor de specifieke les) zijn geformuleerd. Enkele leerkrachten noemen het combineren van methodedoelen met een eigen niet-methodegebonden opdracht. Anderen geven de toelichting dat zij de nadruk leggen op een specifiek element van de lesdoelen uit de methode. Een niet methodegebonden voorbeeld van het gebruik van lesdoelen is aandacht voor één specifieke schrijfgeregule per les (bijvoorbeeld gebruik hoofdletters of opbouw tekst), aandacht voor een specifieke tekstsoort ("het doel is te leren een [soort] brief op te stellen"), of het aanleren van specifieke grammaticaregels.

Ook uit Tabel 6.36 blijkt dat lesdoelen vooral via de taalmethode kenbaar worden gemaakt. Het lijkt er echter op dat sbo-leerkrachten relatief iets meer zelf de doelstellingen kenbaar maken aan het begin van de les (doelen benoemen en/of noteren op het (digi)bord). Specifieke lesdoelstellingen communiceren via de weekplanner is iets dat slechts sporadisch voorkomt.

Tabel 6.36: Manier van gebruik van lesdoelen (interviews leerkrachten)

Gebruik lesdoelen	bo (N=98)		sbo (N=32)	
	N	%	n	%
Via de taalmethode	40	40.8%	9	28.1%
Ik noem het doel, aan het begin van de les	57	58.1%	24	75.0%
Ik noteer het doel op het (digi)bord	41	41.8%	14	43.8%
Via de weekplanner	2	2.0%	0	-

NB: meerdere antwoorden mogelijk

Aanvullend wordt door vier leerkrachten aangegeven dat ze exploratief/gezamenlijk met de leerlingen de lesdoelen vaststellen: "soms vraag ik het aan de kinderen, wat zou het doel van deze les kunnen zijn". Eén leerkracht geeft aan dat hij de leerlingen de methodedoelen laat herformuleren naar een ik-doel: "na deze les kan ik...". Op één school wordt verder genoemd dat de leerdoelen per les in de laptops beschikbaar zijn (alle kinderen werken met een Chromebook in de schrijfles).

Zoals in Tabel 6.30 te zien is, wordt gemiddeld 1.4% van de geobserveerde lestijd besteed aan het benoemen van het lesdoel. Het bleek dat in 41 van de lessen (41%) leerkrachten gedurende

minimaal een tijdsblok van 20 seconden primair gericht waren op het benoemen van het lesdoel. Hier zien we een duidelijk verschil tussen sbo en bo: het bespreken van het lesdoel is in twee lessen (6.7%) in het sbo en in 39 lessen (56%) in het bo gescoord. Overigens kan uit deze resultaten niet met zekerheid worden geconcludeerd dat het lesdoel niet is genoemd in de overige lessen, waarin het lesdoel niet is gescoord. Dit kan hebben plaatsgevonden buiten de periode waarin de observator scoorde. In dergelijke gevallen heeft de leerkracht het lesdoel in ieder geval niet uitgebreid besproken.

Soorten teksten

Leerkrachten geven in de vragenlijst aan dat in de schrijflessen vooral wordt gewerkt aan een spannend of grappig verhaal, of een verslag van een gebeurtenis. Minder vaak geschreven teksten zijn dialogen en ingezonden brieven (naar bijvoorbeeld de krant). Het schrijven van een e-mail of een zakelijke brief is in het sbo minder gebruikelijk dan in het basisonderwijs. Tabel 6.37 laat verder zien dat veel soorten teksten slechts één keer per jaar in groep 8 aan bod komen in de schrijflessen. Groen gearceerde vakken in de tabel geven aan dat het overgrote deel van de leerkrachten regelmatig een tekstsoort behandelt in de schrijfles. Oranje gearceerde vakken geven aan dat het merendeel van de (sbo) leerkrachten aangeeft nooit een tekstsoort te behandelen.

Tabel 6.37: Soorten teksten en het percentage leerkrachten dat deze laat schrijven (leerkrachtvragenlijst)

	bo (N=124)			sbo (N=67)		
	Nooit	1 keer per jaar	Regelmatig	Nooit	1 keer per jaar	Regelmatig
Een spannend of grappig verhaal	-	16%	84%	6%	15%	79%
Een gedicht	2%	40%	58%	9%	48%	43%
Een dialoog	21%	56%	23%	55%	24%	21%
Verslag van gebeurtenis of activiteit	-	19%	81%	5%	28%	67%
Een oproep of uitnodiging	15%	51%	33%	33%	41%	26%
Een persoonlijke brief	10%	50%	40%	18%	58%	24%
Een zakelijke brief	11%	59%	30%	55%	31%	13%
Een instructietekst of recept	12%	53%	35%	23%	47%	30%
Een ingezonden brief	24%	56%	20%	56%	36%	8%
Een betoog ter onderbouwing van een mening	15%	37%	48%	33%	35%	32%
Een uitleg over een situatie of proces	20%	39%	41%	46%	20%	35%
Een werkstuk	2%	50%	47%	10%	57%	33%
Een e-mail of een webtekst	18%	38%	43%	39%	22%	40%

regelmatig: meerdere keren per jaar tot maandelijks.

Ten behoeve van de meerniveau-analyses (hoofdstuk 9) is voor elke leerkracht een index van de genoemde items (aandacht voor tekstsoorten) in de vorm van een gemiddelde somscore berekend (zie

Tabel 6.38).⁹¹ Deze is in het bo (M=11.48) hoger dan in het sbo (M=9.07). Leerkrachten in het basisonderwijs laten leerlingen dus meer soorten teksten (genres) schrijven dan hun collega's in het speciaal basisonderwijs. Wanneer we specifiek kijken naar de scholen in de gewichtencategorieën zien we dat er geen grote verschillen bestaan ten opzichte van het gemiddelde van 11.48.

Tabel 6.38: Door leerlingen geschreven soorten teksten (index tekstsoorten, gemiddelde somscore - leerkrachtvragenlijst)

	N	M	SD	Min	Max
Basisonderwijs	115	11.48	1.91	5	13
S1 (0%)	7	12.29	0.95	11	13
S2 (max 10%)	79	11.19	2.08	5	13
S3 (10-25%)	18	12.11	1.27	9	13
S4 (>25%)	11	12.00	1.61	8	13
Speciaal basisonderwijs	64	9.08	2.63	2	13

Informatieoverdracht en publiek

De leerkrachten is vervolgens gevraagd welke doelstellingen zij nastreven met de ingezette opdrachten voor het schrijfonderwijs. Met andere woorden: welk communicatieve doelen worden nagestreefd met de door leerlingen geschreven teksten. Tabel 6.39 bevat een overzicht van de algemene doelen die dienen te worden bereikt met de leerlingteksten. Het grootste deel van de leerkrachten laat leerlingen teksten schrijven die gericht zijn op alle doelen. 'Een verzoek doen aan de lezer' en 'De lezer aanwijzingen geven hoe iets moet worden uitgevoerd' zijn echter tekstdoelen die sbo-leerlingen niet (of amper) hoeven te bereiken. Wanneer de frequentie van de items dichotoom gescoord worden (0=nooit aandacht voor, 1=wel aandacht voor) zien we dat één (bo) respectievelijk vijf (sbo) leerkrachten aangeven leerlingen geen teksten te laten schrijven waarmee zij de genoemde schrijfdoelen moeten bereiken. De gemiddelde somscore van de genoemde doelstellingen bedraagt 4.53 voor het bo en 3.75 voor het sbo (min=0, max=5). Dit betekent dat in het bo leerlingen teksten met meer verschillende (beoogde) schrijfdoelen schrijven dan in het sbo.

⁹¹ Hierbij is een onderscheid gemaakt tussen 0=nooit en 'pas ik toe'=1

Tabel 6.39: Doelen die leerlingen met de schrijfp opdrachten moeten bereiken (index soorten doelen, gemiddelde somscore - leerkrachtvragenlijst)

Schrijfdoel		N	Nooit	één keer per jaar - regelmatig*
De lezer amuseren	Bo	125	1.6%	98.4%
	Sbo	66	19.7%	80.3%
De lezer informeren door iets te beschrijven	Bo	125	2.4%	97.6%
	Sbo	67	1.9%	98.1%
De lezer aanwijzingen geven hoe iets moet worden uitgevoerd	Bo	125	13.6%	86.4%
	Sbo	66	28.4%	71.6%
Een verzoek doen aan de lezer	Bo	125	18.4%	81.6%
	Sbo	67	36.4%	63.6%
De lezer overtuigen (met standpunten en argumenten)	Bo	124	13.7%	86.3%
	Sbo	67	18.4%	81.6%
		N	M	S.D.
Gemiddelde somscore 'doelen van teksten'	Bo	120	4.53	.88
	Sbo	65	3.75	1.38

* een keer per jaar, per kwartaal of minimaal 1 keer per maand

Tijdens de **lesobservaties** hebben de observatoren genoteerd welke tekstsoorten (met bijbehorende doelen) centraal stonden in de les, zoals weergegeven in Tabel 6.40. Zoals te verwachten op basis van de antwoorden van de leerkrachten over tekstsoorten en doelen in de vragenlijst (Tabellen 6.37 en 6.39) werden de leerlingen in beide schooltypen het vaakst gevraagd een verhalende tekst te schrijven, bedoeld om de lezer te amuseren, en ook informerende teksten werden door de leerlingen relatief vaak geschreven. Beduidend minder vaak kwamen persuasieve en argumentatieve teksten aan bod. In twee geobserveerde lessen in het sbo kwamen poëtische teksten aan de orde; in het bo was dit niet het geval.

Tabel 6.40: Soorten teksten die in de lessen centraal stonden - lesobservaties

Tekstsoort	bo (N=60) ⁹²		sbo (N=30)	
	n	%	n	%
Verhalend	27	45%	16	53%
Informerend	22	37%	7	23%
Mening beïnvloedend	7	12%	3	10%
Mening verdedigend	4	7%	-	-
Poëtisch	-	-	2	7%
Naar eigen keuze	-	-	2	7%

⁹² Voor 10 lessen was niet genoteerd welke tekstsoort centraal stond.

In de **lesobservaties** is ook gekeken naar kenmerken van de opdrachten die werden gegeven (Tabel 6.41). Op voorhand waren er zes categorieën gedefinieerd. Drie van deze categorieën gaven kenmerken weer van (relatief) vrije taken. Het ging om taken die vrijheid kenden (1. vrij schrijven, 2. schrijven over eigen ervaringen en 3. gebruik van eigen voorkennis). In een ruim kwart van de lessen mochten leerlingen vrij schrijven, zonder gerichte opdracht en in het sbo mochten leerlingen in één op de vijf lessen schrijven over hun eigen ervaringen. In (ruim) een derde van de lessen werd daarentegen aan één specifiek onderwerp gewerkt. Gebruikmaking van een schrijfschema kwam aanzienlijk minder voor. Observatoren hebben zowel in bo als sbo enkele extra categorieën toegevoegd, zoals simulatie van alledaagse functies (waarbij leerlingen vermoedelijk vanuit een bepaalde functie een tekst moesten schrijven), deel van een tekst (waarbij b.v. enkel aan inleiding of slot werd gewerkt) en situaties waarin leerlingen mochten kiezen uit diverse onderwerpen om de tekst over te schrijven.

Tabel 6.41: Kenmerken van de schrijfoopdrachten die aan leerlingen zijn gegeven - lesobservaties

		bo (N=68) ⁹³		sbo (N=30)	
		N	%	n	%
Vrije opdrachten	Vrij schrijven	20	29%	5	17%
	Eigen ervaringen	6	9%	6	20%
	Weinig voorkennis	-	-	-	-
Richtinggevende opdrachten	Eén onderwerp	27	40%	10	33%
	Zinnen en alinea's volgens instructie	-	-	2	7%
	Volgens schrijfschema	4	6%	3	10%
Additionele labels	Simulatie alledaagse functies	-	-	1	3%
	Deel van een tekst	4	6%	3	10%
	Keuze uit diverse onderwerpen	7	10%	-	-

Een communicatief effectieve tekst is niet enkel gericht op het bereiken van een schrijfdoel, maar houdt ook rekening met het lezerspubliek. Voor leerlingen is het belangrijk een publiek voor ogen te houden aan wie de tekst gericht is. De leerkrachten is daarom ook gevraagd hoe vaak leerlingen teksten moeten schrijven voor een specifieke lezer/lezerspubliek (Tabel 6.42).

In het bo en sbo schrijven leerlingen vaker een tekst aan bekenden dan aan onbekenden. Bo-leerlingen schrijven vaker teksten aan onbekenden dan in het sbo. Teksten voor vertegenwoordigers van officiële instanties of bedrijven worden in beide onderwijstypen het minst geschreven. Dit geldt vooral voor het sbo, waar bijna 80% van de leerkrachten aangeeft dat hun leerlingen nooit een tekst aan deze relatief abstracte instanties te schrijven. Driekwart van de bo-leerkrachten en 62% van de sbo-leerkrachten laat leerlingen minimaal een keer per jaar teksten schrijven waarbij het publiek niet nader gespecificeerd is.

Wanneer de frequentie van de items dichotoom gescoord wordt (0=niet laten schrijven, 1=wel laten schrijven in lesjaar) zien we dat één (bo) respectievelijk vijf (sbo) leerkrachten aangeven dat leerlingen voor geen enkele van de gespecificeerde publieksoorten teksten hoeven te schrijven. De gemiddelde somscore van de genoemde publieksvarianten bedraagt 3.97 voor het bo en 2.73 voor het

⁹³ Voor 2 lessen was niet genoteerd welke kenmerken de opdracht had.

sbo. Het lijkt erop dat in het bo aan een meer divers publiek aandacht wordt besteed in de schrijflessen dan in het sbo.

Tabel 6.42: Beoogd lezerspubliek in de geschreven teksten (index soorten publiek, gemiddelde somscore – leerkrachtvragenlijst)

		N	Nooit	één keer per jaar - regelmatig*
Bekende leeftijdsgenoten (<i>klasgenoten, vriend(inn)en</i>)	bo	124	1.6%	98.4%
	sbo	65	12.3%	87.7%
Onbekende leeftijdsgenoten	bo	123	26.8%	73.2%
	sbo	65	55.4%	44.6%
Bekende volwassenen (<i>bijv. familieleden</i>)	bo	124	12.9%	87.1%
	sbo	66	15.2%	84.8%
Onbekende volwassenen	bo	123	26.8%	73.2%
	sbo	65	66.8%	33.2%
Vertegenwoordigers van officiële instanties of bedrijven (<i>bijv. politici, journalisten, managers</i>)	bo	125	36.0%	64.0%
	sbo	66	78.8%	31.2%
Niet gespecificeerd (<i>er staat niet in de opdracht voor welke lezers geschreven wordt</i>)	bo	119	25.2%	74.8%
	sbo	66	37.9%	62.1%
		N	M	SD
Gemiddelde somscore 'diversiteit publiek'	bo	122	3.97	1.18
	sbo	62	2.73	1.46

* een keer per jaar, per kwartaal of minimaal 1 keer per maand

Ook tijdens de **lesobservaties** dienden de observatoren te noteren of het publiek was gespecificeerd en zo ja, voor welk publiek de tekst diende te worden geschreven. In Tabel 6.43 staan de resultaten weergegeven. Het blijkt dat in het bo het publiek ongeveer even vaak nabij wordt gedefinieerd (andere leerlingen in de klas of de eigen leerkracht) als in de wereld daarbuiten (mensen buiten de klas). Ook is er in het bo in een bijna een kwart van de lessen geen publiek gedefinieerd, terwijl dit percentage in het sbo zelfs ruim 60% is. In bijna één op de vijf lessen schrijven sbo-leerlingen een tekst voor de wereld buiten de klas.

Tabel 6.43: Schrijven voor een al dan niet gespecificeerd publiek - lesobservaties

	bo (N=55) ⁹⁴		sbo (N=29)	
	n	%	n	%
Zonder specifiek publiek	12	22%	18	62%
Algemeen publiek	11	20%	3	10%
De schrijver zelf	4	7%	-	-
Mensen in de klas	13	24%	3	10%
Mensen buiten de klas	15	27%	5	17%

⁹⁴ In het bo was voor 15 lessen niet genoteerd of er een publiek was gespecificeerd en in het sbo voor 1 les.

De uitkomsten over de soorten teksten, opdrachtkenmerken en publieksoriëntatie in de geobserveerde schrijflessen in het bo kunnen we vergelijken met de bevindingen van Van Gelderen en Blok, zo'n 20 jaar geleden. Zij concluderen op basis van zowel hun observaties als interviews: 'we kunnen stellen dat de grote meerderheid van de stelopdrachten bestaat uit het schrijven van verhalende teksten over een gegeven onderwerp, zonder indicaties over het lezerspubliek' (Van Gelderen & Blok, 1991, p. 167). Hoewel verhalende teksten (over een specifiek onderwerp) nog steeds het meest worden geschreven door de bo-leerlingen, lijkt er wat betreft de specificatie van het publiek wel wat veranderd. Waar Van Gelderen en Blok in 81% van de geobserveerde lessen zagen dat er geen publiek gespecificeerd was voor de schrijfoopdracht is dat in de binnen deze peiling nog maar het geval in 27% van de schrijfoopdrachten in het bo. Daarnaast is ook het schrijven van informerende teksten in deze peiling vaker gesignaleerd (37% van de lessen in plaats van 26% bij Van Gelderen en Blok).

Kennis van teksten en schrijfprocessen

Kennis van soorten teksten en over schrijfprocessen is bevorderend voor de ontwikkeling van schrijfvaardigheid. Het is dus van belang dat een leerkracht in de schrijfles hier aandacht aan besteedt. Tabellen 6.44 en 6.45 bevatten de frequenties van verschillende kennisaspecten over het schrijven van een tekst (tekstsoort, passende woorden, formuleren en taalverzorging) en kennis over schrijfprocessen (informatieverzameling, plannen, zinnen schrijven, reviseren).

Ongeveer twee derde van de leerkrachten in het bo én sbo besteedt minimaal eens per maand aandacht aan verschillende tekstsoorten. De overige leerkrachten doen dat vier keer per jaar of minder (inclusief een zeer beperkt aantal leerkrachten dat aangeeft er 'nooit' aandacht aan te besteden). 'Passende woorden' voor een tekst komt bij ongeveer de helft van de leerkrachten minimaal eens per maand aan bod. Aan taalverzorgingsaspecten als interpunctie wordt zowel in bo als sbo veel aandacht besteed. In beide onderwijstypen rapporteren circa 80% van de leerkrachten dat zij er minimaal 1 keer per maand aandacht aan besteden.

Tabel 6.44: Aandacht voor kennis over teksten, in percentages (leerkrachtvragenlijst)

		N	1 keer per kwartaal of minder*	minimaal een keer per maand
Kennis over verschillende tekstsoorten (<i>bijv. verhalen, informatieve teksten, betogen</i>)	bo	125	36.0%	64.0%
	sbo	66	33.3%	66.7%
Kennis over passende woorden voor een tekst (<i>bijv. signaalwoorden of formeel taalgebruik</i>)	bo	125	50.0%	50.0%
	sbo	66	47.0%	53.0%
Kennis over het formuleren van zinnen in een tekst	bo	124	59.2%	40.8%
	sbo	65	51.5%	48.5%
Kennis over het gebruik van leestekens en hoofdletters in een tekst	bo	125	16.1%	83.9%
	sbo	65	21.2%	78.8%

* nooit, een enkele keer of '1 keer per kwartaal'

Tabel 6.45: Aandacht voor kennis over schrijfprocessen (leerkrachtvragenlijst)

		N	1 keer per kwartaal of minder*	minimaal een keer per maand
Hoe je informatie verzamelt voor een tekst.	bo	126	60.8%	39.2%
	sbo	66	50.0%	50.0%
Hoe je een plan maakt voor een tekst.	bo	125	72.6%	27.4%
	sbo	66	78.8%	21.2%
Hoe je goede zinnen maakt voor een tekst.	bo	126	62.7%	37.3%
	sbo	66	50.0%	50.0%
Hoe je een tekst herziet en verbetert.	bo	125	65.9%	34.1%
	sbo	66	78.8%	21.2%
Hoe je feedback op je tekst verwerkt.	bo	125	69.8%	30.2%
	sbo	66	78.8%	21.2%
Hoe je een tekst geschikt maakt voor het beoogde lezerspubliek.	bo	124	75.8%	24.2%
	sbo	66	89.4%	10.6%
		N	M	SD
Schaalscore 'mate van aandacht voor schrijfproces'	bo	125	3.02	.58
	S1	7	3.29	.47
	S2	85	2.95	.54
	S3	22	3.13	.67
	S4	12	3.22	.62
	sbo	66	2.84	.46

* nooit, een enkele keer of '1 keer per kwartaal'

Tabel 6.45 bevat de mate waarin leerkrachten aandacht besteden aan kennis over schrijfprocessen. Leerkrachten besteden relatief de meeste aandacht aan kennis over het verzamelen van informatie en over het schrijven van goede zinnen kunnen schrijven. Aan kennis over het geschikt maken van een tekst voor een bepaald publiek wordt de minste aandacht besteed. De genoemde items zijn eveneens omgezet in een schaalscore, waarbij de vier antwoordcategorieën telkens gescoord zijn van 1 (nooit) tot 4 (minimaal maandelijks). Gemiddeld gezien besteden leerkrachten een keer per maand aandacht aan kennis over (aspecten van) het schrijfproces in de schrijfles.

Inzet leeractiviteiten

De leerkrachten van groep 8 en van schoolverlaters is verder gevraagd welk soort activiteiten ze inzetten in de schrijflessen en in welke mate ze dit doen (Tabel 6.46). In deze tabel zijn de items behorend bij de drie schalen *prewriting activiteiten*, *revisie activiteiten* en *leren van anderen* respectievelijk lichtblauw, groen en geel gearceerd weergegeven. Wat betreft de prewriting activiteiten zien we dat de meeste leerkrachten leerlingen de tijd geven om informatie te verzamelen en/of een schrijfplan te maken voorafgaande aan het daadwerkelijke schrijven van de tekst. Een schrijfplan maken als activiteit is in het sbo echter minder gebruikelijk. Leerkrachten in het sbo lichten echter wel vaker de schrijfaanpak mondeling toe.

Qua revisie activiteiten zien we dat dat bo-leerkrachten gemiddeld vaker feedback laten verwerken en de eerste versie van een tekst laten herlezen en herzien dan sbo-leerkrachten. De schaal revisie-activiteiten bevat tevens items die gericht zijn op de publicatie van de geschreven tekst, te weten de mate waarin ze leerlingen een versie van een geschreven tekst voor een beoogd publiek laten

verzorgen, of waarin de geschreven teksten ook daadwerkelijke door het beoogde publiek laten lezen. Relatief gezien spenderen de bo- en sbo-leerkrachten hier weinig aandacht aan.

Activiteiten die vallen onder de schaal 'leren van anderen' zijn gericht op modellering, waarbij bijvoorbeeld (goede) voorbeeldteksten kunnen worden ingezet. Dit bespreken van voorbeeldteksten komt in zowel bo als sbo meer voor dan het klassikaal bespreken van de geschreven opdrachten van leerlingen. Dit laatste wordt door een op de vijf leerkrachten vaker dan één keer per maand gedaan. De activiteit die het minst vaak voorkomt in de lessen is dat leerlingen aan elkaar hun schrijfproces *modelleren*, oftewel expliciet aan elkaar voordoen en dit bespreken. Het overgrote deel van zowel bo- als sbo-leerkrachten geeft aan dit één keer per kwartaal of minder aan bod komt, waarvan 66% (bo) en 76% (sbo) aangeeft dat leerlingen nooit aan andere leerlingen te laten voordoen hoe je moet schrijven. De leerlingen zijn in de interviews ook gevraagd naar modelleren; het observeren van andere leerlingen die voordoen hoe te schrijven.

Tabel 6.46: Activiteiten rondom schrijfopdrachten en frequentie van inzet (leerkrachtvragenlijst)

Activiteiten	bo		sbo	
	1 keer per kwartaal of minder*	minimaal een keer per maand	1 keer per kwartaal of minder*	minimaal een keer per maand
Pre-schrijven: informatie verzamelen	54.8%	45.2%	62.1%	37.9%
Pre-schrijven: een schrijfplan maken	68.3%	31.7%	83.3%	16.7%
Mondeling toelichten van de schrijfaanpak door leerkracht	69.4%	30.6%	56.1%	43.9%
Schrijven a.d.h.v. een concreet leerdoel	55.2%	44.8%	66.7%	33.3%
Eerste versie herlezen	50.8%	49.2%	78.1%	21.9%
Leerlingen feedback laten verwerken	61.0%	39.0%	76.6%	23.4%
Herschrijven van de eerste versie	61.9%	38.1%	81.8%	18.2%
Verzorgen van de tekst t.b.v. publicatie voor publiek	79.7%	20.3%	83.1%	16.9%
Leerlingen laten hun tekst lezen door beoogd publiek	88.8%	11.4%	89.4%	10.6%
Klassikaal bespreken van voorbeeldteksten	63.5%	36.5%	68.2%	31.8%
Klassikale bespreking teksten leerlingen	78.6%	21.4%	81.5%	18.5%
De leerlingen modelleren het schrijfproces voor elkaar	96.8%	3.2%	92.4%	7.6%

* nooit, een enkele keer of '1 keer per kwartaal'; nb: n=124-126 (bo) en n= 65-66 (sbo)

Vergeleken met de antwoorden op de vragenlijsten laten de **lesobservaties** zien dat er in respectievelijk 13% en 12% van de lessen is gescoord dat leerlingen werkten aan een schrijfplan of aan het verzamelen van informatie. Deze activiteiten zijn dus in weinig lessen geobserveerd. Als de prewriting activiteit ‘informatie verzamelen’ ruimer wordt opgevat en ook andere activiteiten bevat waarbij leerlingen betrokken zijn, zoals brainstorm of discussie over het onderwerp of een speciale voorbereiding, dan zien we dat hieraan in 70 lessen aandacht wordt besteed (70%). Wanneer we kijken in hoeverre de geobserveerde lessen aandacht voor revisie, feedback geven/evaluatie en publicatie bevatten, dan zien we dat er in 47% van de lessen revisie-activiteiten zijn gescoord (range 0-55% van de lestijd), dat er in 70% van de lessen (enige) aandacht was voor evaluatiefase (range 0-55% van de lestijd) en dat er in 13% van de lessen (enige) aandacht was voor het publiceren van de tekst (range 1-19% van de lestijd).

De schaal prewriting activiteiten geeft de schaalscore op de vier blauw gearceerde items weer (Tabel 6.47). Deze schaalscore loopt van minimumscore 1 tot maximumscore 4⁹⁵ en de gemiddelde schaalscore in het bo 3.14 en in het sbo 3.00. laten zien dat leerkrachten gemiddeld (iets meer dan) een keer per kwartaal voorbereidende activiteiten (laten) uitvoeren. De gemiddelde schaalscore van de oranje gearceerde items (behorend bij de schaal revisie activiteiten) in het bo en sbo ($M_{bo}=2.92$; $M_{sbo}=2.38$) laten zien dat leerkrachten deze activiteiten gemiddeld tussen één keer per jaar en één keer per kwartaal (laten) uitvoeren, waarbij dit vaker het geval is bij de bo-leerkrachten. De laatste gemiddelde schaalscore is gemaakt voor de mate van activiteiten in de les waarbij leerlingen leren van anderen, in Tabel 6.46 grijs gearceerd. De gemiddelde schaalscore liggen voor het bo ($M=2.49$) en het sbo ($M=2.41$) dicht bij elkaar: leerkrachten in beide typen onderwijs rapporteren dit tussen één keer per jaar en één keer per kwartaal te doen.

Tabel 6.47: Schaalscores activiteiten rondom schrijfopdrachten (leerkrachtvragenlijst)

	Schooltype (n)	M	SD
Schaalscore ‘prewriting’	bo (126)	3.14	.59
	sbo (66)	3.00	.65
Schaalscore ‘revisie’	bo (125)	2.92	.66
	sbo (66)	2.38	.88
Schaalscore ‘leren van anderen’	bo (125)	2.49	.68
	sbo (66)	2.41	.73

Hoewel de gehanteerde lesfasen in de **lesobservaties** niet helemaal vergelijkbaar zijn, laten de resultaten zien dat in alle lessen aandacht werd besteed aan voorbereiding, hetzij aan het bespreken van lesdoel en lesoverzicht, van het onderwerp, het geven van uitleg en/of voorbereidende activiteiten waarbij leerlingen een actieve rol hebben. Aandacht voor revisie werd, zoals hierboven beschreven, in 47% van de lessen revisie-activiteiten gescoord (range 0-55% van de lestijd). Zowel de lesobservaties als de schaalscores laten zien dat revisie-activiteiten blijkbaar minder vaak onderdeel van de schrijfles vormen.

Ook de leerlingen is in de interviews gevraagd in hoeverre modelering door de leerkracht, het ‘voordoen door meester of juf hoe je moet schrijven’, voorkomt in de schrijflessen (Tabel 6.48). 18

⁹⁵ Dit geldt ook voor de schalen revisie-activiteiten en leren van anderen.

procent van de bo-leerlingen en maar liefst 42 procent van de sbo-leerlingen geeft aan dat zijn of haar leerkracht niet modelleert tijdens de schrijflessen. In 29% (bo) en 21% (sbo) van de gevallen geven leerlingen aan dat de juf/meester dat juist vaak doet. Op de vraag wat de leerkracht dan voordoet (laat zien) komen uiteenlopende (combinaties) van antwoorden. In het bo gaat het vooral om het verzamelen van ideeën en informatie, het maken van een tekstplan en het formuleren van zinnen, terwijl in het sbo het formuleren van zinnen er het meest uitspringt.

Tabel 6.48: Modelering (voordoen) door leerkracht, volgens leerlingen (interviews leerlingen)

Frequentie	bo (N=198)		sbo (N=72)	
	n	%	n	%
Nooit	36	18%	30	42%
Soms	105	53%	27	38%
Vaak	57	29%	15	21%
Wat laat juf/meester dan zien?	n	%	n	%
Hoe je ideeën kunt verzamelen/ordenen	60	30%	11	15%
Hoe je een tekstplan maakt	37	19%	4	6%
Hoe je goede zinnen formuleert	48	24%	22	31%
Hoe je alineaverbanden laat zien	4	2%	0	-
Hoe je een inleiding schrijft	23	12%	3	4%
Hoe je een slot schrijft	10	5%	4	6%
Hoe je je eigen tekst kunt verbeteren	22	11%	5	7%
anders	36	18%	14	19%

De categorie anders bevat circa 15 antwoorden die lastig of niet zijn te interpreteren. Voorbeelden die wel duidelijk zijn: 'het geven van een voorbeeld of een voorbeeldtekst (zes keer genoemd), het voordoen van het juiste op woord-, zins- en alinea-niveau (spelling, interpunctie, opbouw tekst – ongeveer 20 keer genoemd) en variaties op het "laten zien wat je moet doen als je vastloopt" (acht keer genoemd).

Samenwerken

Met behulp van de vragenlijst is gevraagd naar de frequentie waarop leerlingen samenwerken in de verschillende fasen van het schrijven (plannen, feedback geven en reviseren). Bijna 40% van de bo-leerkrachten rapporteert dat de leerlingen minimaal eens per maand samenwerken door elkaar van feedback te voorzien. Samenwerking bij het plannen van de tekst en bij de revisie zegt (ruim) een kwart van de bo-leerlingen te doen. In het sbo vindt samenwerking het vaakst in de planningsfase plaats (bij

bijna één op de vijf leerkrachten) (Tabel 6.49). Wanneer we de items omzetten in een schaalgemiddelde (1= nooit, 2= een enkele keer, 3=een keer per kwartaal, 4=is minstens een keer per maand) dan zien we dat de gemiddelde scores voor samenwerken in zowel bo als sbo minder dan 3 is. Gemiddeld gezien wordt er dus minder dan één keer per kwartaal samengewerkt in de schrijfles.

Tabel 6.49: Samenwerken tijdens de schrijfopdracht per fase (leerkrachtvragenlijst)

		een keer per kwartaal of minder	een keer per maand of meer
Leerlingen werken samen bij het plannen van een tekst (<i>b.v. schrijfplan</i>).	bo (124)	72.0%	28.0%
	sbo (65)	81.8%	18.2%
Leerlingen geven elkaar feedback op elkaars tekst	bo (124)	60.8%	39.2%
	sbo (64)	83.6%	16.4%
Leerlingen werken samen bij het reviseren van de tekst	bo (124)	75.2%	24.8%
	sbo (64)	88.1%	11.9%
	Schooltype	M	SD
Schaalscore 'samenwerken tussen leerlingen'	Bo (124)	2.94	.73
	Sbo (64)	2.36	.94

Bij de *lesobservaties* diende de observator te noteren gedurende hoeveel tijd de leerlingen samenwerkten. In 36 lessen (51%) in het bo werkten de leerlingen samen, terwijl dit in zes lessen (20%) in het sbo voorkwam. In de lessen waarin samenwerking tussen leerlingen is gesignaleerd, werkten leerlingen in het bo gemiddeld 44% van lestijd samen, terwijl dit in het sbo gemiddeld gedurende 49% het geval was (Tabel 6.50). De duur van de samenwerking verschilde echter sterk per les.

Tabel 6.50: Gemiddelde percentage lestijd waarin leerlingen samenwerken

		gem	sd
% tijd waarin leerlingen samenwerken	bo (36)	43.6%	30.9
	sbo (6)	48.8%	37.0

In de interviews is nader gesproken over de momenten in de schrijfles waarop leerlingen samenwerken en op welke manieren zij dat doen. Tabel 6.51 geeft inzage in de genoemde antwoorden.⁹⁶

⁹⁶ Bij deze vraag komt het voor dat wel de hoofdcategorie is aangekruist (bijv. 'publiceren') maar dat vervolgens geen van de aspecten binnen die categorie is aangekruist.

Tabel 6.51: Manieren om samen te werken, per schrijffase (interviews leerkrachten)

Activiteit		bo (N=99)		sbo (N=36)	
		n	%	n	%
Plannen		74	75%	24	67%
	Ideeën genereren via gesprek in duo/groepje	52	70%	11	46%
	Samen informatie verzamelen	24	32%	8	33%
	Samen een tekstplan maken	22	30%	4	17%
	Observeren hoe een andere leerling werkt	1	1%	0	-
	Anders	2	3%	1	4%
Formuleren (schrijven)		58	59%	23	64%
	Samen één tekst formuleren	47	81%	12	52%
	Elkaar helpen met taalverzorging	12	21%	7	30%
	Anders	4	7%	1	4%
Feedback en reviseren		76	77%	19	53%
	Schriftelijk feedback geven op elkaars tekst	19	25%	0	-
	Mondeling feedback geven op elkaars tekst	51	67%	11	58%
	Tekst bespreken in duo/groepje	21	28%	2	11%
	Samen een tekst reviseren	13	17%	2	11%
	Anders	2	3%	0	-
Publiceren		33	33%	15	42%
	Verzorgen	9	27%	2	13%
	Samen lay-out verzorgen	8	24%	3	20%
	Samen muurkrant maken	11	33%	4	27%
	Anders	9	27%	1	7%

NB: per fase konden meerdere activiteiten aangekruist worden

In de planningsfase van de schrijfoopdracht zegt drie kwart (n=74) van de bo-leerkrachten en twee derde van de sbo-leerkrachten (n=24) de leerlingen wel eens samen te laten werken. Het gaat vooral om leerlingen samen ideeën laten genereren (“brainstormen”) of samen daadwerkelijk informatie te verzamelen. Zoals ook uit de vorige tabel bleek is het samenwerken aan een schrijfplan in het bo vaker aan de orde dan in het sbo. Het observeren van andere leerlingen in deze fase wordt door slechts één leerkracht genoemd. Het laten maken van een taakverdeling in het plannen en informatie verzamelen voor een gezamenlijk werkstuk (voor bijv. wereldoriëntatie) is een voorbeeld die een leerkracht geeft om leerlingen (goed) samen te laten werken.

Voor de fase dat leerlingen de (eerste versie) van hun tekst moeten schrijven, geeft bijna drie vijfde van de leerkrachten in het bo aan dat wel eens samengewerkt, met name bij het formuleren van de tekst. In het sbo wordt in ruim drie vijfde van de gevallen wel eens samengewerkt in deze schrijffase. Wanneer gevraagd wordt naar een specificatie noemen de sbo-leerkrachten echter maar in een derde van de gevallen dat dit gebeurt tijdens het formuleren. Elkaar helpen bij de taalverzorging van de tekst wordt veel minder genoemd door beide groepen leerkrachten. De toelichtingen die de leerkrachten aanreiken voor samenwerking in deze fase van het onderwijs verwijzen vaak naar de frequentie en de intensiteit: “de hele tekst samen schrijven”, “Ieder schrijft een eigen stukje, dat wordt samengevoegd en dan kijken we gezamenlijk of het goed is”. Samenwerking in deze fase is soms ook gekoppeld aan het type opdracht, bijvoorbeeld bij *Nieuwsbegrip*, waar in tweetalen aan de opdracht gewerkt dient te worden, bijvoorbeeld bij het maken van de samenvatting van het nieuwsbericht.

Samenwerken tijdens feedback en reviseren wordt door iets meer dan drie kwart (bo) en ruim de helft van de sbo-leerkrachten genoemd. Leerlingen geven elkaar met name mondeling feedback op geschreven teksten. Het schriftelijk feedback geven of het bespreken van de tekst in een groepje komen in het sbo niet of nauwelijks voor. Een op de vier leerkrachten in het bo geeft echter aan dat ze opdrachten wel eens door leerlingen van feedback laat voorzien. Het gezamenlijk aanpassen van de eerste versie van een tekst gebeurt ook heel weinig. Dertien, respectievelijk twee leerkrachten uit het basis- en speciaal basisonderwijs geven aan dit wel eens toe te passen. Uit de toelichting blijkt dat dit ook vooral gebeurt (of dient te gebeuren) wanneer de schrijfpodracht een gezamenlijke was, bijvoorbeeld het samen maken van een werkstuk, een inzending voor de schoolkrant of een gezamenlijk verslag van een schoolgebeurtenis: leerlingen presenteren bijvoorbeeld “in tweetallen wat ze samen hebben gedaan”.

Het lijkt erop dat in de fase waarin de geschreven teksten publicabel dienen te worden gemaakt, relatief minder wordt samengewerkt. In het bo zegt ongeveer één op de drie leerkrachten dat dit plaatsvindt, en in het sbo is dat ruim 40 procent. Gevraagd naar onderdelen/aspecten van samenwerken in deze fase geven niet alle leerkrachten voorbeelden. Het meeste genoemd in beide schooltypen is het gezamenlijk maken van een Muurkrant.

Uit de gehele set interviews wordt duidelijk dat voor elk van de vier fasen in het schrijfproces steeds een kwart tot twee derde van de leerkrachten aangeeft dat samenwerken niet voorkomt in die fase van de les. Bij publiceren is dat het meest (66% bo, 58% sbo). Uit negen (bo, 9%) en vier (sbo, 11%) interviews wordt duidelijk dat de leerkracht de leerlingen helemaal niet laat samenwerken in de schrijflessen. Gevraagd naar de redenen waarom, geven twee leerkrachten aan daar geen tijd voor te hebben, vijf leerkrachten geven aan geen prioriteit aan samenwerken te hechten (of “er niet aan gedacht hebben”); twee benoemen dat samenwerken ontbreekt in de methode en drie leerkrachten duiden op de (te) grote niveaoverschillen in de klas: “samenwerken is daarvoor niet geschikt”.

In de gesprekken met leerlingen is eveneens aandacht besteed aan samenwerken. Tabel 6.52 bevat de frequenties van het aantal keren dat leerlingen zeggen samen te werken en de momenten in de schrijfles waarop ze dat doen. Het percentage leerlingen dat zegt helemaal nooit samen te werken, respectievelijk 28% (bo-leerlingen) en 36% (sbo-leerlingen), is in relatie tot het aantal leerkrachten dat aangeeft nooit samen te laten werken in de schrijfles, opvallend hoog te noemen. Leerlingen die soms of vaak samenwerken in de schrijfles geven aan dat dit met name plaatsvindt aan het begin, wanneer ze ideeën moeten verzamelen (maar veelal niet bij het maken van een schrijfplan), en tijdens het schrijven waar ze samen zinnen formuleren, of wanneer ze medeleerlingen hulp bieden bij het schrijven. Leerlingen noemen vergelijkbare onderdelen als de leerkrachten waarbij ze samenwerken. Het verschil in de mate waarin leerkrachten en leerlingen zeggen samen te (laten) werken is niet goed te verklaren. Leerlingen lijken zich in ieder geval weinig bewust te zijn van eventuele samenwerking tijdens de schrijflessen. Samenwerking lijkt daarmee in ieder geval geen structureel onderdeel van de schrijfles te zijn. De lesobservaties lijken deze veronderstelling te bevestigen, omdat er in de helft van de bo-lessen en in een vijfde van de sbo-lessen sprake was van samenwerking tussen leerlingen .

Tabel 6.52: Samenwerken volgens leerlingen (interviews leerlingen)

Frequentie	bo (N=199)		sbo (N=72)	
	n	%	n	%
Nooit	55	28%	26	36%
Soms	105	53%	31	43%
Vaak	34	17%	15	21%
Onbekend	5	3%	0	-
Momenten*	n	%	n	%
Aan het begin, om ideeën te verzamelen	67	34%	12	17%
Aan het begin, om een tekstplan/schrijfplan te maken	19	10%	3	4%
Tijdens het schrijven om samen zinnen te formuleren	42	21%	12	17%
Tijdens het schrijven om elkaar hulp te geven	40	20%	18	25%
Na het schrijven, om de tekst te verbeteren	34	17%	7	10%
Anders	25	13%	14	20%

* NB: meerdere antwoorden mogelijk

Motivatatiebevordering

Het samenwerken/samen laten schrijven werkt volgens enkele geïnterviewde leerkrachten extra motiverend in de schrijfles. Uit de resultaten van de vraag in de interviews hoe leerkrachten op een motiverende wijze hun leerlingen activeren in de schrijfles, blijkt dat dit slechts door een zeer beperkt deel (6%) van de leerkrachten naar voren is gebracht. Aansluiting bij de beleavingswereld van leerlingen is – net als bij opdrachten in andere domeinen van het onderwijs – volgens de meeste leerkrachten essentieel om de motivatie voor de schrijfles/de schrijfopdracht te vergroten. Creativiteit en de inzet van een gevarieerd palet aan werkvormen worden ook genoemd. Een belangrijk verschil in de motivatiebevordering van bo- en sbo-leerkrachten lijkt te zitten in de authenticiteit van het publiek en het doel waarvoor leerlingen moeten schrijven. Een vijfde van de bo-leerkrachten vindt dit belangrijk, terwijl in het sbo slechts enkele leerkrachten dit benoemen. Tabel 6.53 geeft een volledig overzicht van de genoemde manieren om motivatie te bevorderen/activeren.

Tabel 6.53: Manieren om leerlingen te motiveren en activeren tijdens de schrijfles (interviews leerkrachten)

Manieren	bo (N=99)		sbo (N=36)	
	n	%	n	%
Aansluiten bij de belevingswereld	70	71%	23	64%
Schrijven voor echt publiek (authentieke taken)	21	21%	1	3%
Schrijven met echt doel (authentieke taken)	23	23%	2	6%
Nadruk op functionele context (bijv. koppeling wereldoriëntatie)	9	9%	3	8%
Laten publiceren/presenteren	19	19%	3	8%
Variatie in werkvormen	19	19%	4	11%
Creatief schrijven	19	19%	6	17%
Thematisch werken	7	7%	3	8%
Samen schrijven	6	6%	2	6%
Anders	42	42%	19	53%

NB: meerdere antwoorden mogelijk

Relatief veel leerkrachten geven ook een toelichting op de wijze waarop ze leerlingen motiveren in of voor de schrijfles (voor een deel behoort dit tot de categorie ‘anders’ in Tabel 6.45). Meerdere leerkrachten geven een voorbeeld van een schrijfo opdracht die aansluit bij de belevingswereld van de leerlingen, bijvoorbeeld aansluiting bij vakanties of feestdagen of “het opstellen van een brief aan de burgemeester om een uitkijktoren in het bos bij de school [te bouwen]”. Het overgrote deel van de toelichting bij deze vraag behoort, zoals een leerkracht opmerkt ‘tot de kernwaarden van vrijwel al het onderwijs’, namelijk zelf enthousiast zijn over de les of de opdracht (elf keer genoemd), leerlingen stimuleren; beloningen en ‘complimenten’ geven (zes keer genoemd), positieve feedback geven, voorkennis activeren of reflecteren op eerdere lessen (vier keer genoemd). Andere voorbeelden zijn: in de verwerking van de opdracht gebruik maken van nieuwe media (YouTube filmpjes, sociale media, Kahoot-quiz); zelf als rolmodel fungeren: voordoen (specifiek: “ik schrijf als leerkracht ook zelf mee met de opdracht”) en stimuleren door “in het begin van de opdracht veel te helpen en voor te doen”. Een leerkracht geeft aan dat het meteen feedback geven in dezelfde les motiverend werkt voor leerlingen. Enkele leerkrachten noemen de methode als zeer stimulerend/motiverend, een ander beschrijft juist dat zij buiten de methode werkt omdat deze “geheel niet bij de belevingswereld van kinderen past”.

6.5.3 Feedback, differentiatie, beoordeling en systematisch volgen

Net als bij andere onderwijsprocessen is het van groot belang te evalueren of leerlingen de gestelde doelen van het schrijfonderwijs behalen. Idealiter gaan leerkrachten tijdens, maar zeker aan het eind van iedere les bij de leerlingen na of het centraal gestelde leer-/schrijfdoel is bereikt en ook over de langere termijn is dienen de leer- en prestatiedoelen gemonitord en geëvalueerd te worden. Leerkrachten is gevraagd hoe zij schrijfproducten beoordelen en of (en hoe) zij de schrijfvaardigheid systematisch registreren. Een proces dat daaraan voorafgaat is het begeleiden van leerlingen in het schrijfproces, het hulp bieden bij de taakaanpak, en het formatief en summatief feedback geven op het schrijfproduct en het schrijfproces. De vraag is of de resultaten van het monitoren, evalueren en het geven feedback in het schrijfonderwijs leidt tot veranderingen in het onderwijsleerproces. Gevraagd is in welke mate en op welke manieren leerkrachten differentiëren tussen leerlingen.

Monitoren en evaluatie schrijfproces

Op de interviewvraag of leerkrachten tijdens het schrijven de taakaanpak (het schrijfproces) van de leerlingen ook monitoren, antwoordt een ruime meerderheid bevestigend. Zes (6%) van de bo-leerkrachten en 1 (3%) van de sbo-leerkrachten geeft aan dit niet te doen. Het monitoren van het schrijfproces gebeurt op een aantal manieren (Tabel 6.54). De meest genoemde manieren om het schrijfproces te monitoren is het rondlopen en observeren in de klas en het aanschuiven bij individuele leerlingen om (eventueel) hulp en ondersteuning te bieden.

Tabel 6.54: Monitoren van het schrijfproces (interviews leerkrachten)

Manieren van monitoren	bo (N=99)		sbo (N=36)	
	n	%	n	%
Ik loop rond in de klas en kijk hoe het gaat	82	83%	29	81%
Ik ga bij individuele leerlingen zitten	52	53%	20	56%
Ik ga bij groepjes leerlingen zitten	11	11%	8	22%
Ik observeer het schrijfgedrag van leerlingen	36	36%	18	50%
Anders	13	13%	3	8%

NB: meerdere antwoorden mogelijk

Monitoring komt, volgens de leerkrachten, vaak voor in een combinatie van aspecten. Een leerkracht helpt na de instructie bijvoorbeeld enkele individuele leerlingen op weg en gaat dan ook rondes lopen: “ideeën delen, in gesprek gaan”. Manieren waarop zij te werk gaan zijn de volgende: “Ik doe looprondes waarbij ik opmerkingen maak over opbouw, doel, spelling, interpunctie”; “[ik] laat op het digibord goede voorbeelden (stukje tekst) zien van leerlingen die bezig zijn met de schrijftaak”, “ik monitor of de juiste bronnen worden gebruikt en help ze”. Een leerkracht ziet op zijn digitale dashboard wat de leerlingen doen en monitort ze op die manier. Een voorbeeld van uitgestelde monitoring zijn weektaken: “leerlingen werken vanuit weektaken [en] geven zelf aan of hulp noodzakelijk is: uitgestelde aandacht [dus]”.

Evaluatie van de taakaanpak

Uit Tabel 6.55 blijkt dat de meeste geïnterviewde leerkrachten aangeven dat zij naast het **monitoren** van het schrijfproces tijdens de les het schrijfproces aan het eind van de les in enige mate **evalueren**. Bijna de helft (47%) van de bo- leerkrachten geeft aan dat vaak (in ongeveer twee derde van de lessen) of altijd te doen. In het sbo is dit zelfs drie kwart (75%). Slechts een zeer klein gedeelte van de gesproken leerkrachten geeft aan het schrijfproces aan het eind van de les nooit te evalueren.

Tabel 6.55: Mate waarin en wijze waarop leerkrachten schrijfproces evalueren (interviews leerkrachten)

	bo (N=99)		sbo (N=36)	
	n	%	n	%
Nooit	4	4%	1	3%
Weinig (ongeveer een derde van de lessen)	17	17%	3	8%
Soms (ongeveer de helft van de lessen)	28	28%	5	14%
Vaak (ongeveer twee derde van de lessen)	33	33%	18	50%
Altijd	14	14%	9	25%
Onbekend	3	3%	0	-
Manieren*	bo (N=99)		sbo (N=36)	
	n	%	n	%
Reflectieopdracht taalmethode	2	2%	3	8%
Klassikaal gesprek na elke les	46	47%	17	47%
Individueel gesprek na elke les	20	20%	6	17%
Zo nu en dan, met de hele klas	29	29%	8	22%
Zo nu en dan, individueel	10	10%	3	8%
Geschreven vragen en feedback	5	5%	1	3%
Anders	5	5%	1	3%

* NB: meerdere antwoorden mogelijk

Evalueren van het schrijfproces, zo blijkt uit de toelichtingen die leerkrachten geven, is zeker niet altijd standaard onderdeel van de les en is ook niet altijd even goed mogelijk. Dat heeft met de tijdsdruk te maken (te weinig tijd aan het eind van de les), maar ook met de behoefte (of ontbreken daarvan) die leerlingen hebben aan een klassikale of individuele evaluatie. Individuele leerkrachten geven aan dat ze aan het eind van de les graag zouden willen evalueren maar dat het “er soms bij inschiet”, en het minder “plaatsvindt dan bij andere taaldomeinen”, of dat het “niet op de dag zelf lukt”. Een leerkracht

geeft aan dat hij het schrijfproces niet evalueert met de leerlingen omdat in de methode is opgenomen dat leerlingen dat onderling moeten doen (“tips en tops geven”). Leerkrachten die nooit of weinig het schrijfproces evalueren geven een aantal keren aan dat dit komt doordat in de les (en in de methode) de focus ligt op het product van de schrijfles.

De evaluatie van het schrijfproces, zo blijkt uit Tabel 6.48, vindt voornamelijk plaats in een klassikale setting: in ongeveer de helft van de gevallen na elke les (bo en sbo) of zo nu en dan (29% bo, 22% sbo). Individuele bespreking/evaluatie van het proces komt minder voor, toch zegt nog ongeveer één op de vijf leerkrachten dat elke schrijfles te doen. Het aantal leerkrachten dat het schrijfproces evalueert via reflectieopdrachten of via geschreven vragen/feedback is zowel in het bo als sbo vrij laag.

Het evalueren van het schrijfproces op klassikale manier wordt vanuit meerdere invalshoeken gedaan: door een voorbeeldproces te bespreken, door per les het proces van leerlingen die er uit springen (positief en negatief) te behandelen in een klassikale setting, of door ervoor te zorgen dat gedurende het jaar het schrijfproces van elke leerling een keer aan de beurt komt.

Het evalueren van het schrijfproces kan onder andere plaatsvinden door een expliciete link te leggen met de leerdoelen van de schrijfopdracht. Gevraagd of en in welke mate leerkrachten dat doen antwoordt ruim drie kwart (bo: 78.8%, sbo:77.8%) dit in twee derde of meer van de lessen te doen (vaak). Slechts vier geïnterviewde leerkrachten (4%) in het bo geeft aan deze link nooit te leggen (in sbo zijn er geen leerkrachten die deze optie aankruisten).

Feedback op schrijfproducten

In zowel de leerkrachtvragenlijst als in interviews was er aandacht voor de wijze waarop en de frequentie van feedback op schrijfproducten (leerlingwerken) van leerlingen. Tabel 6.56 bevat de frequentie van soorten feedback, alsmede twee schalen die hieruit voortkomen.

Tabel 6.56: Frequentie van soorten feedback (leerkrachtvragenlijst)

		Nooit of soms	Meestal of vaak
Opbouw van de tekst (alinea's, logische volgorde)	bo (n=126)	20%	80%
	sbo (n=65)	28%	72%
Samenhang van de tekst (b.v. begrijpelijkheid door gebruik van signaalwoorden)	bo (n=125)	25%	75%
	sbo (n=66)	32%	68%
Stijl (toon, beleefdheid, stijlfiguren)	bo (n=126)	37%	63%
	sbo (n=65)	38%	62%
Spelling	bo (n=126)	10%	90%
	sbo (n=65)	34%	66%
Taalgebruik (zinsbouw, interpunctie en grammatica)	bo (n=126)	3%	97%
	sbo (n=66)	14%	86%
Handschrift	bo (n=126)	27%	73%
	sbo (n=66)	26%	74%
	Schooltype	M	SD
Schaalscore: Feedback op inhoud (communicatieve effectiviteit)	bo (n=126)	2.99	.61
	sbo (n=65)	2.93	.78
Schaalscore: Feedback op technische aspecten	bo (n=126)	1.60	.53
	sbo (n=65)	1.84	.65

NB: meerdere antwoorden mogelijk

De eerste drie items uit de tabel vormen samen een gemiddelde schaalscore 'feedback op inhoud' van de opdracht, terwijl de laatste drie, gehercodeerde items kunnen worden omgezet naar een schaalscore voor 'feedback op technische aspecten'. Dit betekent dat voor de schaalscore voor feedback op communicatieve inhoud geldt: 1= nooit (Min), 2= soms, 3= meestal, 4=vaak (Max), terwijl dit voor de schaalscore feedback op technische aspecten omgekeerd is (1=vaak, 2=meestal, 3=soms en 4=nooit). In het bo en sbo zijn de gemiddelde schaalscores vergelijkbaar voor de feedback op het communicatieve aspect (M bo=2.99; M sbo=2.93). Zij zeggen dus (bijna) meestal feedback op de schrijfproducten te geven. De schaalscore op feedback op technische aspecten is voor het sbo (M=1.84) hoger ten opzichte van het bo (M=1.60). De leerkrachten in beide onderwijstypen geven meestal tot vaak feedback op technische schrijfaspecten, zoals correctheid en netheid. De bo-leerkrachten doen dit vaker dan de sbo-leerkrachten (gezien de hergecodeerde schaalscore).

Van Gelderen (2010) benadrukt het belang van het geven van een schrijfopdracht met een goed omschreven schrijfdoel, omdat een schrijftaak te complex wordt wanneer leerlingen zich met alle aspecten van het schrijven bezig moeten houden. Het is wenselijk dat leerkrachten de feedback die zij in de schrijfles geven vooral richten op de geformuleerde schrijf-/leerdoelen. Op de stelling in de leerkrachtvragenlijst 'Ik beperk mijn feedback tot de leerdoelen die in de betreffende les centraal staan' antwoordt dan ook slechts 5.6% (bo) en 6.1% (sbo) van de leerkrachten dat dit niet (nooit) voor hen geldt. In het bo zijn de bevestigende antwoorden 34.9% (soms), 50.8% (meestal) en 8.7% (vaak), terwijl dat in het sbo als volgt is: 25.8% (soms), 42.4% (meestal) en 25.8% (vaak). In de meerniveau-analyses (hoofdstuk 9) wordt deze variabele als dichotome variabele meegenomen: 'de feedback is

nooit/soms gerelateerd aan de geformuleerde leerdoelen behorende bij de les', versus de feedback is meestal/vaak gerelateerd aan de leerdoelen (respectievelijk 94.4% en 93.9% voor bo en sbo).

In de interviews met leerlingen is op meerdere manieren stilgestaan bij het ontvangen van de feedback van de meester/juf of medeleerling (Tabellen 6.57a tot 6.57c). Op de vraag 'Geef je meester of juf jou ook aanwijzingen om de tekst die je geschreven hebt te verbeteren' geeft 4% (bo) en 10% (sbo) aan dat dit nooit voorkomt. Ongeveer de helft van de leerlingen in beide schooltypen geeft aan dat dit soms voorkomt. Het blijkt dat de directe mondelinge feedback van de leerkracht het meest voorkomt in/rondom de schrijflessen: 69% van de leerlingen in het bo en 81% van de leerlingen in het sbo geeft aan dat dat voorkomt in/rondom de schrijflessen. In het bo ontvangen leerlingen relatief vaker (47% versus 27%) losse opmerkingen bij de geschreven teksten.

Het gebruik van feedbackformulieren wordt door geen enkele leerling genoemd. Wanneer dergelijke formulieren aanwezig zijn in de methodes lijkt het er dus op dat deze niet worden toegepast door leerkrachten. In de categorie 'anders' wordt door leerlingen onder andere nog de volgende manieren benoemd: 'klassikale algemene tips voor schrijven', 'leerkracht stelt vragen', 'feedback a.d.h.v. succescriteria'. Ook benoemen leerlingen in dit stadium van het gesprek waarop de leerkracht dan specifiek (mondelinge) feedback op geeft. Het lijkt vooral te gaan om spelling, interpunctie, handschrift en een enkele keer om de schrijfhouding.⁹⁷

Tabel 6.57a: Ervaren feedback door leerkracht (interviews leerlingen)

Frequentie	bo (N=195)		sbo (N=71)	
	n	%	n	%
Nooit	7	4%	7	10%
Soms	96	49%	34	48%
Vaak	71	36%	26	37%
Altijd	21	11%	4	6%
Manieren*	bo (N=188)		sbo (N=64)	
	n	%	n	%
Losse opmerkingen bij de tekst	89	47%	17	27%
Tips en tops	29	15%	5	8%
Met een formulier uit de taalmethode	0	-	0	-
Met een eigen formulier van de klas	0	-	0	-
Mondeling, de meester of juf vertelt wat er beter kan	130	69%	52	81%
Klassikaal, we bespreken één tekst en verbeteren dan onze eigen tekst	14	7%	9	14%
Anders	13	7%	4	6%

* NB: meerdere antwoorden mogelijk

Leerlingen konden in de interviews ook aangeven of en hoe ze wel eens feedback geven op het schriftwerk van andere leerlingen. Tabel 6.57b laat zien dat in het bo vaker voorkomt dan in het sbo. Bijna twee vijfde van de sbo-leerlingen geeft aan dat ze nooit tips voor verbetering geven aan andere leerlingen. Van de sbo-leerlingen die wel anderen van feedback voorzien gaat dat in de meeste gevallen via mondelinge communicatie (91%) of via het geven van tips en tops (11%). In het bo zijn mondelinge feedback en 'tips en tops' ook de meest voorkomende manieren (69% en 23%). Opvallend

⁹⁷ Al is niet duidelijk wat de leerling (of de interviewer met deze notitie) precies bedoeld.

is dat er 11 leerlingen in het bo aangeven feedback te geven via het gebruik van een feedbackformulier; in tabel 6.50a zagen we dat zij van de leerkracht niet op een dergelijke manier feedback ontvangen.

Tabel 6.57b: Gegeven feedback op klasgenoten (interviews leerlingen)

Frequentie	bo (N-199)		sbo (N=72)	
	n	%	n	%
Nooit	33	17%	27	38%
Soms	109	55%	36	50%
Vaak	43	22%	8	11%
Altijd	14	7%	1	1%
Manieren*	bo (N-166)		sbo (N=45)	
	n	%	n	%
Opmerkingen schrijven bij de tekst	16	9%	3	7%
Tips en tops	38	23%	5	11%
Met een formulier uit de taalmethode	6	4%	0	-
Met een eigen formulier van de klas	5	3%	0	-
Mondeling, ik vertel wat er beter kan	115	69%	41	91%
Mondeling we bespreken de tekst in groepjes	6	4%	0	-
Anders	12	7%	3	7%

*NB: meerdere antwoorden mogelijk

Inhoudelijk lijkt de feedback die wordt gegeven door de leerkracht, volgens de leerlingen vooral te gaan over spelling, interpunctie, zinsbouw/zinsverbanden en handschrift (Tabel 6.57c). In mindere mate wordt de inhoud en de opbouw van de tekst genoemd; in het sbo wordt dit nauwelijks of helemaal niet genoemd. Opvallend is ook dat leerlingen zelf bijna niet zaken als doel, publiek, stijl en lay-out benoemen.

Tabel 6.57c: Aspecten van door de leerlingen ervaren feedback van leerkrachten (interviews leerlingen)

Frequentie	bo (N=188)		sbo (N=65)	
	n	%	n	%
De inhoud van de tekst	31	16%	4	6%
De opbouw van de tekst	29	15%	0	-
Het doel van de tekst	4	2%	0	-
Het publiek van de tekst	2	1%	0	-
Stijl / toon	6	3%	1	2%
Taalverzorging: zinsbouw/zinsverbanden	38	20%	5	8%
Taalverzorging: spelling	87	46%	17	26%
Netheid van het handschrift	37	20%	21	32%
Interpunctie	88	46%	27	42%
Lay-out	6	3%	5	8%
Anders	25	13%	20	31%

NB: meerdere antwoorden mogelijk; procentueel naar totaal aantal leerlingen die (soms, vaak of altijd) feedback ervaren.

In de categorie ‘anders’ worden door leerlingen vooral op een andere manier categorieën uit de tabel omschreven. Omdat dit niet altijd eenduidig is te herleiden naar deze aspecten zijn de antwoorden niet gehercodeerd. Het gaat met name om: het formuleren van zinnen en zinsopbouw; de verschillende schrijfwijze van woorden (in casu spelling) en manieren om de opdracht visueel er beter uit te laten zien (handschrift, schrijfhouding aan tafel, netjes werken, “niet krassen, maar doorstrepen”). Ook worden er dingen benoemd die raken aan de opbouw van de tekst: ‘meer spanning in de tekst’ en minder opsommend/herhalend schrijven (“en toen en toen.... liever niet”).

Reviseren en verbeteren

Na de - schriftelijke of mondelinge - feedback van de leerkracht zegt 86% van de bo-leerkrachten (n=93) en 54% van de sbo-leerkrachten (n=35) dat leerlingen na de eerste versie gelegenheid krijgen hun tekst te reviseren/verbeteren op inhoud. Opvallend is dat 15 (43%) van de sbo-leerkrachten aangeeft dat dit niet plaatsvindt (bo: 14%). Op de vraag ‘Hoe begeleidt u uw leerlingen bij het reviseren/verbeteren van teksten?’ geeft het overgrote deel van de leerkrachten meerdere antwoorden (Tabel 6.57). Tabel 6.58 laat zien dat begeleiding door leerkrachten in de revisiefase met name via een individueel gesprek over de aanpak gaat. In mindere mate wordt klassikaal de aanpak van het verbeteren/reviseren besproken. Revisieopdrachten (uit de methode) of modelleren – de leerkracht doet voor hoe je de geschreven tekst moet aanpassen – komen zowel in bo als sbo het minste voor.

Tabel 6.58: Manieren van begeleiding bij het reviseren en verbeteren van teksten (interviews leerkrachten)

Manieren	Bo (N=81)		Sbo (N=21)	
	n	%	n	%
Aan de hand van revisieopdrachten in de taalmethodode	9	11%	1	5%
Modelling (voordoen hoe je een tekst reviseert)	10	12%	3	14%
Klassikaal gesprek over de aanpak	28	35%	5	24%
Individueel gesprek over de aanpak	41	51%	10	48%
Individuele begeleiding tijdens het reviseren	33	41%	7	33%

NB: meerdere antwoorden mogelijk

Differentiatie

Uit de vorige peiling (Kuhlemeijer e.a., 2013) bleek dat er in het schrijfonderwijs veel onduidelijkheid was over de mate waarin en de manieren waarop leerkrachten differentiëren tussen leerlingen van verschillende (schrijf)niveaus. In de huidige peiling konden leerkrachten in de leerkrachtvragenlijst aangeven op welke algemene manieren zij bij de schrijflessen differentiëren naar niveau (Tabel 6.59). Differentiatie in het beoordelen (naar niveau) blijkt het meeste voor te komen in zowel bo als sbo. Daarop volgen differentiatie in begeleiding (instructie), en tijd. Het geven van verschillende opdrachten voor verschillende niveaus wordt door één vijfde van de bo-leerkrachten en een derde van de sbo-leerkrachten aangegeven.

Tabel 6.59: Manieren van differentiëren (leerkrachtvragenlijst)

Differentiatie: algemeen	bo (N=126)		sbo (N=66)	
	Niet	Wel	Niet	Wel
Alle leerlingen krijgen tegelijkertijd dezelfde schrijfoopdracht. Bij de beoordeling houd ik rekening met verschillen in niveau van leerlingen	13%	87%	21%	79%
Leerlingen krijgen verschillende schrijfoopdrachten afhankelijk van hun niveau	76%	24%	64%	36%
Ik geef zwakkere leerlingen meer tijd om hun schrijfoopdrachten te maken.	37%	63%	32%	68%
Ik begeleid zwakkere leerlingen meer bij het schrijven en laat de sterkere leerlingen zelfstandig werken.	32%	68%	21%	79%
Anders	85%	15%	86%	14%

NB: meerdere antwoorden mogelijk

In beide schooltypen geeft ongeveer één op de zeven leerkrachten aan dat zij (ook) op een andere manier differentiëren ('anders, namelijk'). Een groot aantal van deze manieren gaan echter over het uitzetten van verschillende opdrachten naar niveau, of naar het uitzetten van opdrachten in verschillende omvang:

- opdrachten gebruiken uit eerdere groepen (6-7);
- zelfde type opdrachten, maar het onderwerp is voor zwakke schrijvers anders;
- de hoeveelheid (deel)onderwerpen in opdracht aanpassen voor zwakkere schrijvers;
- verwerkingsopdracht schrijven naar aanleiding van gekozen boek, in het boek zitten al niveauverschillen;
- de eisen van de opdracht aanpassen (bijvoorbeeld meer of juist minder argumenten gebruiken in de opdracht);
- op niveau een stappenplan laten maken en kijken wat haalbaar is voor de leerling.

Andere toevoegingen in de categorie 'anders' gaan over het gebruik van de pc of tablets voor zwakkere schrijvers (of leerlingen met motorische beperkingen), of dyslectici (schrijven met programma 'textaid'). Met het digitale transcriptiemedium kunnen leerlingen dan allerlei instellingen gebruiken (lettergrootte, spellingscorrector, grammaticahulp). Verder worden genoemd: leerlingen laten samenwerken via coöperatieve werkvormen (differentiëren in samenstelling niveaus, drie keer genoemd), en leerlingen zelf laten plannen (via weektaken en montessorionderwijs).

Wanneer we kijken naar de observaties van de schrijflessen, die bij 70 klassen/leerkrachten in het bo hebben plaatsgevonden, blijkt differentiatie zeer beperkt in de schrijflessen aanwezig te zijn. Bij zes leerkrachten/lessen is een vorm van differentiatie gezien. In één van deze lessen werd gedifferentieerd in benodigde tijd voor de opdracht, in drie lessen in de gegeven schrijftaak en in drie lessen in instructie. In de 30 geobserveerde sbo-lessen kwam differentiatie vaker voor: in totaal is in dertien lessen een differentiatievorm geobserveerd. In zes van deze lessen differentieerde de leerkracht naar tijd, in negen lessen naar de gegeven schrijftaak en in zeven lessen in instructie.

Op het moment dat leerkrachten beoordelingen van schrijfproducten en de schrijfvaardigheidsontwikkeling van leerlingen als uitgangspunt nemen, kunnen zij beter, want duidelijker gericht op de behoeften van leerlingen, differentiëren. De leerkrachten is daarom in de interviews gevraagd of zij dit doen (en zo niet, waarom niet) en hoe zij vervolgens omgaan met verschillen tussen leerlingen, ofwel hoe zij zwakkere en/of sterkere leerlingen bedienen in het schrijfonderwijs.

Ruim drie vijfde van de bo-leerkrachten en bijna drie kwart van de sbo-leerkrachten geeft tijdens de interviews aan dat ze, om te differentiëren, gebruik maken van beoordelingen van schrijfproducten en andere manieren om de schrijfvaardigheidsontwikkeling van leerlingen te volgen. Gevraagd naar de reden(en) waarom leerkrachten hier geen gebruik van maken zeggen de meeste leerkrachten dat differentiatie op basis van informatie over leerlinguitkomsten geen prioriteit heeft. Aanvullend aan de genoemde opties in Tabel 6.60 wordt in de categorie 'anders' overigens geen extra informatiebron genoemd op basis waar de leerkracht differentieert.⁹⁸

⁹⁸ Wel wordt (opnieuw) het ontbreken van differentiatiemogelijkheden in de methode genoemd als ook dat alle leerlingen in de klas over hetzelfde niveau beschikken en differentiatie daarom niet nodig is.

Tabel 6.60: Redenen voor het niet gebruik maken van beoordelingen en inzicht in schrijfvaardigheid om te differentiëren (interviews leerkrachten)

	bo (N=37)		bbo (N=10)	
	n	%	n	%
Leerkracht heeft hier geen tijd voor	3	8%	4	40%
Leerkracht geeft hier geen prioriteit aan	12	32%	6	60%
Leerkracht heeft onvoldoende kennis en vaardigheden	5	14%	1	10%
Leerkracht heeft geen instrumenten voor het volgen van schrijfvaardigheid	6	16%	1	10%
Leerkracht heeft onvoldoende beoordelingsmomenten	2	5%	1	10%
Anders	6	16%	1	10%
Reden(en) onbekend	8	21%	0	-

NB: meerdere antwoorden mogelijk

Gevraagd naar de methodes om sterke en zwak schrijvende leerlingen in de klas beter te bedienen geven de leerkrachten uiteenlopende antwoorden. Uit Tabel 6.61 blijkt dat sterke schrijvers vooral meer uitdagende opdrachten krijgen van de leerkrachten. Zwakker schrijvers krijgen vooral meer instructie (verlengd), intensievere begeleiding in de taakaanpak, en - met name in het sbo - eenvoudigere opdrachten. De helft van de sbo-leerkrachten – die zegt te differentiëren (n=26) – doet dat door de schrijfproducten van sterke en zwakke schrijvers op verschillende manieren te beoordelen. In het bo (N=62) geeft één op de vijf leerkrachten (21%) aan dit te doen.

Tabel 6.61: Differentiëren naar sterke en zwakke schrijvers (interview leerkrachten)

Differentiatiemethodes	bo (N=62)		sbo (N=26)	
	n	%	n	%
Sterke schrijvers krijgen uitdagender opdrachten	19	31%	8	31%
Sterke schrijvers krijgen meer schrijftaken	3	5%	4	15%
Sterke leerlingen schrijven andere tekstsoorten	1	2%	0	-
Sterke schrijvers gebruiken een	0	-	0	-
Zwakke schrijvers krijgen eenvoudiger opdrachten	11	18%	8	31%
Zwakke schrijvers krijgen verlengde instructie	15	24%	4	15%
Zwakke schrijvers krijgen meer tijd voor	9	15%	2	8%
Zwakke schrijvers krijgen intensievere begeleiding op	16	26%	7	27%
Zwakke schrijvers krijgen extra hulpmiddelen	4	6%	2	8%
Ik geef zwakke schrijvers vaker feedback	8	13%	2	8%
Ik geef zwakke schrijvers meer feedback op	6	9%	0	-
Ik geef zwakke schrijvers meer feedback op	5	8%	0	-
Ik geef zwakke schrijvers meer feedback op inhoud	7	11%	0	-
Zwakke schrijvers gebruiken een tekstverwerker / PC	4	6%	1	4%
Ik beoordeel schrijfproducten op verschillende manieren	13	21%	13	50%

NB: meerdere antwoorden mogelijk

Beoordelingsmethoden en registratie vaardigheden

In de leerkrachtvragenlijst is gevraagd op welke manieren de leerkracht schrijfopdrachten voorzien van een (cijfermatige) beoordeling. De verdeling van de genoemde antwoorden is opgenomen in Tabel 6.62⁹⁹.

Tabel 6.62: Wijze van beoordelen (leerkrachtvragenlijst)

		nooit – soms	meestal – altijd
Ik geef een oordeel over de kwaliteit van de tekst als geheel.	bo (n=126)	13.5%	86.5%
	sbo (n=66)	16.7%	83.3%
Ik geef aparte oordelen voor aspecten van teksten (bijv. inhoud, opbouw en taalgebruik).	bo (n=126)	59.5%	40.5%
	sbo (n=66)	61.5%	38.5%
Bij het beoordelen maak ik gebruik van rubrics of schalen waarin verschillende niveaus van kwaliteit worden omschreven.	bo (n=126)	84.8%	15.2%
	sbo (n=66)	94.0%	6.0%

Het zogenoemde holistisch beoordelen (oordeel over de gehele kwaliteit) wordt door bijna alle leerkrachten in het bo en sbo gedaan. Slechts vier van de 193 deelnemende leerkrachten geeft aan dit nooit te doen. Het beoordelen van specifieke aspecten van de tekst wordt vaak tot altijd gedaan door ongeveer twee vijfde van de leerkrachten (in beide groepen). Het gebruik van rubrics of schalen in de beoordeling is iets wat maar weinig leerkrachten aangeven te doen: 19 leerkrachten (15%) in het bo en vier leerkrachten in het sbo (6%).

In de interviews is nader ingegaan op het beoordelen van schrijfproducten en het volgen van de schrijfvaardigheid van leerlingen. Ruim drie kwart (77%) van de bo-leerkrachten (n=92) geeft aan regelmatig complete schrijfproducten van leerlingen te beoordelen via een vorm van summatieve evaluatie. In het sbo is dat 44% (n=34).¹⁰⁰ Dat betekent dat 23 (bo), respectievelijk 56 procent (sbo) van de leerkrachten aangeeft geheel niet aan summatieve toetsing te doen.

Van de leerkrachten die beoordelen geven de meesten aan dit met een woordschaal of met concrete cijfers te doen. Daarbij valt op dat de helft van de bo-leerkrachten dat met een woordschaal doet, tegenover en slechts een vijfde van de sbo-leerkrachten. Ongeveer één op de acht leerkrachten geeft aan voor verschillende onderdelen van de tekst een deelcijfer of oordeel te geven. In de categorie anders vallen onder meer: een mondeling oordeel; een mondeling oordeel i.c.m. geschreven feedback; punten per onderdeel in een portfolio; terugkoppeling van de opdracht naar de lesdoelen; en 'beperking beoordeling (met woorden) tot inzet en schrijfproces'. Vijf leerkrachten benoemen dat ze schrijfvaardigheid alleen bij een groot werkstuk (meestal één keer per jaar) beoordelen. Dat werkstuk wordt dan gemaakt voor de zaakvakken (wereldoriëntatie, aardrijkskunde, geschiedenis). Het gebruik van schrijftoetsen en rubrics worden nauwelijks genoemd (in absolute aantallen).

⁹⁹ In de meerniveau-analyses (hoofdstuk 9) worden de twee onderste variabelen (oordeel op aparte aspecten, en gebruik rubrics) als dichotome variabelen gebruikt (0= nooit-soms/ 1= meestal/vaak). Ten aanzien van het eerste item is sprake van te weinig variatie: bijna alle leerkrachten geven een oordeel over het geheel.

¹⁰⁰ Voor zeven bo- en twee sbo-leerkrachten is geen informatie beschikbaar t.a.v. deze interviewvraag.

Tabel 6.63: Summatieve beoordelingsindicatoren/methoden (interviews leerkrachten)

	bo (N=78)		sbo (N=15)	
	n	%	n	%
Een cijfer voor de tekst als geheel	20	26%	2	13%
Aparte cijfers/oordelen voor verschillende aspecten van een tekst (bijv. inhoud en taalgebruik)	9	12%	2	13%
Een cijfer via een methode gebonden schrijftoets	2	3%	3	20%
Een beoordeling via een woordschaal (bijv. voldoende/goed)	39	50%	3	20%
Een rubriek (rubric) met omschrijvingen van prestaties	7	9%	1	7%
Anders	21	27%	4	27%

NB: meerdere antwoorden mogelijk

De mate waarin leerkrachten schrijfproducten beoordelen varieert sterk, van 1 keer per jaar tot elke week (circa 40 keer per jaar, zie Tabel 6.64). In het geval van 1 à 2 keer per jaar gaat het meestal, zoals hierboven al genoemd, om het beoordelen van het schrijfwerk voor werkstukken die voor de zaakvakken gemaakt worden. In het bo beoordeelt 58% van de leerkrachten schrijfproducten minimaal maandelijks of vaker (elke week). In het sbo is dat 50%. Zeven leerkrachten geven een andere frequentie van beoordelen ('anders'), bijvoorbeeld: 'alle producten worden bekeken, maar niet beoordeeld'; 'elke twee weken'; 'alle producten worden beoordeeld, maar frequentie onduidelijk'; en '4 keer per jaar, [maar niet in alle kwartalen]'.

Tabel 6.64: Beoordelingsfrequentie schrijfproducten (interviews leerkrachten)

	N	1 à 2 keer per jaar	4-6 keer per jaar *	Elke maand	Elke week	Anders
Bo	77	4%	32%	48%	10%	5%
Sbo	16	19%	25%	19%	31%	19%

* eens per kwartaal / tweemaandelijks

Op de vraag of de leerkracht de vorderingen van zijn/haar leerlingen op een systematische manier registreert, antwoord 64% van de bo-leerkrachten (64 personen) en 42% van de sbo-leerkrachten (15 personen) bevestigend. In het bo wordt 20 keer de groepsmap genoemd (31% van 64); 23 keer via het leerlingvolgsysteem (36% van 64) en 22 keer een andere manier ("anders namelijk") Voor het sbo is dit respectievelijk 7 keer (47%) de groepsmap, 5 keer (33%) het leerlingvolgsysteem en 4 keer (27%) een andere manier. De manieren die onder andere bij 'anders' worden genoemd zijn: een eigen of andere registratiesysteem dan het LLVS, het leerlingportfolio, scorelijsten die in de methode zijn opgenomen, de applicatie Snappet, en via de beoordeling in rapporten (bijvoorbeeld: onvoldoende, voldoende, goed).

Formatieve evaluatie

Het (systematisch) volgen en beoordelen van leerlingen in het onderwijs om de ontwikkeling van een leerling in een bepaald domein in kaart te brengen, kan voor leerkrachten een uitdaging zijn. Dat geldt

vermoedelijk ook voor het schrijfonderwijs. Op de vraag in ‘Heeft u het idee dat u een goed beeld heeft van de schrijfvaardigheidsontwikkeling van individuele leerlingen, zodat u in uw onderwijs hierop kunt inspelen (formatieve evaluatie)?’ antwoordt 29 procent van de geïnterviewde bo-leerkrachten en 39 procent van de sbo-leerkrachten negatief. Geen prioriteit en geen tijd zijn de meest genoemde redenen (Tabel 6.65a).

De overige leerkrachten¹⁰¹ (bo 68%; sbo 61%) hebben wel het idee dat zij een goed beeld hebben van de ontwikkeling van leerlingen als het gaat om schrijven, en wel om de redenen die zijn genoemd in Tabel 6.65b. Het observeren van leerlingen tijdens de schrijfles als ook het regelmatig beoordelen van schrijfproduct en -proces zijn zaken die volgens leerkrachten een goed inzicht in de schrijfvaardigheidsontwikkeling geven. In het bo wordt het gebruik van een schrijfportfolio en informatieoverdracht tussen de groepen genoemd, terwijl daar in het sbo geen sprake van lijkt te zijn.

Tabel 6.65a: Genoemde redenen onzeker beeld m.b.t. schrijfvaardigheidsontwikkeling

	bo (N=29)	sbo (N=14)
Geen tijd	6	6
Geen prioriteit	14	4
Onvoldoende kennis en vaardigheden	6	0
Geen instrumenten	6	1
Anders	6	5

NB: meerdere antwoorden mogelijk

Tabel 6.65b: Genoemde redenen voor een goed beeld m.b.t. schrijfvaardigheidsontwikkeling

	bo (N=68)		sbo (N=22)	
	n	%	n	%
Door regelmatige beoordeling schrijfproducten	40	59%	13	60%
Door regelmatige evaluaties van het schrijfproces (klassikaal)	21	31%	6	27%
Door gesprekken met individuele leerlingen	19	28%	6	27%
Aan de hand van een schrijfportfolio (en gesprekken)	10	15%	0	-
Door overdracht van informatie over vorderingen binnen de school	2	3%	0	-
Door leerlingen tijdens het schrijven te observeren	41	60%	13	60%
Anders	6	9%	2	9%

NB: meerdere antwoorden mogelijk

¹⁰¹ Er missen gegevens over dit onderwerp voor twee bo-leerkrachten

6.6 Resultaten vragenlijst leerlingen

De vragenlijst voor de leerlingen bevatte vragen naar de achtergrond, schrijfervaring en attitude voor schrijven van leerlingen. De resultaten worden telkens apart gepresenteerd voor bo en sbo. Deze paragraaf eindigt met informatie uit de leerlinginterviews over hoe zij te werk gaan tijdens verschillende fasen van het schrijfproces. Onderscheiden worden de aanpak, het schrijven zelf en de revisie.

6.6.1 Achtergrondinformatie

Verreweg de meeste deelnemende leerlingen zijn in Nederland geboren (zie Tabel 6.66). Van de bo-leerlingen die de vraag hebben ingevuld is 95% in Nederland geboren en 5% niet, terwijl het percentage in het sbo dat in Nederland geboren is iets lager ligt (93%). Gemiddeld waren de leerlingen die niet in Nederland zijn geboren 5 jaar oud toen zij naar Nederland kwamen (zowel in bo als sbo). Leerlingen waren in bo en sbo bij aankomst in Nederland tussen 0 en 11 jaar oud.

Tabel 6.66: Land van herkomst deelnemende leerlingen (leerlingvragenlijst)

Type school	Schoolgewicht	Aantal	Geboren in Nederland	Niet geboren in Nederland	
			%	%	Gem. leeftijd naar NL (S.D.)
bo	S1	149	96	4	2.8 (2.7)
	S2	1706	97	3	4.4 (3.6)
	S3	360	91	9	6.2 (3.4)
	S4	203	90	10	5.3 (3.2)
bo	Totaal	2418	95	5	4.9 (3.5)
sbo	Totaal	833	93	7	5.4 (3.2)

Om zicht te krijgen op de taalsituatie van leerlingen is een drietal vragen gesteld. Allereerst is leerlingen gevraagd in welke taal ze hebben leren praten (Tabel 6.67). Leerlingen konden hier kiezen tussen drie opties: 1) het Nederlands, 2) een andere taal en 3) in twee talen (waaronder Nederlands). Ongeveer driekwart van de bo- en sbo-leerlingen heeft in het Nederlands leren praten, hoewel we zien dat met name in S4-scholen de meeste leerlingen in twee talen hebben leren praten (61%). De S4-scholen bevat tevens de grootste groep leerlingen die in een andere taal dan het Nederlands heeft leren praten (10%). Ten tweede is gevraagd naar de taal die zij met hun ouders spreken. De antwoordopties zijn teruggebracht naar drie antwoordcategorieën: 1) Nederlands (eventueel in combinatie met dialect/Fries/andere taal), 2) dialect of Fries en 3) een andere taal¹⁰². De overgrote meerderheid van de leerlingen (ca 90% in zowel bo als sbo) spreekt in ieder geval Nederlands met de ouders. Bij de S4-groep ligt dit percentage lager (82%) en ligt het percentage leerlingen dat in een andere taal met ouders spreekt hoger (15%). In het sbo wordt het hoogste percentage leerlingen gevonden dat enkel in dialect of Fries met de ouders spreekt. Tenslotte is gevraagd naar de taal die ouders onderling

¹⁰² Er is voor dezelfde indeling als bij de meerniveau-analyses gekozen. Voor de driedeling was met name het onderscheid tussen het spreken van Nederlands (al dan niet in combinatie met een andere taal/dialect) van belang.

spreken, waarvoor dezelfde indeling wordt gehanteerd. Zoals Tabel 6.68 laat zien spreken ouders onderling minder vaak Nederlands en vaker in een andere taal.

Tabel 6.67: Taalsituatie van de leerlingen (leerlingvragenlijst)

Taal waarin leerling heeft leren praten								
Type school	schoolgewicht	Aantal	NL		Andere taal		NL en andere taal	
		N	N	%	n	%	n	%
Bo	S1	149	127	85%	1	1%	21	14%
	S2	1710	1489	87%	18	1%	203	12%
	S3	359	243	68%	14	4%	102	28%
	S4	206	60	29%	21	10%	125	61%
Bo	Totaal	2424	1919	79%	54	2%	451	19%
Sbo		894	664	74%	38	4%	192	22%
Taal waarin leerling met ouders praat								
Type school	schoolgewicht	Aantal	Nederlands (én evt. dialect/ Fries/andere taal)		Alleen dialect of Fries		Alleen andere taal	
		N	N	%	n	%	n	%
bo	S1	149	146	98%	1	1%	2	1%
	S2	1713	1586	93%	99	6%	28	2%
	S3	360	325	90%	8	2%	27	8%
	S4	206	168	82%	7	3%	31	15%
bo	Totaal	2428	2225	92%	115	5%	88	4%
sbo		894	789	88%	62	7%	43	5%

Tabel 6.68: Taal waarin ouders onderling spreken (leerlingvragenlijst)

Type school	Schoolgewicht	Aantal	Nederlands (en evt ook dialect, Fries)		Alleen dialect of Fries		Nederlands en een andere taal		Niet van toepassing (bij 1 ouder wonend)	
			n	%	n	%	n	%	n	%
bo	S1	148	125	85%	4	3%	17	12%	2	1%
	S2	1711	1363	80%	145	9%	169	10%	34	2%
	S3	353	233	66%	12	3%	100	28%	8	2%
	S4	205	41	20%	7	3%	145	71%	12	6%
bo	Totaal	2417	1762	73%	168	7%	431	18%	56	2%
sbo		892	589	66%	77	9%	189	21%	37	4%

6.6.2 Ervaring met schrijven

Voor de ontwikkeling van de schrijfvaardigheid is het belangrijk dat leerlingen veel schrijven. Om zicht te krijgen op de mate waarin leerlingen op school en in hun vrije tijd schrijven, zijn vragen gesteld over het schrijven van teksten op de mobiele telefoon, op pc/tablet of papier en is geïnventariseerd hoe vaak leerlingen inschatten dat zij een tiental teksten schrijven, hetzij in hun vrije tijd, hetzij op school.

Schrijven in de vrije tijd

Veel leerlingen aan het eind van het basisonderwijs beschikken over een mobiele telefoon, waarop berichtjes en teksten kunnen worden geschreven. Gevraagd is naar het aantal minuten dat zij de voorgaande week op hun mobiel hebben geschreven (Tabel 6.69). Opmerkelijk is dat de mate waarin leerlingen van beide schooltypen teksten op hun mobiel schrijven uiteenloopt. Bij bo-leerlingen heeft 35% van de leerlingen maximaal 10 minuten op de mobiel geschreven, terwijl bijna een kwart van hen dit meer dan 60 minuten heeft gedaan. Ook binnen het sbo lopen deze percentages sterk uiteen. Leerlingen in het sbo besteden de minste tijd aan het schrijven op de mobiel: in het sbo geeft 47% van de leerlingen aan maximaal 10 minuten te hebben geschreven op de mobiel.

Tabel 6.69: Aantal minuten dat de afgelopen week op de mobiel teksten zijn geschreven (leerlingvragenlijst)

		Aantal	0 minuten	10 minuten	30 minuten	60 minuten	> 60 minuten	Geen mobiel
		N	%	%	%	%	%	%
Bo	S1	148	7%	25%	29%	14%	23%	3%
	S2	1713	5%	30%	27%	14%	21%	3%
	S3	359	7%	27%	23%	13%	26%	5%
	S4	206	10%	25%	13%	11%	36%	5%
	Totaal	2426	6%	29%	25%	14%	23%	4%
sbo		895	14%	33%	13%	8%	26%	6%

Daarnaast is de leerlingen gevraagd hoe vaak in de week zij inschatten een tekst op hun mobiel te schrijven. Afgezien van een klein percentage leerlingen dat aangeeft niet over een mobiel te beschikken (maximaal 6%), blijkt in het bo de helft van de leerlingen een paar keer per dag tekstberichten te schrijven op de mobiel, terwijl dit in het sbo 39% van de leerlingen betreft. De minderheid van de leerlingen in zowel bo als sbo schrijft niet tot een paar keer per week (maximaal 21% in het sbo).

In 2009 zijn aangaande het gebruik van sociale media twee vragen gesteld: de mate waarin leerlingen in een week 1) chatten¹⁰³ en 2) sms'en. Van een vergelijking met de huidige peiling wordt afgezien, omdat sociale media-kanalen, zoals Whatsapp, die nu populair zijn toen nog niet beschikbaar waren, de toen wel beschikbare kanalen vandaag de dag minder vaak gebruikt worden en naar alle waarschijnlijkheid veel minder leerlingen in 2009 een mobiele telefoon hadden.

Tabel 6.70: Schrijven op de mobiele telefoon (per week) (leerlingvragenlijst)

		N	nooit	Paar keer per week	Bijna elke dag	Paar keer per dag	Heb geen telefoon
2019							
bo	S1	149	0%	9%	24%	64%	3%
	S2	1718	1%	13%	29%	54%	3%
	S3	360	1%	14%	31%	49%	5%
	S4	205	4%	10%	42%	40%	5%
bo	Totaal	2432	1%	13%	30%	52%	4%
sbo		894	5%	16%	34%	39%	6%

¹⁰³ Wat precies onder chatten wordt verstaan is niet nader gespecificeerd.

Daarnaast is leerlingen gevraagd naar een inschatting van het aantal minuten dat zij de voorgaande week in hun vrije tijd op papier, een computer of een tablet hebben geschreven. In Tabel 6.71 staan de resultaten weergegeven. De groep bo- en sbo-leerlingen dat meer dan een uur heeft geschreven in de vrije tijd is lager dan bij de mobiel: gemiddeld 17 à 18% van de leerlingen zegt dit te doen; in S4 geldt dit voor meer dan een kwart van de leerlingen. Voor zowel de bo- als sbo-leerlingen geldt dat respectievelijk bijna de helft (49%) en meer dan de helft (62%) dit maximaal 10 minuten doet. Overigens loopt de mate waarin leerlingen schrijven uiteen: 24% van de bo-leerlingen zegt de afgelopen week niet te hebben geschreven in de vrije tijd, terwijl 18% van hen zegt dit meer dan een uur te hebben gedaan.

Tabel 6.71: Schrijven in de vrije tijd (op papier, computer of tablet) (leerlingvragenlijst)

		aantal	0 min	10 min	30 min	60 min	> 60 min
		N	%	%	%	%	%
Bo	S1	149	28%	32%	15%	9%	16%
	S2	1714	23%	25%	22%	13%	17%
	S3	358	23%	24%	22%	12%	19%
	S4	206	26%	21%	19%	6%	27%
Bo	Totaal	2426	24%	25%	21%	12%	18%
Sbo		894	36%	26%	14%	8%	17%

Naast een inschatting van de tijd die leerlingen de afgelopen week hebben besteed aan het schrijven, is gevraagd naar de mate van schrijven in hun vrije tijd over een jaar. Ruim veertig procent (bo 42% en sbo 46%) zegt bijna iedere dag te schrijven op papier, computer of laptop. Met name in het sbo is de diversiteit in schrijffrequentie groot: bijna een kwart van de sbo-leerlingen zegt nooit te schrijven en 46% van hen zegt dit (bijna) dagelijks te doen. In 2009 is dezelfde vraag gesteld Tabel 6.72 laat zien dat het percentage leerlingen dat in 2009 zei bijna dagelijks te schrijven lager is dan in 2019. De percentages leerlingen die nooit zeggen te schrijven leken in 2009 ook hoger, in bo en sbo.

Tabel 6.72: Mate waarin leerlingen in hun vrije tijd schrijven (leerlingvragenlijst)

2019		n	nooit	1 of 2* per jaar	1* per maand	1* per week	Bijna iedere dag
Bo	S1	149	11%	12%	12%	29%	36%
	S2	1715	8%	9%	15%	26%	42%
	S3	360	11%	8%	9%	29%	44%
	S4	206	14%	4%	11%	24%	47%
Bo	Totaal	2429	9%	9%	14%	26%	42%
Sbo		895	23%	7%	6%	18%	46%
2009		n	nooit	1 of 2* per jaar	1* per maand	1* per week	Bijna iedere dag
Bo		797	18%	16%	22%	26%	19%
Sbo		247	31%	5%	5%	28%	31%

Leerlingen is van tien soorten teksten gevraagd hoe vaak ze die op school of in hun vrije tijd schrijven. Leerlingen konden antwoorden aan de hand van de antwoordcategorieën nooit, soms en vaak. De resultaten geven een door de leerlingen gegeven inschatting van de frequentie weer (zie Tabel 6.73). Een stukje voor de schoolkrant, in een dagboek, voor internet en een liedtekst worden in 2019 door het merendeel (meer dan 60%) van de bo- en sbo-leerlingen nooit geschreven. De meest geschreven teksten zijn een verhaal, een briefje/kaartje aan familie of vrienden en (met name in het bo) een werkstuk. Leerlingen in het bo zeggen daarbij over het algemeen vaker dat zij deze teksten vaak schrijven dan sbo-leerlingen.

In 2009 is dezelfde vraag aan de leerlingen voorgelegd, Wel was er een lichte wijziging in antwoordcategorieën, zoals het geven van een voorbeeld als 'hyves' bij teksten op internet. Wanneer we de resultaten vergelijken, zien we geen grote verschillen, op enkele uitzonderingen na. Zo lijkt in 2019 minder vaak een stukje voor de schoolkrant te worden geschreven. Deze stukjes worden in 2019 in het bo door 85% en in het sbo door 87% van de leerlingen *nooit* geschreven. Ook noteren leerlingen in 2019 minder vaak een stukje voor internet te schrijven (in 2019 resp. 7% en 10% *vaak* in bo en sbo, terwijl in 2009 resp. 41% en 40% *vaak*). Dit zou mogelijk het gevolg kunnen zijn van het niet noemen van voorbeelden als Facebook of Instagram, waardoor leerlingen niet aan bepaalde sociale media-kanalen denken, terwijl deze voorbeelden in 2009 wel werden gegeven. Ook het percentage leerlingen in het bo en sbo dat zegt nooit een boekverslag te schrijven lijkt hoger in 2019 dan in 2009.

Tabel 6.73: De mate waarin leerlingen verschillende soorten teksten schrijven (op school of in vrije tijd) – in 2019 en 2009 (leerlingvragenlijst)

Tekstsoort		2019, percentages			2009, percentages		
		Nooit	Soms	Vaak	Nooit	Soms	vaak
Een verhaal	bo	8	71	21	5	79	16
	sbo	16	71	13	16	75	9
Een e-mail	bo	48	46	6	9	49	41
	sbo	55	36	9	30	41	29
Een briefje of kaartje aan familie of vrienden	bo	19	65	15	16	68	16
	sbo	23	65	13	20	68	13
Een stukje voor de schoolkrant	bo	85	14	2	42	49	9
	sbo	87	12	1	53	41	6
Een stukje voor internet (b.v. blog of website)	bo	70	23	7	23	36	41
	sbo	61	28	10	24	36	40
Een stukje in mijn dagboek	bo	74	19	7	63	27	10
	sbo	74	17	9	66	23	11
Een werkstuk (b.v. voor geschiedenis of techniek)	bo	9	70	21	5	76	19
	sbo	27	61	12	25	61	14
Een verslag over een boek dat je gelezen hebt	bo	42	52	6	26	66	7
	sbo	54	40	6	45	49	6
Een gedicht	bo	38	56	6	30	60	10
	sbo	37	56	7	33	56	11
Een tekst voor een liedje	bo	65	28	7	-	-	-
	sbo	62	28	10	-	-	-

De mate waarin leerlingen in 2019 de tien verschillende teksttypen zeggen te schrijven is gesommeerd om een indicatie te geven van de schrijfervaring van leerlingen in en buiten school. Deze schaal loopt van 10 tot 30 en kan als volgt worden geïnterpreteerd: hoe hoger de score, hoe meer opgedane schrijfervaring. In het bo is deze somscore voor 2390 leerlingen beschikbaar. Binnen een range lopend van minimumscore 10 tot en met maximumscore 27 is de gemiddelde score voor bo-leerlingen 16.38 (SD=2.71). Van 854 sbo-leerlingen was deze somscore beschikbaar. Voor de sbo-leerlingen loopt de range van minimumscore 10 tot en met maximumscore 26 en is de gemiddelde score 15.96 (SD=2.89).

6.6.3 Schrijfplezier

Eén van de doelen van het schrijfonderwijs is dat leerlingen plezier krijgen in het lezen en schrijven van voor hen bestemde verhalen, gedichten en informatieve teksten (kerndoel 9). Leerlingen die plezier hebben in het schrijven van teksten zullen bovendien vaker schrijven, hetgeen hun vaardigheid positief kan beïnvloeden. De in de huidige peiling gebruikte vragen binnen de schaal voor schrijfplezier zijn afgeleid uit eerder onderzoek naar schrijven binnen het vmbo (De Miliano, 2013). De schaal bleek opgebouwd uit 7 items (zie paragraaf 5.8). In Tabel 6.74 is een overzicht gegeven van de items uit de vragenlijst voor leerlingen die bij de schaal Schrijfplezier horen. Daarna worden twee voorbeelditems gegeven.

Tabel 6.74: Schaal schrijfplezier

Schaalscore	Items
Schrijfplezier	19_1, 19_2, 19_3, 19_4, 19_5, 19_6, 19_7 ¹⁰⁴

Voorbeelditems attitude:

In mijn vrije tijd schrijf ik voor mijn plezier
<input type="checkbox"/> helemaal mee oneens
<input type="checkbox"/> mee oneens
<input type="checkbox"/> niet mee eens, niet mee oneens
<input type="checkbox"/> mee eens
<input type="checkbox"/> helemaal mee eens

Zonder schrijven zou school leuker zijn (gehercodeerd)
<input type="checkbox"/> helemaal mee oneens
<input type="checkbox"/> mee oneens
<input type="checkbox"/> niet mee eens, niet mee oneens
<input type="checkbox"/> mee eens
<input type="checkbox"/> helemaal mee eens

In Tabel 6.75 is de verdeling van scores weergegeven voor de sbo-leerlingen, voor de bo-en sbo-leerlingen. De schaal schrijfplezier loopt van 0-28. Gemiddeld hebben de bo-leerlingen een relatief laag niveau van schrijfplezier ($M=12.61$) en datzelfde geldt voor de sbo-leerlingen ($M=12.00$).

Omdat meisjes in de vorige peiling meer plezier in schrijven hadden, is er ook nu gekeken naar mogelijke verschillen in schrijfplezier tussen jongens en meisjes. In het bo hebben meisjes gemiddeld significant meer plezier in schrijven dan jongens (resp. $M=14.82$ en $M=10.39$, $t(2153)=18.2$, $p=0.00$). Voor meisjes in het sbo geldt hetzelfde: zij hebben significant meer plezier in schrijven dan jongens (resp. $M=14.79$ en $M=10.62$, $t(831)=9.61$, $p=0.00$). Hoewel in 2009 schrijfplezier anders is bevraagd, kwam een vergelijkbaar beeld naar voren.

¹⁰⁴ Items 19_5 en 19_6 zijn gehercodeerd.

Tabel 6.75: Gemiddelde gewogen score op schaal Schrijfplezier (leerlingvragenlijst)

Type school	Aantal	gemiddelde	SD	Min	P25	P50	P75	Max
bo	2377	12.61	5.98	0	8	13	17	28
sbo	863	12.00	6.26	0	8	12	16	28

Ook in de vorige peiling (2009) is leerlingen gevraagd naar schrijfplezier. Hiervoor is in 2009 geen schaal gebruikt, maar is leerlingen gevraagd wat zij van schrijven vonden. Hierbij waren twee antwoordopties: leuk en niet zo leuk. Van de bo-leerlingen in groep 8 bleek in 2009 54% schrijven leuk te vinden en 46% niet zo leuk, terwijl in het sbo 60% schrijven leuk vond en 40% niet zo leuk. Hoewel in 2009 de indeling in groepen scholen naar formatiegewicht verschilde van die in 2019, bleken ook in 2009 de bo-leerlingen op scholen met veel leerlingen met een formatiegewicht (S3 in 2009 en S4 in 2019) schrijven leuker te vinden dan op scholen met weinig leerlingen met een formatiegewicht.

6.6.4 Self efficacy

Een hogere beleving van de eigen competentie in een bepaald vakgebied van een leerling (self efficacy), gaat over het algemeen gepaard met betere prestaties (Pajares & Valiante, 1997). Leerlingen die er weinig vertrouwen in hebben dat ze adequate teksten kunnen schrijven, zullen minder geneigd zijn veel te schrijven en zich hiervoor in te spannen dan leerlingen die dit vertrouwen wel hebben, en vice versa (leerlingen die denken dat zij weinig vaardig zijn hebben vaak ook weinig vertrouwen in hun eigen kunnen). Dit kan negatieve gevolgen hebben voor hun vaardigheid.

In Tabel 6.76 is een overzicht gegeven van de vijf items uit de leerlingvragenlijst die bij de schaal self efficacy horen. Daaronder worden twee voorbeelditems getoond.

Tabel 6.76: Schaal self efficacy (beleving eigen competentie in schrijven)

Schaalscore	Items
Self efficacy	19_8, 19_9, 19_10 (rec), 19_11, 19_12 (rec)

Voorbeelditems Self efficacy

Ik ben goed in schrijven op school
<input type="checkbox"/> helemaal mee oneens
<input type="checkbox"/> mee oneens
<input type="checkbox"/> niet mee eens, niet mee oneens
<input type="checkbox"/> mee eens
<input type="checkbox"/> helemaal mee eens

Voor een schrijfpodracht op school ben ik altijd zenuwachtig
<input type="checkbox"/> helemaal mee oneens
<input type="checkbox"/> mee oneens
<input type="checkbox"/> niet mee eens, niet mee oneens
<input type="checkbox"/> mee eens
<input type="checkbox"/> helemaal mee eens

In Tabel 6.77 is de verdeling van scores weergegeven voor de sbo-leerlingen, voor de bo- en sbo-leerlingen. Op de schaal self efficacy, die van 0 tot 20 loopt scoren bo-leerlingen in Nederland gemiddeld 13.12 en de sbo-leerlingen 12.14. Dit betekent dat de zelfwaardering voor schrijven in beide groepen vrij positief is en dat het geloof in eigen 'schrijf-kunnen' een punt hoger is bij de bo-leerlingen.

Gekeken is of jongens en meisjes verschillen in hun geloof in eigen competentie op het gebied van schrijven. In het bo hebben meisjes gemiddeld een significant hogere beleving van de eigen competentie in schrijven dan jongens (resp. $M=13.40$ en $M=12.89$, $t(2153)=3.67$, $p=0.00$). In het sbo wordt dit onderscheid tussen jongens en meisjes niet gevonden, hun beleving van de eigen competentie voor schrijven is vergelijkbaar ($M_j=12.18$ en $M_m=12.21$, $t(831)=0.12$, $p=0.91$). In 2009 bleken jongens en meisjes in zowel groep 8 als in het sbo niet van elkaar te verschillen met betrekking tot hun beleving van de eigen competentie, ook al was deze beleving toen anders bevestigd.

Tabel 6.77: Gemiddelde score op schaal self efficacy (leerlingvragenlijst)

Type school	Schoolgewicht	Aantal	M	SD	Min	P25	P50	P75	Max
bo	Totaal	2396	13.12	3.22	1	11	13	15	20
sbo		864	12.14	3.85	0	9	12	15	20

Ook in de vorige peiling (2009) is leerlingen gevraagd naar de beleving van de eigen competentie in schrijven. Hiervoor is in 2009 geen schaal gebruikt, maar is leerlingen gevraagd hoe goed zij waren in schrijven. Hierbij waren er twee antwoordopties: goed en niet zo goed. Van de bo-leerlingen in groep 8 vond in 2009 55% zichzelf goed in schrijven en 45% niet zo goed, terwijl in het sbo 59% zichzelf goed vond in schrijven en 41% niet zo goed. Kijkend naar de gevonden verschillen tussen bo-scholen met verschillende percentages leerlingen met een formatiegewicht, die in 2009 weliswaar anders zijn ingedeeld, bleken in 2009 de leerlingen op scholen met het hoogste percentage leerlingen met een leerlinggewicht een hogere self efficacy voor schrijven te hebben dan de leerlingen op scholen met minder of (bijna) geen gewichtenleerlingen. Deze bevinding wordt in 2019 niet gevonden, zoals hierboven beschreven. In 2019 wordt juist gevonden dat de leerlingen in S1 een hogere self efficacy hebben dan die op de andere scholen.

6.6.5 Resultaten interviews: het schrijfproces beschreven

Met behulp van de interviews zijn leerlingen bevraagd over hoe zij te werk gaan tijdens het schrijven. Er is een onderscheid gemaakt tussen de verschillende fasen in het schrijfproces. Zo zijn leerlingen bevraagd over de manier waarop zij een opdracht aanpakken, waar ze tegenaan lopen tijdens het schrijven en of en hoe zij teksten verbeteren.

Aanpak: hoe beginnen leerlingen aan een concrete opdracht?

Gedurende het interview zijn de leerlingen bevraagd over hun schrijfproces. In totaal zijn 199 bo-leerlingen geïnterviewd en 72 sbo-leerlingen. Er is gevraagd naar hun aanpak bij aanvang van de taak, ervaren moeilijkheden en gebruik van hulpmiddelen tijdens het schrijven en of en zo ja, op welke manier zij hun teksten herzien. Om zicht te krijgen op de manier waarop leerlingen een schrijfopdracht aanpakken is hen een de hieronder afgebeelde, concrete opdracht voorgelegd, waarin leerlingen een brief aan een directeur moeten schrijven waarin zij hem moeten aan de hand van minimaal drie argumenten moeten overtuigen op schoolreis te gaan naar een pretpark. Gevraagd is hoe de leerlingen met deze opdracht zouden beginnen.

Schrijfpdracht

Er is een nieuw pretpark in Nederland met een geweldige achtbaan! Jullie willen hier graag naar toe met schoolreisje, maar het pretpark is best ver weg en ook duur. Schrijf een brief aan de directeur/directrice van jullie school waarin je vraagt of jullie tóch met schoolreisje naar dit pretpark mogen. Geef in elk geval drie goede redenen, waarmee je de directeur kunt overtuigen om naar dit pretpark te gaan. Samen met andere kinderen heb je al een goed plan bedacht om het geld bij elkaar te krijgen.



Op basis van de antwoorden van de leerlingen heeft de interviewer één of meer van zeven vooraf vastgestelde categorieën aangekruist. Deze zijn weergegeven in Tabel 6.78. In totaal zijn voor 197 bo-leerlingen en 72 sbo-leerlingen gegevens ingevuld bij deze vraag. Een kleine helft van de bo-leerlingen en een derde van de sbo-leerlingen geeft aan eerst na te denken over argumenten die ze kunnen gebruiken in hun brief. Naast deze inhoudelijke focus wordt ook de aanhef, een vormaspect van de brief, relatief vaak genoemd (bo: 25% en sbo: 32%). Een relatief groot deel van de leerlingen (36% van de bo-leerlingen en 29% van de sbo-leerlingen) begint gewoon direct met schrijven. De door de interviewers genoteerde steekwoorden laten verder zien dat veertien leerlingen meerdere versies willen maken. Ze zeggen eerst een kladversie en dan een nette versie te gaan schrijven/typen. Of leerlingen dan de gemaakte tekst een-op-een overnemen of daadwerkelijk reviseren is op basis van de steekwoorden niet duidelijk. Ook overleg met klasgenoten over de taak en het gebruik van formeel taalgebruik (in de aanhef of gebruik van formele woorden) wordt door leerlingen genoemd (resp. 9 en 8 leerlingen). Enkele leerlingen geven aan welke acties ze gaan ondernemen om het geld voor de schoolreis te verzamelen, zoals een sponsorloop/geldinzamelingsactie organiseren of het – naast het versturen van de brief – ook nog in gesprek gaan met de directeur. Een aantal leerlingen gaat in op correctheidsaspecten waar ze extra aandacht aan gaan besteden, zoals spelling, interpunctie en hoofdletters.

Tabel 6.78: Resultaten over aanpak van een schrijfpodracht (interviews leerlingen)

	bo (n=197)		sbo (n=72)	
	n	%	n	%
Ideeën (argumenten) verzamelen	87	44%	22	31%
Een tekstplan maken	32	16%	5	7%
Kijken hoe een brief er uit moet zien	6	3%	3	4%
De inleiding schrijven	17	9%	0	0%
Gelijk beginnen met schrijven	72	37%	21	29%
Adres opschrijven	1	1%	0	0%
De aanhef opschrijven	49	25%	23	32%
Anders	18	9%	6	8%

Tijdens het schrijven: ervaren moeilijkheden

Als een leerling heeft nagedacht over hoe hij/zij een opdracht aan wil pakken, begint de fase van het schrijven zelf. Gevraagd is of leerlingen tijdens het schrijven tegen dingen aanlopen die ze lastig of moeilijk vinden en of ze daar twee dingen/voorbeelden van kunnen noemen. De spontane antwoorden van de leerlingen zijn door de interviewers vertaald naar de in Tabel 6.79 gepresenteerde categorieën. Voor zes bo-leerlingen zijn geen antwoorden bij deze vraag ingevuld. De door zowel bo- als sbo-leerlingen meest genoemde moeilijkheden richten zich op de inhoud van de tekst (bo: 28%, sbo: 24%) en de correctheidsaspecten spelling (bo: 18% en sbo: 21%) en interpunctie (bo: 19% en sbo: 32%). In beide groepen is de categorie *anders* ook regelmatig aangekruist (bo: 31% en sbo: 40%). Op basis van de genoteerde steekwoorden kan worden opgemaakt dat 29 leerlingen geen moeilijkheden noemen (*'ik vind schrijven niet moeilijk'*). Ook wordt het netjes schrijven van een tekst door beide groepen leerlingen met regelmaat (n=17) genoemd, soms in combinatie met het moeten schrijven van lange teksten. Sommige leerlingen noemen dat het ze vooral lastig vinden om een tekst lang genoeg te maken of vertellen dat ze specifieke fonologische problemen hebben (zoals het structureel door elkaar halen van f en b of dyslexie). Andere genoemde problemen zijn het zelf veel moeten bedenken tijdens het schrijven, dat ze het moeilijk vinden om hun fantasie te gebruiken, met name bij verhalen, dat ze moeilijk zelfstandig op gang kunnen komen. Ook vinden enkele leerlingen sommige tekstsoorten lastig, zoals informatieve teksten, verhalen of samenvattingen.

Tabel 6.79: Aspecten die leerlingen tijdens het schrijven lastig vinden (interviews leerlingen)

	bo (n=193)		sbo (n=72)	
	n	%	n	%
Inhoud van de tekst	55	28%	17	24%
Schrijven met een bepaald doel	3	2%	0	0%
Schrijven voor een publiek (lezer)	6	3%	1	1%
Opbouw van de tekst	19	10%	1	1%
Begin van de tekst (inleiding)	22	11%	1	1%
Juiste woorden kiezen	17	9%	6	8%
Woorden goed schrijven (spelling)	35	18%	15	21%
Goede zinnen maken	23	12%	6	8%
Het middenstuk	4	2%	1	1%
Verbanden tussen zinnen maken	7	4%	1	1%
Het slot van de tekst	8	4%	3	4%
Interpunctie/leestekens	37	19%	23	32%
Op tijd klaar zijn	6	3%	0	0%
Anders	62	31%	29	40%

Bij de beschrijving van het gebruik van hulpmiddelen zijn de resultaten van de leerlingen aan bod geweest. Zoals uit Tabel 6.21b en de bijbehorende tekst kan worden afgeleid zegt twee derde van de bo-leerlingen hulpmiddelen te gebruiken, terwijl dit voor een kwart van de sbo-leerlingen het geval is.

Revisie: het herzien van teksten

Na het schrijven van een eerste versie is het belangrijk dat leerlingen hun tekst herlezen, hierop reflecteren en eventuele wijzigingen aanbrengen. De mate waarin leerlingen hun eigen teksten herzien en de manier waarop zij bij het herzien van teksten te werk gaan is tijdens het interview bevraagd. Voor het laatste is leerlingen een voorbeeld van een door een leerling geschreven voorbeeldtekst voorgelegd (door de interviewer, die de tekst ook voorlas). Dit is een door een fictieve leerling geschreven instructie voor het spelen van het spel Triominos (zie hieronder). De leerlingen is gevraagd of zij vinden dat de tekst verbeterd moet worden. Zo niet, dan dienden ze aan te geven waarom niet. Als de leerling wél vond dat de tekst aanpassing behoefde, is hem/haar gevraagd twee verbeterpunten te noemen. Op voorhand bleken er 26 missende waarden op de hoofdvraag te zijn (Vind je dat de tekst verbeterd moet worden?). De ingevulde subcategorieën lieten echter zien dat alle 26 missende waarden geïdentificeerd konden worden als 'wel verbetering nodig'.

Triomios

Eerst op je het spel uit de kast.

Dan doe je de doos open.

~~En de stadesdjes~~

Er kunnen twee tot vijf spelers mee doen

Je opent de stadesdjes er uit en elke speler moet vijf ~~to~~ blokjes met rijvers er op hebben.

En moet één van de spelers vragen of iemand drie de zelfde rijvers op één blokje ~~heeft~~ heeft staan.

Bijvoorbeeld driedrieën. nog wie dan ~~to~~ twee drieën en ga nog een rijver dan kan dat blokje er tegen aan legen en dan is de volgende.

De meerderheid van de sbo-leerlingen (78%) en de overgrote meerderheid van de bo-leerlingen (94%) vindt dat de aan hen voorgelegde tekst verbeterd zou moeten worden (zie Tabel 6.80). Wanneer ze dat niet vonden, werd met name naar voren gebracht de tekst 'helemaal goed zo' is of dat de 'uitleg duidelijk' is. Leerlingen die wel vonden dat de tekst herschreven zou moeten worden, benoemden vooral de netheid van de tekst (bo: 46%, sbo: 43%) en correctheidsaspecten als spelling (bo: 45%, sbo: 24%), correctheid van zinnen (bo: 38%) en het verbeteren van leestekens (bo: 14%). Ook het verbeteren van de inhoud wordt regelmatig genoemd, zowel bo- als sbo-leerlingen. De genoteerde steekwoorden laten zien dat leerlingen meermaals noemen dat het niet de bedoeling is te krassen in de tekst als toelichting dat ze vinden dat de leerling netter zou moeten schrijven. Meerdere leerlingen vinden dat de eerste (twee) zin(nen) niet sterk of zelfs overbodig zijn en zouden ze dan ook aanpassen of weghalen. Ook wordt regelmatig genoemd dat de uitleg onvoldoende is ('Ik snap het spel nog niet, dus de uitleg moet duidelijker'). Daarbij wordt het slot meermaals genoemd: dit noemen ze rommelig en het bevat volgens hen teveel informatie. Een heel ander type verbeterpunt wordt door twee leerlingen aangedragen: zij vinden de tekst saai en vinden dat het verlevendigd moet worden.

Tabel 6.80: Benodigde revisie van de voorbeeldtekst (interviews leerlingen)

	bo		sbo	
	n	%	n	%
Nee	12	6%	16	22%
	n	%	n	%
Ja	187	94%	56	78%
<i>Op welke manier?</i>				
Doelgerichter schrijven	27	14%	2	3%
Publiekgerichter schrijven	6	3%	0	0%
Structuur/opbouw verbeteren	36	18%	3	4%
Nummering gebruiken voor de stappen	12	6%	0	0%
Een tekening erbij maken	2	1%	1	1%
Inhoudelijk verbeteren (b.v. aangeven wanneer iemand wint of verliest)	49	25%	15	21%
Inhoud inleiding verbeteren/beter begin	28	14%	3	4%
Inhoud van het slot verbeteren/betere afsluiting	5	3%	4	5%
Netter schrijven	92	46%	31	43%
Overtypen	10	5%	5	7%
Spelfouten verbeteren	90	45%	17	24%
Zinnen verbeteren	75	38%	11	14%
Leestekens verbeteren	28	14%	7	10%
Anders	9	5%	5	7%

Ook is gevraagd of en hoe vaak leerlingen hun eigen teksten herzien. De resultaten staan weergegeven in Tabel 6.81. Ongeveer 90% van de leerlingen in beide schooltypen zegt soms of vaker de tekst te verbeteren. In het bo zegt zelfs één op de vijf leerlingen dit altijd te doen en in het sbo wordt dit ook door 14% gerapporteerd. Gekeken naar de toelichting in steekwoorden blijkt dat 19 leerlingen zeggen dat ze een klad- en nette versie maken, 34 leerlingen refereren aan correctheidsaspecten (spelling, zinnen en interpunctie) die ze moeten/willen veranderen en 16 leerlingen zeggen dat zij van de leerkracht een tweede versie moeten maken. Vaak is niet gespecificeerd welke veranderingen leerlingen dan moeten aanbrengen, maar één leerling geeft aan hoe ze te werk gaan: *Eerst maak ik een kladversie. Die bespreken we onderling [in duo's], dan maak ik een nette versie en die gaat naar de leerkracht. Die geeft dan commentaar.* Tien leerlingen herlezen hun eigen tekst kritisch en brengen dan veranderingen aan. Dit kan gaan om fout geschreven woorden, maar ook om meer inhoudelijke wijzigingen. Een voorbeeld daarvan wordt gegeven door deze leerling: *Ik spring soms van de hak op de tak en verbeter dan mijn eigen teksten.* Een andere leerling geeft aan richtlijnen te hebben voor het herzien van de tekst, *dan moet je terugkijken naar je tips en tops van de vorige les.* Leerlingen in die klas moesten tips voor de volgende keer noteren en hiervan gebruik maken bij de revisie van de nieuwe tekst. Redenen om niet de eigen tekst te verbeteren zijn *Ik ben meestal wel tevreden of meestal heb ik het al goed.*

Tabel 6.81: Mate waarin leerlingen hun eigen teksten verbeteren (interviews leerlingen)

		N	Nooit		Soms		Vaak		Altijd	
			n	%	n	%	n	%	n	%
Leerling verbetert eigen tekst	bo	197	22	11%	80	40%	57	29%	38	19%
	sbo	72	7	10%	39	54%	16	22%	10	14%

Hoofdstuk 7. Standaardbepaling

Zoals in hoofdstuk 5 is beschreven zijn de schrijfprestaties van de leerlingen af te beelden op een vaardigheidsschaal voor schrijfvaardigheid. Om deze vaardigheidsschaal beter te kunnen interpreteren en de vaardigheidsscores te kunnen relateren aan de referentieniveaus 1F en 1S/2F is een standaardbepaling uitgevoerd. Deze heeft in januari 2020 plaatsgevonden, onder regie van het College voor Toetsen en Examen (CvTE). Deze standaardbepaling heeft een advies aan het CvTE opgeleverd dat in het voorjaar 2020 door het CvTE is overgenomen. In dit hoofdstuk wordt het door Cito opgestelde rapport integraal weergegeven (Lampe & Hemker, 2020). Na een korte inleiding (paragraaf 7.1) volgt informatie over de voorbereidingsfase (paragraaf 7.2), de methode en selectie van de schrijftaken (paragraaf 7.3) en de procedure en resultaten (paragraaf 7.4). Paragraaf 7.5 bevat het advies dat is uitgebracht. Tevens bevat dit hoofdstuk enkele bijlagen behorend bij het rapport.

7.1 Inleiding

In de maanden april, mei en juni 2019 is de peiling Schrijfvaardigheid uitgevoerd. Onderdeel van het peilingsonderzoek was een standaardbepaling. Dit rapport beschrijft de uitkomsten van deze standaardbepaling. Voor nadere informatie over het peilingsonderzoek zelf verwijzen wij naar het technische rapport dat door het consortium wordt opgesteld en het publieksrapport dat door de Inspectie van het Onderwijs wordt geschreven. Deze rapporten zijn naar verwachting begin 2021 beschikbaar.

Binnen het peilingsonderzoek is met behulp van twaalf verschillende schrijfoopdrachten bij leerlingen in groep 8 van reguliere basisscholen en schoolverlaters in het sbo de schrijfvaardigheid in het Nederlands getoetst. In totaal namen van 134 scholen deel. Drie van de twaalf opdrachten (T01, T03 en T08) fungeerden als ankers waarmee de resultaten van de leerlingen in 2019 vergeleken konden worden met de leerlingprestaties tijdens de vorige peiling Schrijfvaardigheid in 2009. De overige negen taken waren nieuw ontwikkelde taken. In Tabel 7.1 staat een overzicht van enige (inhoudelijke) kenmerken van de twaalf taken.

Iedere taak is door 857 tot 1066 leerlingen gemaakt. In totaal zijn per taak tussen de 964 en 1195 beoordelingen voor analyse geselecteerd, inclusief circa 10% dubbele beoordelingen. 1.8% van de totale hoeveelheid taken kon niet worden geselecteerd voor de analyses (zie Tabel 5.3).

Tabel 7.1: Schrijftaken in Peil.Schrijfvaardigheid

Taak	Thema	Genre	Publiek	Doel	Teksthandeling
T01 (anker)	Wim	Verhaal	Leeftijdgenoten	Amuserend	Vertellen
T02	Verboden boek	Verhaal	Ongespecificeerd	Amuserend	Vertellen
T03 (anker)	Unicef	Samenvatting/stukje op basis van tekst	Leeftijdgenoten	Informatief	Beschrijven
T04	Nederland	Brief	Leeftijdgenoot uit ander land	Informatief	Beschrijven
T05	Nepnieuws	Tijdschriftartikel	Leeftijdgenoten	Informatief	Uitleggen
T06	Vlinder	'Biologietekst'	Iemand zonder plaatjes erbij	Informatief	Uitleggen
T07	Pannenkoeken	Recept	Leeftijdgenoten uit ander land	Instructief	Instrueren
T08 (anker)	Poekie	Flyer over verdwenen huisdier	Buurtgenoten	Instructief	Instrueren
T09	Interviewhulp	Brief	Leeftijdgenoot	Directief	Verzoeken
T10	Speelveld	Brief	Burgemeester	Directief	Verzoeken (en argumenteren)
T11	Kortere vakantie	Brief	Schooldirecteur	Persuasief	Argumenteren
T12	Frisdrank en snoep	Stuk in de krant	Krantenlezers	Persuasief	Argumenteren

Bij iedere taak zijn de schrijfproducten van leerlingen beschreven in observaties over de inhoud, opbouw en formulering. Voorbeelden van dergelijke observaties zijn 'Komt een bepaald cruciaal inhoudselement in de tekst voor?', 'Is de kern van het verhaal chronologisch opgebouwd?' of, bij formulering, 'Laat de schrijver je meeleven door intensiveerders, uitroepen of vragen?'. De observaties door beoordelaars zijn vervolgens gescoord als items, waarmee de score voor de taak bepaald is.

Om de resultaten op deze vaardigheden van deze leerlingen in perspectief te plaatsen is een standaardbepalingsonderzoek uitgevoerd. De procedure en verslaglegging van de standaardbepaling is uitgevoerd door het consortium dat het peilingsonderzoek Schrijfvaardigheid po uit voert, met het CvTE als gedelegeerd opdrachtgever. Het betrof een standaardbepaling voor de referentieniveaus 1F en 1S/2F, waardoor het mogelijk was om een inschatting te maken hoeveel leerlingen in groep 8 van het reguliere basisonderwijs (bo) en het speciaal basisonderwijs (sbo) in 2019 deze referentieniveaus haalden. Dit was mogelijk omdat in het peilingsonderzoek dezelfde twaalf taken in het bo en in het sbo zijn gemaakt. In deze rapportage beschrijven wij hoe deze standaardbepaling is uitgevoerd en wat de resultaten zijn.

Voorbehoud

Het is van belang te benadrukken dat dit de eerste keer is waarbij er een standaard voor schrijfvaardigheid wordt bepaald op basis van de referentieniveaus taal en rekenen. Alle mogelijke vergelijkingen met andere of eerdere niveaumetingen moeten om die reden daarom met de nodige voorzichtigheid gebeuren.

Omdat het hier productieve taken betreft, is bij deze standaardbepaling gebruik gemaakt van een taakgerichte methode om de cesuren te verkrijgen. Leerlingen hadden slechts een deel van de twaalf schrijfopdrachten gemaakt. Desalniettemin wilden we op basis van de prestaties een uitspraak doen over schrijfvaardigheid in het algemeen zoals geoperationaliseerd met deze taken. De aanname was dat de gevoerde operationalisatie zodanig was dat de standaarden ook betekenis hebben buiten de twaalf opgevoerde taken.

Hierbij is het belangrijk om ook rekening te houden met een beperking van de gehanteerde definitie van de schrijfvaardigheid. Binnen het peilingsonderzoek wordt schrijfvaardigheid primair gekarakteriseerd door de communicatieve effectiviteit van de teksten, die kan worden weergegeven door de drie aspecten inhoud, formulering en opbouw. Deze definitie bevat geen onderdelen die gaan over "correctheid". Daarom moeten deze standaarden met voorzichtigheid worden geïnterpreteerd.

Leeswijzer

De standaardbepaling is via een zorgvuldige procedure uitgevoerd door een expertpanel op 17 en 24 januari 2020. De standaardbepaling nam twee dagen in beslag. In paragraaf 7.2 beschrijven we het voorbereidende werk voor de standaardbepaling en de samenstelling van het panel. Paragraaf 7.3 gaat over de keuze voor een methode van standaardbepaling. Deze zal eerst algemeen worden besproken, waarna de stappen om tot de standaard te komen beschreven worden. Paragraaf 7.4 betreft vervolgens de procedure en de resultaten van de standaardbepaling op 17 en 24 januari 2020. Tot slot staat in paragraaf 7.5 het advies voor de standaarden van de referentieniveaus voor wat betreft schrijfvaardigheid op 1F en op 1S/2F aan het College voor Toetsen en Examen.

7.2 Voorbereidende activiteiten

7.2.1 Het expertpanel

Iedere vorm van standaardbepaling heeft een panel van experts nodig. Om tot een goede, reële en gedragen uitkomst te komen moet de samenstelling van zo'n panel aan een aantal eisen voldoen. Zo is er een voldoende aantal leden nodig, met voldoende spreiding over relevante kennis, achtergrond en ervaring.

Het panel dat voor het zetten van de standaarden schrijfvaardigheid is samengesteld, bestond in totaal uit 16 personen. Van deze 16 personen waren tien deskundig op het gebied van (onderzoek naar) schrijfvaardigheid. Het panel bestond niet alleen uit experts op het gebied van onderzoek naar schrijfvaardigheid, maar ook uit deskundigen werkzaam in het basisonderwijs en onderbouw voortgezet onderwijs (het ontvangende onderwijs), om zo de prestaties in perspectief te houden. Dit betrof zes velddeskundigen met ervaringen met leerlingen in (groep 8 van) het bo, in het laatste leerjaar van het sbo en in de onderbouw van het voortgezet onderwijs. Twee van deze zes panelleden hebben in de zomer van 2019 de schrijfopdrachten van de leerlingen beoordeeld: zij hebben enig zicht op werkelijke prestaties. De beschrijving van de standaarden bleek namelijk zoveel ruimte te laten

voor interpretatie, dat een realistisch perspectief op de leerlingprestaties voor de standaardbepaling noodzakelijk lijkt (zie Tabel 2). De ervaring leert dat taalexperts vaker aan de (te) hoge kant zitten met hun standaarden. Om die reden nemen ook personen aan de standaardbepaling deel die ervaringen met leerlingen hebben. Zij zullen wellicht minder bekend zijn met de referentieniveaus, maar zijn veel meer bekend met wat de leerlingen kunnen. Deze experts zijn hier aangeduid als velddeskundigen.

De procedure voor en de uiteindelijke selectie van uit te nodigen experts en velddeskundigen is afgestemd met het CvTE. Het CvTE is namelijk eindverantwoordelijk voor de samenstelling van het expertpanel. Voor het CvTE is het relevant rekening te houden met verschillende visies uit het veld met het oog op de kwaliteit en het draagvlak van het eindproduct.

7.2.2 De te gebruiken schalen om standaarden op te bepalen

Zoals in paragraaf 7.1 aangegeven is, bestond de vaardigheidsmeting uit twaalf taken. Een drietal van deze taken waren in 2009 ontwikkeld. Dit waren de ankertaken. Hoewel de drie ankertaken in 2009 niet waren ontwikkeld op grond van kenmerken van de referentieniveaus, waren voor deze drie taken wel standaarden te bepalen, omdat de beoordeling van deze taken ondergebracht kon worden op de te onderscheiden kenmerken inhoud, opbouw en formulering/stijl binnen de beoordelingsmodellen. Dat was zeker mogelijk bij de nieuw ontwikkelde taken. Bij het opstellen van de beoordelingsmodellen is rekening gehouden met een koppeling met de referentieniveaus. De beschrijving van de *referentieniveaus 1F en 2F* voor schrijven, zoals gedefinieerd in 2009¹⁰⁵ staat in Tabel 7.2. In het algemeen komt inhoud aan de orde in Tabel 7.2 bij de afstemming op doel. De opbouw refereert naar samenhang, naar de basisconventies bij een brief, en naar de leesbaarheid. Formulering/stijl betreft de afstemming op publiek en woordgebruik en woordenschat.

Daarnaast zijn alle taken ook beoordeeld op correctheid. Dit staat in Tabel 7.2 beschreven als spelling, interpunctie en grammatica. In Tabel 7.2 wordt verwezen naar taalverzorging, omdat dit kenmerk van het schrijfproduct niet is meegenomen in de evaluatie van de schrijfvaardigheid. Dat betekende dat correctheid geen onderdeel van de vaardigheidsschaal was, en dat de experts dit niet met de standaardbepaling meenamen. De beoordeling van spelling, interpunctie en grammatica is wel gebruikt als controlerende variabele. Bij een ordening van producten op de scores waar de standaard op bepaald is (zeg X), had een product met een hogere score X nooit een lagere score op correctheid dan een product met lagere score X.

¹⁰⁵ https://slo.nl/publish/pages/5901/referentiekader_taal_en_rekenen_referentieniveaus.pdf

Tabel 7.2: Beschrijvingen van de referentieniveaus behorend bij niveau 1F en 2F

	Niveau 1F	Niveau 2F
Kenmerken van de taakuitvoering		
Samenhang	<p>De informatie is zodanig geordend, dat de lezer de gedachtegang gemakkelijk kan volgen en het schrijfdoel bereikt wordt.</p> <p>De meest bekende voegwoorden (en, maar, want, omdat) zijn correct gebruikt, met andere voegwoorden komen nog fouten voor. Fouten met verwijswwoorden komen voor. Samenhang in de tekst en binnen samengestelde zinnen is niet altijd duidelijk.</p>	<p>Gebruikt veel voorkomende verbindingswoorden (als, hoewel) correct.</p> <p>De tekst bevat een volgorde; inleiding, kern en slot.</p> <p>Kan alinea's maken en inhoudelijke verbanden expliciet aangeven.</p> <p>Maakt soms nog onduidelijke verwijzingen en fouten in de structuur van de tekst.</p>
Afstemming op doel		<p>Kan in teksten met een eenvoudige lineaire structuur trouw blijven aan het doel van het schrijfproduct.</p>
Afstemming op publiek	<p>Gebruikt basisconventies bij een formele brief: Geachte/ Beste en Hoogachtend/Met vriendelijke groet.</p> <p>Hanteert verschil informeel/ formeel</p>	<p>Past het woordgebruik en toon aan het publiek aan.</p>
Woordgebruik en woordenschat	<p>Gebruikt voornamelijk frequent voorkomende woorden.</p>	<p>Fouten met voorzetsels en idiomatische uitdrukkingen komen nog voor.</p> <p>Varieert het woordgebruik.</p>
Spelling, interpunctie en grammatica	<p>Zie niveaubeschrijving Taalbeschouwing/Taalverzorging.</p>	<p>Zie niveaubeschrijving Taalbeschouwing/Taalverzorging.</p>
Leesbaarheid	<p>Hanteert titel.</p> <p>Voorziet een brief op de gebruikelijke plaats van datering, adressering, aanhef en ondertekening.</p> <p>Besteedt aandacht aan de opmaak van de tekst (handschrift, bladspiegel, eventueel beeldende elementen en kleur).</p>	<p>Gebruikt titel en tekstkopjes.</p> <p>Heeft bij langere teksten (meer dan twee A4) ondersteuning nodig bij aanbrengen van de lay-out.</p>

Om op basis van prestatie van leerlingen op enkele taken een uitspraak te kunnen doen over hun schrijfvaardigheid op alle taken, was een design opgezet waarbij leerlingen verschillende sets van taken maakten. Het was hierdoor mogelijk om met Item Response Theory (IRT) een schaling op te zetten om de taken te kunnen equivaleren. In dit onderzoeksdesign hebben leerlingen in het basisonderwijs vier van de twaalf taken gemaakt, waarbij zes combinaties van taken een “gelinkt” design opleverden. De sbo-leerlingen kregen twee van de taken, ook in een onderling gelinkt design, zodat het in principe mogelijk was om twee losse schalen voor bo en voor sbo te maken. Omdat de schaling geen verschil liet zien tussen bo en sbo, konden de beoordelingen van alle twaalf taken voor beide groepen op een enkele schaal worden geplaatst.

De schaalconstructie –inclusief de vertaling van observaties naar scores– is elders beschreven maar voor de standaardbepaling waren een aantal zaken van belang. Ten eerste dat het mogelijk bleek om een schaal te maken van alle observaties die de kwaliteit van de opbouw (O), formulering (F) en inhoud (I) meten: de OFI-schaal. De drie onderdelen O, F en I hingen onderling voldoende samen, waarbij met name de samenhang tussen O en F zeer sterk bleek te zijn. Het heeft een groot voordeel dat deze onderdelen op een schaal passen, omdat er zo duidelijk over een eenheid schrijfvaardigheid gesproken kan worden.

Een tweede belangrijk punt is dat alle twaalf taken op de schaal pasten: er werden bij enkele taken slechts een paar observaties niet meegenomen, omdat deze observaties bij de analyse geen samenhang vertoonden met de schrijfvaardigheid van een leerling. Het niet meenemen van deze observaties leverde een betere meting van het construct. Door toepassing van Item Response Theory is het bij iedere taak mogelijk om de standaard die op een taak bepaald is, te vertalen naar een score op de onderliggende latente schaal voor schrijfvaardigheid. Door IRT zijn ook de leerling-prestaties – en de bijbehorende geschreven teksten die als voorbeelden dienen –op een en dezelfde schaal te plaatsen. Door de gemiddelde scores op de schaal voor schrijfvaardigheid te berekenen bij de standaarden, bepaald met voorbeelden van prestaties op de taken, zijn ten slotte de twee standaarden voor schrijfvaardigheid als geheel bepaald.

Hierbij moet opgemerkt worden dat sommige taken de vaardigheid iets consistentier bleken te meten dan andere taken. Er zijn wat verschillen in betrouwbaarheid die voor een belangrijk deel veroorzaakt werden door verschillen in aantallen scorepunten per taak. Alle taken maten dus hetzelfde, alleen sommige taken deden dat iets beter (consistentier) dan andere taken.

Als gekeken wordt naar de mogelijke scores die gebruikt kunnen worden op een taak, dan was te zien dat niet de gehele scoreschaal gebruikt werd. De prestatie van meer dan 95% van de leerlingen kon beschreven worden door minder dan 50% van de mogelijke scores op de scoreschaal¹⁰⁶. Ook was te zien dat de taken 06, 11 en 12 vooral de onderkant van de vaardigheidsschaal beschreven en de taken 04, 05 en 08 de bovenkant van de schaal. De taken 06, 11 en 12 zijn dan ook gebruikt voor de bepaling van referentieniveau 1F en de taken 05 en 08 voor referentieniveau 1S/2F (zie Tabel 7.3). Over het algemeen leken de taken meer geschikt om de minder vaardige leerlingen te onderscheiden dan de meer vaardige leerlingen¹⁰⁷. Op basis van deze uitkomsten zijn taken geselecteerd om met name referentieniveaus 1F, 2F dan wel beiden te bepalen.

¹⁰⁶ Kijken we per taak dan zie we 95% van de leerlingen scores halen die 40% tot 60% van de scoreschaal van de taak dekken.

¹⁰⁷ Ruim 60% van de relevante verwachte scorepunten ligt in het deel van de vaardigheidsschaal waar de 50% slechtst presterende leerlingen zit, en minder dan 40% in het gebied van de schaal dat de vaardigheid van de 50% best presterende leerlingen beschrijft.

7.3 Standaardbepaling: methode en selectie schrijfproducten

7.3.1 Methode van de standaardbepaling

Voor het bepalen van een standaard zijn twee zaken cruciaal: een goede omschrijving van de standaard, en een operationalisatie van de vaardigheid die we willen meten. Bij de standaardbepaling komen deze zaken samen. Dit proces is weergegeven in Figuur 7.1. Het vormen van een expertpanel en de beschrijving van Tabel 7.2 gaan over de rechterrouten in Figuur 7.1, namelijk dat de standaarden bekend moeten zijn bij de experts. Aangezien de beschrijvingen van Tabel 7.2 redelijk algemeen zijn, was het belangrijk om de referentieniveaus door middel van training te concretiseren.

Ter voorbereiding kregen de experts daarom voorafgaand aan de standaardbepaling een beschrijving van de referentieniveaus toegestuurd en een handreiking voor de koppeling met de observatiemodellen met de opdracht om deze beschrijving en handreiking grondig te bestuderen. De standaardbepaling startte met een zogenaamde ‘familiarisatiefase’: dit was een inhoudelijke bespreking van de kenmerken van de referentieniveaus 1F en 1S/2F, in combinatie met twee oefeningen waarbij de experts voorbeelden van schrijfproducten (op papier) in oplopende volgorde van schrijfvaardigheid beoordeelden. Na elke oefening vond een plenaire bespreking plaats van de resultaten van de oefening en de dilemma’s die de experts hierbij hebben ervaren. Deze fase beoogde het kennisniveau van de experts over de referentieniveaus te verhogen. Voor de familiarisatie is het beoordelingsformulier in bijlage 1 van dit hoofdstuk gebruikt.

Figuur 7.1: Stappen richting het bepalen van een standaard



De beschrijving van de manier hoe een schaal gemaakt is op basis van de scores van de schrijfproducten van leerlingen (de linker route in Figuur 1) is hiervoor en ook elders beschreven. Deze schaal werd zichtbaar tijdens de standaardbepaling door de schrijfproducten van leerlingen in oplopende volgorde van schrijfvaardigheid aan de experts aan te bieden. Daarnaast kregen de experts na elke eerste beoordeling een grafische weergave van de posities van de voorbeelden op de schaal in de vorm van een figuur (zie bijvoorbeeld Figuur 4).

De beoordelaars registreerden de waarneembare kenmerken of eigenschappen van het schrijfproduct. Deze registraties werden vervolgens omgezet in scores, die een totaalscore op die taak opleverden. Voor de experts waren die afzonderlijke kenmerken te lastig als eenheid te beoordelen vanwege de grote onderlinge samenhang (en soms afhankelijkheid) van de verschillende kenmerken. Het schrijfproduct van een leerling is de kleinst mogelijke eenheid die een expert kon beoordelen: het product dat voortkomt uit de schrijftaak vormde namelijk een geheel. Bij de standaardbepaling vormde daarom het schrijfproduct, dat wil zeggen de leerling-respons, de te beoordelen eenheid.

De schrijftaak was dan ook minder goed op te delen in afzonderlijke en eenvoudig te observeren kenmerken, zoals bijvoorbeeld inhoud, opbouw en stijl/formulering: de moeilijkheid van elk afzonderlijk (analytisch) kenmerk (of item) was daardoor niet goed in te schatten. Een itemgerichte vorm van standaardbepaling was dan ook niet toepasbaar. De standaard is daarom niet bepaald op grond van de evaluatie van afzonderlijke kenmerken, maar op basis van de gehele schrijfproducten van de leerlingen. Dit betekende dat de gebruikte standaardbepaling te kenschetsen was als een holistische productgerichte methode.

Een methode om standaarden te bepalen voor productieve taken is de Body of Work methode (Kingston, Kahl, Sweeney, & Bay, 2001)¹⁰⁸. Bij deze holistische standaardbepalingsmethode worden alle werken van een kandidaat meegenomen om tot een oordeel te komen over het beoogde niveau. Deze methode diende als inspiratie voor de toegepaste methode maar had als onoverkomelijk nadeel dat de methode specifiek gericht is op alle taken die de leerling heeft gemaakt. Hoewel het evident is dat de algemene schrijfvaardigheid van de individuele leerling beter te bepalen is als er meer gegevens van de leerling worden betrokken, was er echter een aantal redenen om dat in dit geval niet te doen.

De belangrijkste reden was dat het voor een expert lastig is om de verschillende prestaties op de verschillende taken te wegen. Leerlingen maken immers niet alle taken precies even goed. Als de prestaties van een leerling variëren over taken kan dat de algehele beoordeling beïnvloeden. Ervaringen hierbij suggereren dat een relatief slechtere prestatie het algemene beeld van de leerling negatief beïnvloedt, meer dan dat een relatief goede prestatie het beeld doet bijstellen. Ook volgorde-effecten waarbij de eerste en de laatste schrijftaak zwaarder meetellen, of de gevallen waarin een halo-effect eventueel een rol kan spelen als de eerste taak juist heel goed is gemaakt (en een horn-effect wanneer deze juist heel slecht is) kunnen optreden bij de beoordelingen van de verschillende schrijfproducten van een leerling. Bij de reguliere beoordelingen was dat geen probleem aangezien de beoordelingen gesegmenteerd per taak plaatsvonden: de producten van de leerlingen zijn telkens door verschillende beoordelaars beoordeeld.

Een andere reden waarom het niet wenselijk was dat de panelleden alle producten van een leerling beoordelen was dat het lezen van de producten waarschijnlijk relatief lang kon duren: een beoordelingstijd van vier taken zou in het kader van de standaardbepaling naar verwachting tussen 10 en 15 minuten duren. Het blijven onthouden wat men van de eerder gelezen teksten vond en het correct wegen van alle teksten is een complexe mentale taak. Om dat vervolgens consequent op dezelfde wijze te doen voor een groot aantal leerlingen (geheugen-effecten) is zonder extreem lange training schier onmogelijk.

Aangezien we de taak voor de experts niet complexer wilden maken dan nodig was en zo de kans op systematisch goede beoordelingen zo groot mogelijk wilden maken, hebben de panelleden de standaardbepaling in eerste instantie per taak uitgevoerd. Een reeks voorbeelden van schrijfproducten

¹⁰⁸ Kingston, N. M., Kahl, S. R., Sweeney, K., & Bay, L. (2001). Setting performance standards using the Body of Work method. In G. J. Cizek (Ed.), *Standard setting: Concepts, methods, and perspectives* (pp. 219-248). Mahwah, NJ: Erlbaum.

bij een taak zijn steeds in oplopende volgorde van schrijfvaardigheid aangeboden, waarbij de experts moesten aangeven bij welk schrijfproduct voor het eerst het bevroegde niveau gehaald is.

De werkwijze lijkt op een bookmark-methode, die ook bij de centrale examens wordt gebruikt¹⁰⁹. Bij een bookmark methode worden de opgaven geordend naar moeilijkheid en geven de experts aan op welke positie in deze rangorde de leerling presteert wanneer deze net het beoogde niveau heeft. Bij de hier gebruikte methode zijn niet de opgaven geordend maar de voorbeelden van de schrijfproducten en geven de experts aan vanaf welk schrijfproduct leerlingen het beoogde niveau tonen. Bij een bookmarkmethode zijn de oordelen per object (item, schrijfproduct) niet onafhankelijk van elkaar. Wanneer een panellid een voorbeeld van een schrijfproduct als (nét) voldoende beoordeelt voor het beoogde niveau, dan zijn de schrijfproducten met hogere scores dat ook, en als een schrijfproduct als onvoldoende ingeschat wordt dan geldt dat ook voor de producten met lagere scores.

Deze werkwijze zou tot een conflict kunnen leiden wanneer een expert zou vinden dat een hoger gescoord product van lagere kwaliteit was dan een lager gescoord product, bijvoorbeeld door elementen in de tekst die te maken hebben met correctheid (grammatica, interpunctie, spelling) die niet meetellen in de evaluatie. Het voordeel van de bookmarkmethode is in dit geval dan ook dat de ordening van teksten volledig uitgevoerd is op basis van de relevante aspecten, en dat de juiste dimensies van oordelen gebruikt zijn. Er gaan daardoor geen omwisselingen plaats vinden waardoor een lager gescoord product wel als van voldoende niveau beoordeeld wordt en een hoger gescoord product niet.

De ongewogen totaalscores vormden de basis van de ordening in de standaardbepaling. De maximaal te behalen totaalscore verschilde per schrijftaak en was minimaal 16 en maximaal 50 scorepunten. De methode van standaardbepaling is dan ook 'data-driven': de (empirische) informatie van de reeds verkregen scores van de beoordelaars is direct bij de standaardsetting gebruikt.

Eerst zijn de leerlingprestaties per schrijftaak geclusterd zodat er prototypische voorbeelden geselecteerd konden worden van iedere schrijftaak per (relevant) scorepunt. Het aantal voorbeelden om uit te kiezen kon verschillen per score. De extreme scores zijn niet of nauwelijks geobserveerd, ook niet met onze aantallen producten (rond de 1000 per taak). Voor het bepalen van de standaarden was dit geen cruciaal probleem aangezien de grenzen van 1F en 1S/2F niet direct in deze extremen verwacht werden.

Gemiddeld bleken er per taak tussen 12 en 13 scorepunten relevant. Het relevante gebied is gelegd in de range waar slechts 1% van de sbo-leerlingen lager scoort en 2% van de bo-leerlingen hoger scoort. Voor alle twaalf taken zijn sets van 5 tot 8 voorbeelden geselecteerd om in deze procedure te gebruiken. Bij de selectie van voorbeelden van schrijfproducten is primair rekening gehouden met (1) het beoordelaarseffect en (2) de vaardigheid van de kandidaat, om te kunnen bepalen of de voorbeelden goede typische voorbeelden waren. Secundair is er ook op gelet dat er bij de voorbeelden per taak enige spreiding is wat betreft de kenmerken van de leerling.

De score moet een reflectie zijn van de leerlingprestatie en niet van de beoordelaar. Daarom zijn de schrijfproducten alleen geselecteerd uit de 10% schrijfproducten die door twee onafhankelijke beoordelaars waren beoordeeld. Bij deze producten konden we direct per product evalueren of het oordeel onafhankelijk van de beoordelaar was: wanneer twee beoordelaars tot (bijna) hetzelfde oordeel kwamen konden we aannemen dat dit een goede weergave was van de kwaliteit van het product. De prestaties van leerlingen waar twee beoordelaars het (vrijwel volledig) met elkaar eens

¹⁰⁹ <https://wetten.overheid.nl/BWBR0037590/2017-04-06/>

waren –in ieder geval op totaalscoreniveau en bij voorkeur ook op itemniveau– waren goede voorbeelden omdat we van die leerlingen beter wisten dat die score een goede reflectie was van de schrijfvaardigheid van de leerling dan van leerlingen waar de beoordelaars het (sterk) niet met elkaar eens waren.

Een bijkomend voordeel van de schrijfproducten was dat we niet een enkel product per leerling hadden maar twee tot vier producten. We kenden daarmee niet alleen de prestatie op de taak, maar ook de prestatie van de leerling over alle taken heen. Net zoals bij de Body of Work-methode konden we zo alle gegevens van een kandidaat gebruiken omdat we bij de selectie steeds hebben gekeken of de prestatie op de taak representatief was voor deze leerling. Dit had een aantal voordelen. Ten eerste konden we zo op basis van de beoordeelde prestaties ten behoeve van de standaardbepaling alle beschikbare informatie over de vaardigheid van de kandidaat meenemen. Ten tweede was het door te selecteren op een representatieve score aannemelijk dat deze score ook een reflectie was van de werkelijke vaardigheid: op basis van de overige taken die de leerling gemaakt heeft hadden andere beoordelaars ook een vergelijkbare vaardigheid bij de leerling vastgesteld¹¹⁰. Hiermee hebben we ook het beoordelareffect meegenomen. We konden namelijk naast homogeniteit over taken ook nog voor iedere kandidaat evalueren of de kandidaat homogeen was wat betreft de samenstellende scores O, F en I. Zo kregen we representatieve, evenwichtige voorbeelden.

Merk op dat deze procedure vergelijkbaar was met de aanbeveling bij de ‘Body of Work’-methode door Cizek en Bunch (2007, chapter 9, p. 125) dat de prestatie (score) over de verschillende opdrachten die een (geselecteerde) kandidaat heeft uitgevoerd niet al te veel in kwaliteit mag verschillen. Dat is de werkwijze binnen de Body of Work waarbij alleen kandidaten geselecteerd worden met homogene prestaties op de verschillende taken¹¹¹. Als de prestaties over de taken vergelijkbaar zijn kan evenzogoed een enkele taak beoordeeld worden door de experts. De ordening van de taken zou daarmee identiek zijn, alleen zou de expert niet een taak zien, maar vier taken die allen ook een identieke ordening zouden hebben omdat deze daarop geselecteerd zijn. Aangezien het aantal taken een viervoud is zal de beoordeling ook veel langer duren, terwijl de meerwaarde beperkt zou zijn. Dit zou gemakkelijk een verdubbeling van de kosten opleveren terwijl de zekerheid over de oordelen niet verdubbeld zou worden. Doordat alle taken in de procedure opgenomen zijn, is de mogelijke heterogeniteit van de beoordelingen over de taken ook gedekt.

Per serie voorbeelden van een taak is in tweede instantie gecontroleerd of alle voorbeelden niet afkomstig zijn van leerlingen van dezelfde school of van een te beperkte selectie van scholen. Zo is een regionale spreiding van voorbeelden verkregen. Daar waar er leerlingkenmerken bekend waren –zoals schooltype– zijn die bij de selectie vooraf meegenomen zodat de voorbeelden indien mogelijk representatief verdeeld waren over BO en SBO. Merk op dat deze spreiding minder relevant is bij de beoordeling van schrijfproducten dan bij het beoordelen van spreektaken waarbij de kandidaten zelf ook direct geobserveerd worden.

¹¹⁰ Wanneer een kandidaat vier producten heeft gemaakt dan zijn er in de meeste gevallen ook vier beoordelaars, met enkele uitzonderingen waarbij er twee producten door een beoordelaar beoordeeld zijn wat een totaal van drie beoordelaars oplevert. Bij de kandidaten met twee productie (SBO) zijn er altijd twee verschillende beoordelaars.

¹¹¹ Dit wordt als oplossing gepresenteerd om te voorkomen dat experts afgeleid worden door extreme producten van een kandidaat om tot hun oordeel te komen.

7.3.2 Dubbele controle vooraf van oplopende volgorde schrijfproducten

De standaardbepaling bestond steeds uit twee rondes per referentieniveau: een gezamenlijke ronde (1F/2F), gevolgd door een ronde voor referentieniveau 1F, respectievelijk 1S/2F. Tijdens elke ronde kregen de experts voorbeelden van zes schrijftaken te beoordelen. Bij de selectie van deze voorbeelden is in eerste instantie gestreefd naar een verdeling van de aantallen voorbeelden, zoals weergegeven in de kolom 'Na te streven' in Tabel 7.3. Zo was het de bedoeling dat Taak 01 zowel in Ronde 1 als Ronde 2 voor 1F werd gebruikt en dat de experts tijdens de standaardbepaling in beide rondes vijf voorbeelden van schrijfproducten van Taak 01 te beoordelen zouden krijgen. Het streven was dat een expert in ronde 1 in totaal 36 voorbeelden (gemiddeld zes voorbeelden per taak) voorgelegd zou krijgen en dat de experts in totaal 107 voorbeelden voorgelegd zouden krijgen.

Tijdens de standaardbepaling was het, vanwege de beperkte hoeveelheid tijd, niet wenselijk dat de experts twijfels hadden over de oplopende vaardigheid van de voorgelegde voorbeelden van schrijfproducten. Daarom is de volgorde van de geselecteerde voorbeelden van schrijfproducten drie maal gecontroleerd voorafgaand aan de feitelijke standaardbepaling.

Allereerst hebben vier inhoudelijke experts, werkzaam bij Cito, de oplopende volgorde van de geselecteerde voorbeelden van schrijfproducten beoordeeld. De volgorde van de schrijfproducten is steeds beoordeeld door wisselende combinaties van twee experts. Tijdens een plenaire bespreking op donderdag 6 december zijn de volgordes besproken waar vraagtekens bij werden gezet, en is een selectie gemaakt van die schrijfproducten, waarover beide betrokken experts het eens konden worden.

Vervolgens is gecontroleerd of de geselecteerde producten niet alleen op grond van de som van de ruwe scores, maar ook op grond van de schaalscores op basis van de gewogen somscores (WML-score) op de taak een oplopende volgorde vertoonden. Dit was noodzakelijk om de posities van de voorbeelden van de schrijfproducten op de schaal voor schrijfvaardigheid grafisch te kunnen weergeven zoals bijvoorbeeld in Figuren 7.3 en 7.4.

Tenslotte heeft een groep van twee externe experts en een expert werkzaam bij Cito tijdens een pilot op vrijdag 13 december nogmaals beoordeeld of de ordening van de gekozen voorbeelden inderdaad oplopend was in vaardigheid zoals geëvalueerd door de onderzoekers en de experts werkzaam bij Cito. De voorbeelden, waarvoor geen eenduidigheid bestond dat zij een oplopende volgorde weergaven, zijn vervolgens niet geselecteerd voor de standaardbepaling. Dit resulteerde in een aanpassing van de selectie van de schrijfproducten bij twaalf van de 18 voor te leggen reeksen schrijfproducten. Daarnaast bleek dat de kwaliteit van de scan van 16 schrijfproducten moest worden vergroot.

Tijdens deze pilot is tevens de procedure beproefd van de training tijdens de familiarisatiefase en de beoordeling van de schrijfproducten. Voor de beoordeling van zes reeksen van voorbeelden van schrijfproducten in oplopende volgorde bleken de deelnemers aan de pilot ongeveer anderhalf uur nodig te hebben. Dit betekende dat het niet nodig bleek om het aantal taken in de eerste ronde wat te beperken om meer tijd voor discussie over te houden. De gehele eerste ronde overslaan is onwenselijk omdat het ten eerste een goed zicht geeft op de relatie tussen 1F en 2F, en ten tweede omdat op basis van de resultaten de discussie over de referentieniveaus en de schrijfopdrachten meer inhoudelijk kan worden.

Tabel 7.3: Taken gebruikt per ronde, met na te streven en geselecteerd aantal voorbeelden per taak.

Taak \ Ronde	Na te streven			Geselecteerd		
	R1: 1F/2F	R2: 1F	R2: 2F	R1: 1F/2F	R2: 1F	R2: 2F
Taak 01	5	5		4	5	
Taak 02		5			4	
Taak 03	5		7	4		4
Taak 04	6		7	5		7
Taak 05			5			5
Taak 06		7			4	
Taak 07	8	8		8	6	
Taak 08			5			5
Taak 09	7		7	4		6
Taak 10	5		5	4		4
Taak 11		5			5	
Taak 12		5			4	
Totaal	36	35	36	29	28	31

Tijdens de pilot bleek het lastig om de oplopende volgorde van de voorbeelden bij Taak 03 over Unicef te beoordelen, omdat deze taak een informatieve tekst over Unicef bevatte. Bij leerlingen die grote delen van deze tekst verwerkten in hun schrijfproduct leek de woordenschat erg groot te zijn, maar deze woorden bleken in de tekst in de opdracht voor te komen. Tijdens de standaardbepaling moest bij de beoordeling van de schrijfproducten bij Taak 03 daarom tijd worden ingeruimd voor het maken van afspraken over de beoordeling van de voorbeelden van de schrijfproducten.

7.4 Procedure en resultaten standaardbepaling

De standaardbepaling startte op 17 januari met een kennismakingsronde en een korte training, waardoor de experts konden wennen aan het type taken, het beoordelingskader, en de beoordeling van de schrijfproducten. Hierdoor konden zij ook beter begrijpen hoe de ordening tot stand is gekomen, en welke elementen daarin van belang waren. Met name het niet meenemen van “correctheid” in het oordeel behoeft enige training. Dat is de reden dat er ook voldoende tijd voor het trainen van de experts is uitgetrokken in de procedure van de standaardbepaling.

Vervolgens zijn tijdens de training de referentieniveaus kort toegelicht en heeft een familiarisatie plaats gevonden waarbij de experts twee maal een reeks schrijfproducten in oplopende volgorde kregen voorgelegd met het verzoek om te bepalen welk schrijfproduct voldoende is voor het behalen van referentieniveau 1F, respectievelijk 1S/2F. Een voorbeeld van een blad dat hierbij ingevuld is staat in Bijlage 1.

De resultaten van de oordelen van de twee reeksen schrijfproducten zijn plenair besproken en de argumenten van de experts op grond waarvan ze de schrijfproducten beoordeelden zijn verzameld en besproken. De training en familiarisatiefase bleken in totaal 2,5 uur in beslag te nemen.

De argumenten voor de keuze van referentieniveau 1F bleken te zijn:

- Inhoud is het allerbelangrijkste argument.
- De gedachtegang in de tekst moet te volgen zijn.
- De gedachtegang bevat een probleembeschrijving en overtuigende argumenten.
- Het is belangrijk dat er pogingen in staan om redeneerwoorden, voegwoorden en verwijswaarden te gebruiken. Samenhang is ook belangrijk, maar dit is alleen te beoordelen wanneer een product als een brief of verhaal twee of meer volledige zinnen bevat.

Afgesproken is om coulance te betrachten bij de beoordeling of een leerling alle conventies van een brief goed toepast. Bij een formele brief is een aanhef wel belangrijk, het ontbreken van een afsluiting is minder belangrijk, omdat leerlingen voor het maken van de schrijftaken de instructie hadden gekregen om niet met hun eigen naam te ondertekenen in verband met de privacy-wetgeving (AVG).

Voor de keuze voor referentieniveau 2F waren bovenop de argumenten bij 1F de volgende argumenten belangrijk:

- Bereikt de leerling het doel? Maakt de leerling bijvoorbeeld echt een spannend (echt avontuur, levendig) verhaal, en geen saai verhaal en/of verslag van gebeurtenissen. Het verhaal bevat details en consequenties van tijd.
- Is het verhaal samenhangend? Dat wil zeggen dat het verhaal volledige zinnen bevat, met liefst een opbouw in alinea's, en dat niet alles duidelijk of logisch hoeft te zijn.
- Komen alle onderdelen aan bod, met een duidelijke verhaalstructuur (inleiding-plot-einde)?

Andere, minder belangrijke criteria zijn de variatie in woordgebruik, het gebruik van meer soorten verbindingswoorden en een passende tijdswisseling of perspectiefwisseling in verhaal.

7.4.1 Eerste ronde 'rangefinding'

Vervolgens hebben de experts op 17 januari 's middags een reeks schrijfproducten beoordeeld, waarbij ze moesten aangeven welke schrijfproducten het referentieniveau 1F behaalden en welke schrijfproducten het referentieniveau 1S/2F. De eerste ronde van de standaardbepaling is de ronde die in de Body of Work methode wordt aangeduid als 'rangefinding' (Cizek & Bunch, 2007, p. 125). De experts kregen in deze ronde vijf tot acht voorbeeldproducten van taken voorgelegd in oplopende volgorde van schrijfvaardigheid op basis van de schaalcores op de taak. Op basis van de vijf tot acht voorbeelden moest de expert voor elke afzonderlijke schrijftaak twee absolute oordelen geven. Bijlage 2, aan het einde van dit hoofdstuk, bevat de eerste pagina van het invulformulier in deze ronde.

De eerste vraag luidde: Is het schrijfniveau in het schrijfproduct van deze leerling al op referentie-niveau 1F? De tweede luidde: Is het schrijfniveau in het schrijfproduct van deze leerling al

op referentieniveau 1S/2F? Hierbij speelden twee ordeningen een rol. Ten eerste die van de niveaus. De niveaus 1F en 1S/2F zijn geordend. Als een product niet voldeed aan het niveau 1F, dan voldeed het ook niet aan 1S/2F, en als het voldeed aan 1S/2F dan voldeed het ook aan de eisen van niveau 1F.

Een andere manier waarop de volgorde een rol speelde is de ordening van producten. Het product met het laagste nummer dat voldeed aan het niveau, impliceerde dat het nummer eronder niet voldeed. De expert gaf zo ook het hoogste nummer van het product dat niet voldeed. Dit is ook expliciet bij de standaardbepaling omdat de experts in deze ronde per taak en per referentieniveau een verticale lijn (streepje) moesten zetten op een lijn met de nummers van de beoordeelde schrijfproducten. Dat streepje gaf de plaats weer waar de grens lag tussen de twee schrijfproducten die de expert als onvoldoende en als voldoende beschouwde op het beoogde referentieniveau. Merk op dat de precieze positie tussen de twee nummers niet meegenomen wordt: voor de verwerking is alleen het eerste (=laagste) nummer van het schrijfproduct op referentieniveau 1F, respectievelijk 1S/2F ingevoerd in tabel 4 en 5. Het oordeel van de experts was dus op een discrete schaal, en niet op een continue schaal. De resultaten van de eerste beoordeling staan in Tabel 7.4.

De experts 14, 15 en 16 konden de eerste dag niet bijwonen en hebben zelf thuis de voorbeelden van schrijfproducten op vier van de zes taken beoordeeld. De twee niet beoordeelde taken zijn grijs gearceerd in Tabel 4. Deze experts bleken de schrijfproducten bij referentieniveau 1F milder te hebben beoordeeld dan de dertien experts die wel de bijeenkomst op 17 januari konden bijwonen.

De getallen onder de horizontale lijn geven een aantal samenvattende beschrijvingen van de 16 of minder oordelen per schrijftaak weer. De mediaan is het middelste nummer na een sortering van de 16 of minder oordelen op grootte en P75-P25 is het verschil tussen de scores op 25% en 75% van deze sortering. De modus is het eerste nummer dat het meest als oordeel is gegeven. 'Aantal' geeft het aantal keren weer dat de modus als oordeel is gegeven.

Veel experts gaven het oordeel dat geen van de voorgelegde schrijfproducten referentieniveau 1S/2F weerspiegelde. De groen gemarkeerde cellen in de kolom Referentieniveau 1S/2F in Tabel 7.4 geven dit oordeel weer. Bij al deze beschrijvingen onder de horizontale lijn zijn deze lege cellen meegenomen als groter dan het hoogste nummer ('>4' bij Taak 01, 03, 09 en 10, en '>5' bij Taak 04).

De experts zijn redelijk unaniem in hun eerste oordelen over referentieniveau 1F bij de producten bij Taak 01, 03, 04, en 10: bij elke taak bleken negen of meer experts de grens te leggen bij hetzelfde product. Bij Referentieniveaus 1S/2F zijn 14, 12 en 10 experts gelijk in hun oordeel bij Taak 01, 09 en 10.

Tabel 7.4: Eerste oordelen over eerste nummers schrijfproducten per taak, op referentieniveau 1F, dan wel 1S/2F.¹¹²

Nr.	Referentieniveau 1F						Referentieniveau 1S/2F					
	T01(W)	T03(U)	T04(A)	T08(P)	T09(N)	T10(S)	T01(W)	T03(U)	T04(A)	T08(P)	T09(N)	T10(S)
1	3	2	4	6	4	3	4	4	5	7	>4	4
2	2	3	3	3	3	3	4	3	4	7	>4	4
3	3	3	4	3	4	3	4	>4	5	5	>4	4
4	3	3	4	4	4	3	>4	3	>5	8	>4	>4
5	3	3	4	4	4	4	4	>4	>5	8	>4	>4
6	3	3	4	3	4	3	4	>4	5	6	>4	4
7	3	3	4	5	3	3	4	>4	5	8	4	4
8	2	2	2	4	4	2	4	3	5	6	>4	4
9	3	3	4	5	3	3	>4	4	>5	7	>4	>4
10	3	3	3	4	2	3	4	>4	>5	7	4	>4
11	4	3	4	4	2	3	4	3	>5	6	>4	>4
12	4	3	4	5	4	2	4	>4	>5	8	>4	4
13	3	2	3	6	2	3	4	3	4	7	4	4
14	2		3		1	3	4		4		>4	4
15	2		2		2	3	4		5		4	>4
16	2		3		2	3	4		4		>4	4
mediaan	3	3	4	4	3	3	4	4	5	7	>4	4
P75-P25	1	0	1	1	2	0	0	-	-	2	-	-
modus	3	3	4	4	4	3	4	>4	5*	7	>4	4
aantal	9	10	9	5	7	13	14	6	6	5	12	10

5* = Taak 04 heeft twee modi: de oordelen 5 en >5 zijn door zes experts gegeven.

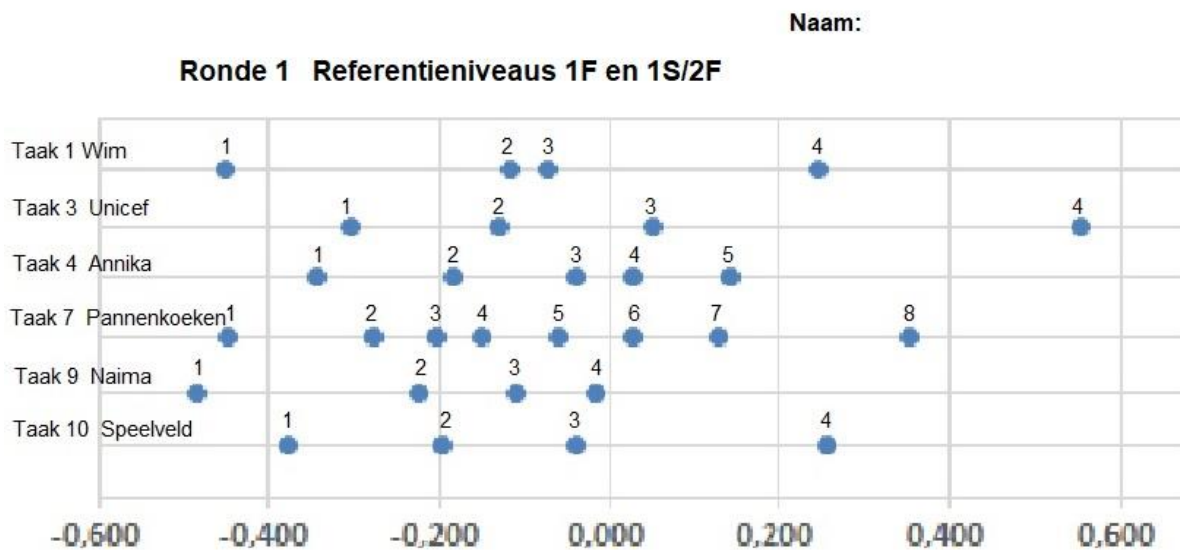
Na de eerste beoordeling van de volgordes van de schrijfproducten op de zes taken kregen de experts informatie over de relaties tussen de prestaties van de leerlingen op de schrijftaken.. Figuur 7.3 is een voorbeeld van de representatie van de empirische data die de experts kregen voorgelegd in de eerste ronde. In dit Figuur komen de nummers op de scoreschalen (=horizontale lijnen) overeen met de nummers van de voorgelegde schrijfproducten tijdens de standaardsetting (bijvoorbeeld de eerste 1 op de bovenste horizontale lijn geeft de positie weer van Schrijfproduct 1 bij Taak01 over Wim). De positie van (de nummers van) elk schrijfproduct op de scoreschalen weerspiegelt de positie van de vaardigheid¹¹³ op de desbetreffende schrijftaak.

In Figuur 7.3 kunnen experts precies zien hoe de schrijfproducten zich tot elkaar verhouden. De experts kunnen op basis van dit figuur ook een indruk hebben over hoe consistent ze zijn. Deze methode heeft als groot voordeel dat de schrijftaken en de empirische resultaten van echte leerlingen op deze taken gecombineerd worden, waardoor de beoordelingen van experts naar verwachting realistischer en consistentere zullen zijn dan bij andere methodes.

¹¹² Nr = Nummer expert; T01(W) = Wim; T03(U) = Unicef; T04(A) = Annika; T07(P) = Pannenkoeken; T08(N) = Naima; T10(S) = Speelveld

¹¹³ Deze vaardigheid is bepaald met een weighted maximum likelihood (WML)-score uit de IRT

Figuur 7.3: De positie van schrijfproducten op de schaal voor schrijfvaardigheid, zoals gepresenteerd tijdens eerste ronde van standaardsetting Peil.Schrijfvaardigheid.



Elke expert kreeg Figuur 7.3, waarbij de gemaakte keuzes voor referentieniveau 1F groen en voor 1S/2F rood waren omcirkeld. Naar aanleiding van de resultaten in Tabel 7.4 en hun eigen Figuur 7.3 hebben de experts gediscussieerd over de gemaakte keuzes. Deze werkwijze gaf hen de mogelijkheid de eigen oordelen te vergelijken met die van de andere experts.

Belangrijk hierbij was dat het resultaat van de eerste ronde met de experts gezamenlijk is besproken aan de hand van hun oordelen in een representatie van de beoordeelde schrijfproducten, zoals in Figuur 7.3. De ervaring leert namelijk dat de oordelen van experts erg van elkaar (en van de empirie) kunnen afwijken op grond van verschillende argumenten. Uit empirische gegevens kan vervolgens blijken dat een schrijftaak een heel andere verdeling van verschillende scores blijkt te hebben voor kandidaten dan verondersteld.

Op grond van de resultaten in Figuur 7.3 kregen de panelleden de mogelijkheid om hun oordelen in hun grafiek bij te stellen. Dit leverde de resultaten in Tabel 7.5 op.

Tabel 7.5: Oordelen na discussie over eerste nummers schrijfproducten per taak, op referentieniveau 1F, dan wel 1S/2F.¹¹⁴

Nr.	Referentieniveau 1F						Referentieniveau 1S/2F					
	T01(W)	T03(U)	T04(A)	T08(P)	T09(N)	T10(S)	T01(W)	T03(U)	T04(A)	T08(P)	T09(N)	T10(S)
1	3	2	4	6	4	3	4	>4	5	7	>4	4
2	2	3	3	4	3	3	4	3	5	7	>4	4
3	3	3	4	5	4	3	4	>4	5	7	>4	4
4	3	3	4	5	4	3	>4	4	>5	8	>4	>4
5	3	3	4	4	4	4	4	>4	>5	8	>4	>4
6	3	3	4	6	4	3	4	>4	5	8	>4	4
7	3	3	4	5	3	3	4	>4	5	8	4	4
8	2	2	2	4	4	2	4	3	5	6	>4	4
9	3	3	4	5	3	3	>4	>4	>5	7	>4	>4
10	3	3	3	4	2	3	4	>4	>5	7	>4	>4
11	3	3	3	4	3	3	4	>4	>5	7	>4	>4
12	4	3	4	5	4	3	4	>4	>5	8	>4	4
13	3	2	3	6	2	3	4	>4	>5	7	>4	4
14	2		3		3	3	4		5		>4	4
15	2		2		2	3	4		5		4	4
16	2		3		2	3	>4				>4	4
mediaan	3	3	3,5	5	3	3	4	>4	5*	7	>4	4
P75-P25	1	0	1	1	1,25	0	0	-	-	1	-	-
modus	3	3	4	5	4	3	4	>4	5*	7	>4	4
aantal	10	10	8	5	7	14	13	10	8	7	14	11

5* = zowel voor mediaan als voor modus geldt, dat de oordelen 5 en >5 door acht experts zijn gegeven.

Duidelijk is dat na de bespreking de beoordelaars over het algemeen meer op dezelfde lijn zaten qua beoordeling: de verschillen in scores tussen het derde (=P75) en het eerste kwart (=P25) van de oplopend gesorteerde scores werden kleiner¹¹⁵ en de aantallen beoordelaars die de modus kozen bleken groter.

De beoordelaars 14, 15 en 16 konden de eerste dag niet bijwonen en hebben zelf thuis de voorbeelden van schrijfproducten op vier taken beoordeeld. Deze beoordelaars bleken de schrijfproducten bij referentieniveau 1F milder te hebben beoordeeld dan de dertien experts die wel de bijeenkomst op 14 januari konden bijwonen. Dit resultaat is met deze drie experts voorafgaand aan de tweede bijeenkomst besproken en vervolgens door hen bijgesteld aan de hand van hun resultaten in Figuur 7.3.

Hoewel de werkelijke standaard niet in deze ronde is gezet, bood het resultaat van de eerste ronde een indicatie van de bandbreedte waarin het niveau van 1F ligt, en een waar het niveau 1S/2F ligt ('rangefinding'; Cizek & Bunch, 2007) op de OFI_WML scoreschaal voor Schrijfvaardigheid. De grensschaalscore ligt voor elk oordeel tussen de schaalscore voor het schrijfproduct, voorafgaand aan het eerste nummer en de schaalscore voor het eerste nummer van een schrijfproduct op referentieniveau 1F of 1S/2F. Dit is berekend door het gemiddelde te nemen van de schaalscores van deze twee producten. Wanneer een expert vond dat alle producten in een reeks nog niet aan referentieniveau 1S/2F voldeden, dan is de schaalscore van het product met het hoogste nummer als ondergrens voor de grensschaalscore gehanteerd.

¹¹⁴ Nr=Nummer expert; T01(W)=Wim; T03(U)=Unicef; T04(A)=Annika; T07(P)=Pannenkoeken; T09 (N)=Naima; T10(S)=Speelveld

¹¹⁵ Dit was ook het geval als andere percentielen gekozen waren zoals P90 en P10 of P67 en P33.

Op grond van de gemiddelde grensschaalscores van de zes taken ligt referentieniveau 1F op -0,116 van de schaal Schrijfvaardigheid en het referentieniveau 1S/2F op 0,070 op deze schaal. Op basis van de scoreverdeling in het basisonderwijs¹¹⁶ van de schaal betekent dit dat gemiddeld 69% van de populatie leerlingen in de peiling op of boven referentieniveau 1F presteert en dat gemiddeld 30% van deze leerlingen op of boven referentieniveau 1S/2F presteert.

Tabel 7.6: Gewogen gemiddelde schaalscores en percentages behaald in (speciaal) basisonderwijs van gekozen schrijfproducten per taak, op referentieniveau 1F en 1S/2F tijdens eerste ronde 'Rangevinding'.

Schrijftaak	Referentieniveau 1F			Referentieniveau 1S/2F		
	OFI_WML score	%BO	%SBO	OFI_WML score	%BO	%SBO
Wim	-0,142	74%	39%	0,086	27%	10%
Unicef	-0,081	62%	29%	0,074	30%	11%
Annika	-0,077	61%	29%	0,084	28%	10%
Pannenkoeken	-0,111	68%	34%	0,132	20%	7%
Naima	-0,169	79%	44%	-0,063	59%	27%
Speelveld	-0,115	69%	35%	0,108	24%	9%
Gemiddeld	-0,116	69%	35%	0,070	30%	12%
Minimum	-0,169	79%	44%	-0,063	59%	27%
Mediaan	-0,113	69%	34%	0,085	28%	10%
Maximum	-0,077	61%	29%	0,132	20%	7%

Het advies voor de standaard is bepaald in de volgende twee rondes op 24 januari 2020.

7.4.2 Tweede en derde ronde 'pinpointing'

De bijeenkomst op 24 januari startte met een presentatie van de resultaten na de eerste ronde 'Rangevinding'. Op grond van de resultaten voor 1F merkten twee panelleden, die niet bij de eerste dag aanwezig waren, op dat het overeengekomen niveau voor 1F te hoog lag. 1F was immers, zoals een van deze twee panelleden zei 'het niveau dat 75% van de leerlingen in groep 8 in 2008 zou moeten halen'.

Na een korte discussie over het feit dat deze dagen de absolute standaard voor 1F wordt bepaald op grond van de beschrijving van de referentieniveaus, en niet op basis van de relatieve verdeling van de leerlingresultaten, kregen de panelleden nieuwe voorbeelden van schrijfproducten in oplopende volgorde op zes schrijftaken ter beoordeling voorgelegd. De vraag was weer: Is het schrijfniveau in het schrijfproduct van deze leerling al op referentieniveau 1F? De resultaten van deze eerste beoordeling staan in Tabel 7.7.

Na de eerste beoordeling van de volgordes van de schrijfproducten op de zes taken kregen de experts ook informatie over de relaties tussen de prestaties van de kandidaten op de schrijftaken zoals gegeven in Figuur 7.4, waarin de door de experts gemaakte keuzes voor referentieniveau 1F waren omcirkeld. In Figuur 7.4 konden experts precies zien hoe de schrijfproducten zich tot elkaar verhouden. De beoordelaars konden op basis van deze figuur ook een indruk krijgen over hoe consistent ze zijn.

¹¹⁶ De schaalscores OFI_WML vertonen voor het BO een gemiddelde van -0,024 en een standaarddeviatie van 0,183 en voor SBO een gemiddelde van -0,206 en een standaarddeviatie van 0,231

Naar aanleiding van de resultaten in Tabel 7.7 en hun eigen Figuur 7.4 hebben de experts gediscussieerd over de gemaakte keuzes. Deze werkwijze gaf hen de mogelijkheid de eigen oordelen te vergelijken met die van de andere experts. Op grond van de resultaten in Figuur 7.4 kregen de panelleden de mogelijkheid om hun oordelen in hun grafiek bij te stellen. Dit leverde de resultaten in de kolom 'Tweede oordeel' in Tabel 7.7 op.

Een aantal experts wilde geen keuze maken tussen twee producten en gaf daarom halve waarden op. Dit zijn de geel gemarkeerde cellen in Tabel 7.7. Aangezien de vraag was om het eerste voorbeeld op niveau 1F aan te geven zijn in de tabel deze halve waarden met een half verhoogd. Daarmee wordt steeds het product met de hoogste schaalscore als ondergrens gehanteerd.

Tabel 7.7: Eerste volgnummers schrijfproducten per taak, op referentieniveau 1F.

Nr.	Eerste oordeel						Tweede oordeel						Derde oordeel					
	W ¹¹⁷	V	VI	P	K	F	W	V	VI	P	K	F	W	V	VI	P	K	F
1	3	3	3	4	3	3	3	3	2	4	3	3	3	3	2	4	3	3
2	2	3	2	6	3	4	2	3	2	6	3	3	2	2	2	2	3	3
3	3	2	4	4	3	4	2	2	3	4	3	4	2	2	2	4	3	3
4	4	2	3	6	4	3	4	2	3	6	4	4	2	2	3	6	3	3
5	4	3	3	4	3	4	3	3	3	4	3	3	3	3	3	4	3	3
6	3	4	4	6	4	4	2	4	4	6	4	3	2	2	2	4	3	3
7	4	3	2	6	3	4	3	3	2	6	3	4	2	3	2	5	3	3
8	3	2	4	6	4	3	3	2	3	6	4	2	2	2	4	6	3	3
9	2	2	3	6	4	4	2	2	3	6	4	3	2	2	3	5	3	3
10	4	3	3	6	3	3	2	3	3	6	3	3	4	3	3	3	3	3
11	2	2	2	5	3	2	2	2	2	5	3	3	2	2	2	3	2	2
12	3	4	4	6	4	3	3	4	3	6	4	3	2	4	3	5	3	3
13	4	3	2	5	3	3	2	2	2	5	3	3	2	2	2	2	2	3
14	2	2	3	2	4	3	2	2	3	2	3	3	2	2	2	2	3	3
15	2	2	3	4	3	3	2	2	2	4	3	3	2	2	2	4	3	3
16	3	2	3	4	3	3	3	2	3	5	3	3	3	2	3	2	2	2
mediaan	3	2,5	3	5,5	3	3	2	2	3	5,5	3	3	2	2	2	4	3	3
P75-P25	2	1	0,5	2	1	1	1	1	1	2	1	0	0,25	1	1	2,25	0	0
modus	3	2	3	6	3	3	2	2	3	6	3	3	2	2	2	4	3	3
aantal	6	8	8	8	10	9	9	9	9	8	11	12	12	11	9	5	13	14

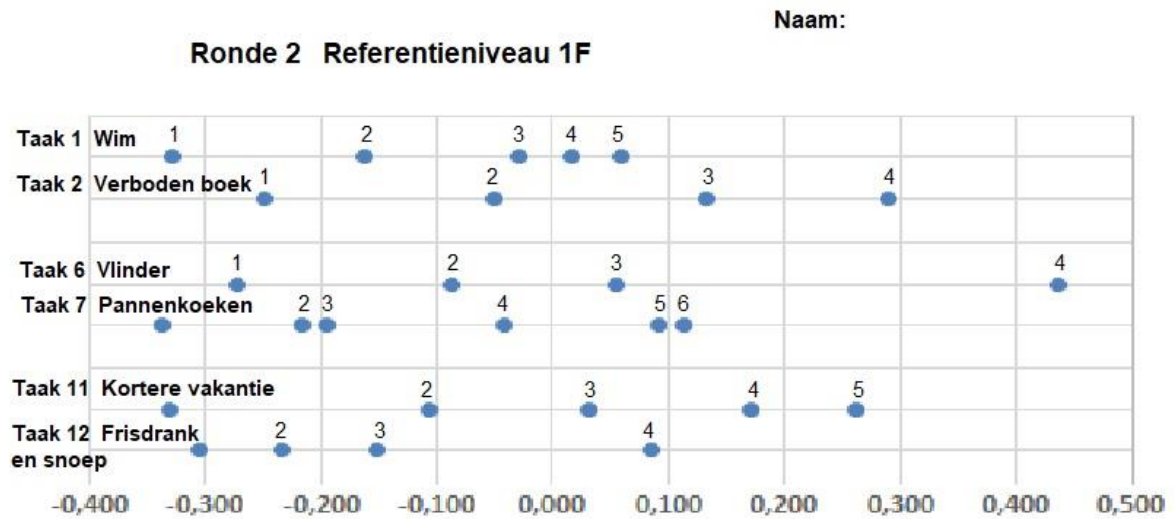
Taak 01 (Wim) was op 17 januari gebruikt ter familiarisatie van referentieniveau 2F (en vervolgens ook gebruikt voor de bepaling van referentieniveau 1F, zie Tabel 7.4 en 7.5). Tijdens de discussie van de eerste beoordeling van referentieniveau 1F op 24 januari bleken enkele experts abusievelijk de afspraken rondom referentieniveau 2F te hebben toegepast bij de beoordeling van Taak 01 (Wim).

Vanwege deze verwarring en de onvrede van twee panelleden met de volgens hen te hoge standaard voor referentieniveau 1F is besloten om eerst duidelijke afspraken te maken over de precieze inschatting van het beoogde referentieniveau 1F (en van het referentieniveau 1S/2F), en vervolgens thuis nogmaals de voorbeelden van schrijfproducten te beoordelen voor de bepaling van referentieniveau 1F. Deze oordelen staan in de kolom 'Derde oordeel' in Tabel 7.7. Duidelijk is de grotere consensus van de panelleden zichtbaar bij de vier taken Wim, Verboden boek, Kortere vakantie en Frisdrank en snoep: elf of meer panelleden kozen hetzelfde schrijfproduct als op referentieniveau 1F. Alleen bij de taak Pannenkoeken bleken de oordelen meer verdeeld dan in de eerste twee rondes. De gemiddelde schaalscore bij deze taak ligt echter het dichtst bij het gemiddelde van alle schrijftaken,

¹¹⁷ Nr = Nummer beoordelaar; W = T01: Wim; V = T02: Verboden boek; VI = T06: Vlinder; P = T07: Pannenkoeken; K = T11: Kortere vakantie; F = T12: Frisdrank en snoep

daarom worden de oordelen op deze schrijftaak wel meegenomen bij het advies aan het College voor Toetsen en Examen.

Figuur 7.4: De positie van schrijfproducten op de schaal voor schrijfvaardigheid, zoals gepresenteerd tijdens ronde 2 van standaardsetting Peil.Schrijfvaardigheid.



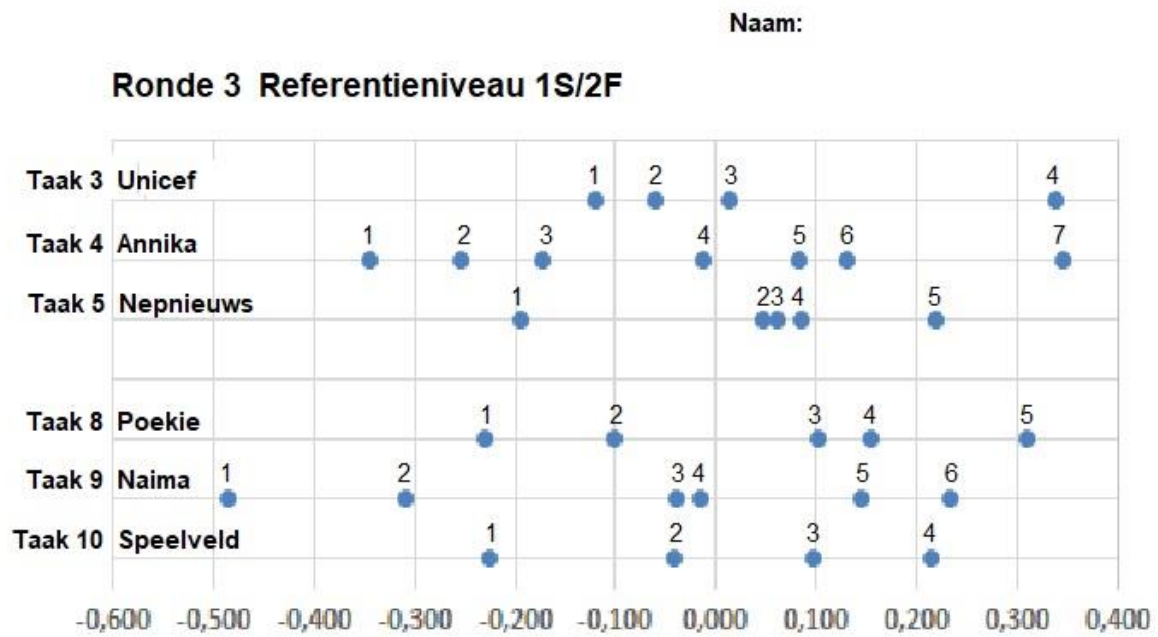
De discussie leverde de volgende inhoudelijke beschrijving van de referentieniveaus 1F en 1S/2F op, zoals weergegeven in Tabel 7.8.

Tabel 7.8: Kenmerken van schrijfproducten, nét op referentieniveau 1F en nét op 1S/2F.

Kenmerken schrijfproduct nét op 1F	Kenmerken schrijfproduct nét op 2F
<p>Inhoud is het meest belangrijk, dat wil zeggen</p> <ul style="list-style-type: none"> • De boodschap moet overkomen • Gedachtegang is te volgen. <p>Tekstsoort (brief, etc) moet herkenbaar zijn.</p> <p>Ondergeschikt hieraan zijn de volgende kenmerken:</p> <ul style="list-style-type: none"> • De meest bekende voegwoorden en, maar, want en omdat zijn correct gebruikt. Eventueel verkeerd gebruik van andere voegwoorden mag niet aangerekend worden. • (Afstemming op) Doel in 1F niet expliciet opgenomen, alleen bij Samenhang. Doel hoeft niet volledig bereikt te worden, juiste richting is al snel goed. • Zinsgrenzen zijn aangegeven (hoofdletter en punten). • Fouten in verwijswaarden mogen voorkomen, maar stelselmatige systematische fouten niet, want dan is gedachtegang niet meer te volgen. • Woordvolgorde binnen zinnen is redelijk correct. • Tekst bevat schrijftaal, geen spreektaal. • Verschil informeel versus formeel is passend bij aanhef en afsluiting (bijvoorbeeld ‘Hoi burgemeester...’). 	<p>Zie 1F, plus</p> <p>Opbouw past bij tekstsoort.</p> <p>Woordgebruik en toon zijn aangepast aan lezer. Bijvoorbeeld, de hele brief is formeel van toon.</p> <ul style="list-style-type: none"> • Bereikt de leerling het doel: bijvoorbeeld maakt de leerling echt een spannend (echt avontuur, levendig) verhaal, en geen saai verhaal of verslag van gebeurtenissen, consequenties van tijd, bevat details • Is de tekst samenhangend? Dat wil zeggen dat het verhaal volledige zinnen bevat, met liefst een opbouw in alinea’s, niet alles hoeft duidelijk of logisch te zijn. • Komen alle onderdelen aan bod, met een duidelijke verhaalstructuur (inleiding-plot-einde)? <p>Andere, minder belangrijke criteria zijn de variatie in woordgebruik, het gebruik van meer soorten verbindingswoorden en een passende tijdswisseling of perspectiefwisseling in verhaal.</p>

Na de discussie hebben de panelleden de voorbeelden van schrijfproducten op referentieniveau 1S/2F beoordeeld. De oordelen zijn vervolgens wederom weergegeven in de vorm van Figuur 7.5 waarin de onderlinge posities van de voorbeelden van de schrijfproducten zijn weergegeven. Na een korte discussie hebben de panelleden hun oordelen bijgesteld. De resultaten van deze bijstelling zijn weergegeven in Tabel 7.9.

Figuur 7.5: De positie van schrijfproducten op de schaal voor schrijfvaardigheid, zoals gepresenteerd tijdens ronde 3 van standaardsetting Peil.Schrijfvaardigheid.



Ook hier zijn bij de halve scores van experts de eerst volgende gehele score gebruikt om aan te geven wat volgens deze experts het eerste voorbeeld was op het niveau 1S/2F. Dit zijn de geel gemarkeerde cellen in Tabel 7.9.

Tabel 7.9: Nummers schrijfproducten per taak, nét op referentieniveau 1S/2F.

Nr.	Eerste oordeel						Tweede oordeel					
	U ¹¹⁸	A	N	P	Na	S	U	A	N	P	Na	S
1	3	5	3	5	6	3	3	5	3	5	5	3
2	3	5		5	6	3	4	5	6	5	6	3
3		7	4	3	5	4	5	5	4	3	5	4
4	4	4	3	5	4	4	4	4	3	5	4	4
5	4	5	5	2	5	4	4	5	5	2	5	4
6	4	5	3	5	6	3	4	5	3	5	6	3
7	4	5	3	4	4	3	4	5	3	4	4	3
8	4	4	2	4	4	3	3	4	3	4	4	3
9	4	5	3	4	5	3	4	5	3	4	5	3
10	4	5	4	3	5	3	4	5	4	3	5	3
11		5	5	3	4	2	5	5	5	3	4	3
12	3	5	4	3	5	3	3	5	4	3	5	3
13	4	5	5	5	5	3	4	5	5	5	5	3
14		5	5	5	5	3	5	5	5	5	5	3
15	3	5	4	4	4	3	3	5	4	4	5	3
16	4	5	5	5	5	3	4	5	5	5	5	3
mediaan	4	5	4	4	5	3	4	5	4	4	5	3
P75-P25	1	0	2	2	1	0	0,25	0	2	2	0,25	0
MODUS	4	5	5	5	5	3	4	5	3	5	5	3
aantal	9	13	5	7	8	12	9	14	6	7	10	13

Na de bespreking zaten de tweede bijgestelde oordelen meer op dezelfde lijn qua beoordeling: de percentielen 75 en 25 lagen dicht bij elkaar en de aantallen experts die de modus kozen bleken groter. Met name bij de schrijftaken Annika, Naima en Speelveld was de overeenstemming tussen de panelleden groot.

7.5 Advies referentieniveaus 1F en 1S/2F

De resultaten van de derde oordelen voor 1F en de tweede oordelen voor 1S/2F vormen de basis voor het advies voor de standaarden van de Referentieniveaus Schrijfvaardigheid 1F en 1S/2F aan het College voor Toetsen en Examens ('pinpointing'; Cizek & Bunch, 2007).

Op grond van het gewogen gemiddelde van de schaalscores van de twee opeenvolgende nummers van de schrijfproducten net onder en op de gekozen grens luidt het advies:

1. Stel de cesuur voor het behalen van het referentieniveau 1F vast bij de score -0,128 van de schaal Schrijfvaardigheid
2. Stel de cesuur voor het behalen van het referentieniveau 1S/2F vast op de score 0,088 op deze schaal.

¹¹⁸ Nr = Nummer beoordelaar; U = Unicef; A = Annika; N = Nepnieuws; P = Poekie; Na = Naima; S = Speelveld

Op basis van de scoreverdeling¹¹⁹ van de schaal betekent dit dat gemiddeld 72% van de populatie leerlingen in het basisonderwijs (BO) in de peiling op of boven referentieniveau 1F presteert en dat gemiddeld 27% van deze leerlingen op of boven referentieniveau 1S/2F presteert. De resultaten staan in Tabel 7.10.

Tabel 7.10: Gewogen gemiddelde schaalscores en percentages behaald in (speciaal) basisonderwijs van gekozen schrijfproducten per taak, op referentieniveau 1F en 1S/2F tijdens tweede en derde ronde 'Pinpointing'.

1F	Taak	schaal	%BO	%SBO	1S/2F	Taak	schaal	%BO	%SBO
T01	Wim	-0,208	84%	50%	T03	Unicef	0,172	14%	5%
T02	Verboden boek	-0,079	62%	29%	T04	Annika	0,019	41%	16%
T06	Vlinder	-0,092	64%	31%	T05	Nepnieuws	0,104	24%	9%
T07	Pannenkoeken	-0,114	69%	34%	T08	Poekie	0,123	21%	8%
T11	Kortere vakantie	-0,071	60%	28%	T09	Naima	0,057	33%	13%
T12	Frisdrank en	-0,202	84%	49%	T10	Speelveld	0,051	34%	13%
	Gemiddeld	-0,128	72%	37%		Gemiddeld	0,088	27%	10%
	Minimum	-0,208	84%	50%		Minimum	0,019	41%	16%
	Mediaan	-0,103	67%	33%		Mediaan	0,080	29%	11%
	Maximum	-0,071	60%	28%		Maximum	0,172	14%	5%

De resultaten van de uiteindelijke oordelen tijdens de 'Pinpointing'-fases in Tabel 7.10 vertonen grote gelijkens met de resultaten van de eerste ronde 'Rangefinding' in Tabel 7.6. Voor 1F is de gemiddelde schaalscore iets omlaag gegaan van -0,116 naar -0,128. Dit betekent dat na pinpointing 72% (in plaats van 69%) van de bo-populatie op en boven referentieniveau 1F presteert. Voor de standaard voor het referentieniveau 1S/2F is de gemiddelde schaalscore iets omhoog gegaan, van 0,070 naar 0,088. Dit betekent dat de schrijfvaardigheid van 27% van de onderzochte leerlingen op of boven referentieniveau 1S/2F ligt. In Tabel 7.6 betrof dit 30% van de onderzochte bo-leerlingen.

Dit advies is ter instemming voorgelegd aan alle experts die hebben deelgenomen aan de standaardbepaling. 15 van de 16 experts konden met dit advies instemmen, een expert bleek niet meer mee te kunnen werken vanwege persoonlijke omstandigheden. Een expert wilde bij haar instemming de volgende kanttekening plaatsen, namelijk dat in onderzoek vaker verschillen over typen taken gezien worden. De expert vroeg zich af in hoeverre het realistisch is om één referentieniveau vast te stellen op basis van deze gegevens en zou liever de referentieniveaus per genre vaststellen.¹²⁰ Hierbij kan echter wel gemeld worden dat bij deze standaardbepaling de niveaus gerelateerd aan de

¹¹⁹ De schaalscores OFI_WML vertonen voor het BO een gemiddelde van -0,024 en een standaarddeviatie van 0,183 en voor SBO een gemiddelde van -0,206 en een standaarddeviatie van 0,231

¹²⁰ Expert: "Ik zie wel dat er grote verschillen zijn tussen schrijftaken, zoals we eigenlijk altijd zien in onderzoek. Ik vraag me dan ook af in hoeverre het realistisch is om één referentieniveau vast te stellen op basis van deze gegevens en of we niet meer recht doen aan de complexiteit van de vaardigheid door de referentieniveaus per genre vast te stellen."

vaardigheidsschaal niet sterk over de taken verschilden. De toepassing van de IRT zorgt er namelijk voor dat verschillen in complexiteit meegenomen worden in de berekening van de schaalscores.

7.6 Bijlagen bij dit hoofdstuk omtrent de standaardsetting

Bijlage 7.1: Formulier Familiarisatie Schrijven 1F

Centraal staat de volgende vraag:

- Welk schrijfproduct beschouwt u als **nét** voldoende op het referentieniveau 1F?

U krijgt vier schrijfproducten in oplopend vaardigheidsniveau bij een opgave.

- Bekijk elk schrijfproduct en bepaal of u het schrijfproduct als (**nét**) voldoende beschouwt op het referentieniveau 1F.
- Omcirkel uw oordeel en geef daarbij een korte toelichting.

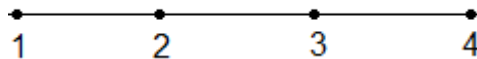
Taak 10 Speelveld	1F-niveau ¹	Toelichting
Schrijfproduct 1	nee / ja	
Schrijfproduct 2	nee / ja	
Schrijfproduct 3	nee / ja	
Schrijfproduct 4	nee / ja	

¹ Omcirkel uw oordeel:

nee = Schrijfproduct is onder 1F-niveau,

ja = Schrijfproduct is **nét** op 1F-niveau (of hoger)

Zet een verticale lijn waar volgens u de grens ligt tussen de twee schrijfproducten die u als onvoldoende en als **nét** voldoende beschouwt op het referentieniveau 1F.



Geef een korte toelichting bij de grens, die u heeft aangegeven.

Bijlage 7.2: Eerste pagina formulier eerste ronde standaardsetting Schrijven (1F,1S/2F)

U krijgt vier tot acht schrijfproducten te zien bij zes verschillende taken.

De schrijfproducten staan in oplopende schrijfvaardigheid (van laag naar hoog).

- Bekijk eerst elk schrijfproduct en bepaal vervolgens welk schrijfproduct u als (nét) voldoende beschouwt op het referentieniveau 1F, en welk schrijfproduct u als (nét) voldoende beschouwt op het referentieniveau 1S/2F. Omcirkel uw oordelen.

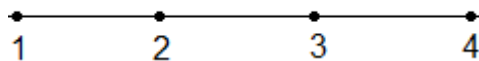
Taak 1 Wim	1F-	1S/2F-	Toelichting en aantekeningen
Schrijfproduct 1	nee / ja	nee / ja	
Schrijfproduct 2	nee / ja	nee / ja	
Schrijfproduct 3	nee / ja	nee / ja	
Schrijfproduct 4	nee / ja	nee / ja	

¹Omcirkel uw oordeel:

nee = Schrijfproduct is onder 1F-niveau, dan wel 1S/2F-niveau;

ja = Schrijfproduct is nét op (of hoger dan) 1F-niveau, dan wel 1S/2F-niveau.

Zet een verticale lijn waar volgens u de grens ligt tussen de twee schrijfproducten die u als onvoldoende en als nét voldoende beschouwt op het referentieniveau 1F. Dito voor 1S/2F.



Geef een korte toelichting bij de grenzen, die u heeft aangegeven.

Hoofdstuk 8. Resultaten

Belangrijk doel van het peilingsonderzoek is om de schrijfvaardigheid van de leerlingen in groep 8 en van de schoolverlaters in het sbo in kaart te brengen. Niet alleen brengen we de huidige resultaten in beeld, maar we vergelijken ook de prestaties van de leerlingen met die van de vorige peiling uit 2009. Om die vergelijking mogelijk te maken was een deel van de taken in beide peilingen gelijk. Dit hoofdstuk gaat in op de resultaten van de leerlingen aan de hand van de beantwoording van de hieronder genoemde onderzoeksvragen. Onderzoeksvragen, die gericht zijn op de relatie tussen de schrijffprestaties en kenmerken van scholen, leerkrachten, leerlingen en het onderwijsleerproces, worden in hoofdstuk 9 beantwoord met behulp van meerniveau-analyses.

1. Hoeveel procent van de leerlingen einde basisonderwijs en schoolverlaters speciaal basisonderwijs beheerst de referentieniveaus (1F en 2F) voor schrijfvaardigheid en wat is de vaardigheidsverdeling op de onderliggende schaal?
2. Hoe kan de prestatie van een leerling worden getypeerd als het gaat om de communicatieve effectiviteit, zoals blijkend uit de kenmerken van de taakuitvoering? Beschrijf de range van leerlingprestaties met specifieke aandacht voor de grenspunten <1F-1F en 1F-2F, de laagvaardige (de 10% laagst scorende leerlingen), de gemiddelde en de hoogvaardige (de 10% hoogst scorende leerlingen) leerlingen door middel van een selectie van voorbeeldprestaties uit leerling producten.
3. Welke verschillen zijn er tussen de leerlingprestaties anno 2019 en de leerlingprestaties in het peilingsonderzoek schrijfvaardigheid van 2009?

Allereerst worden in paragraaf 8.1 de leerling resultaten weergegeven in termen van de standaarden op de referentieniveaus 1F en 1S/2F die ten behoeve van het peilingsonderzoek zijn bepaald (zie hoofdstuk 7). We presenteren percentages leerlingen die onder 1F, op 1F en minimaal op 1S/2F presteren. In paragraaf 8.1 wordt antwoord gegeven op onderzoeksvraag 1. In paragraaf 8.2 wordt ingegaan op de verdeling van de vaardigheid. Vervolgens worden in paragraaf 8.3, meer specifiek in paragraaf 8.3.1, de prestaties van de leerlingen kwantitatief getypeerd. Gekeken wordt naar de beheersing van de opgaven in relatie tot verschillende vaardigheidsniveaus van leerlingen. Een kwalitatieve indruk van twee voorbeeldtaken wordt beschreven in paragraaf 8.4. Aan de hand van voorbeeldteksten geven we een kwalitatief beeld van zes verschillende schrijfvaardigheidsniveaus (p10 sbo, p10 bo, 1F, p50 bo, 2F en p90 bo). Paragraaf 8.5 beschrijft resultaten van analyses die gericht zijn op de algemene (holistische) indruk van de tekst en verzorgingsaspecten als spelling en grammatica, ook in samenhang met de vaardigheid van de leerlingen. Daarna wordt in paragraaf 8.6 een vergelijking gemaakt tussen de huidige en de vorige peiling, de trendvergelijking.

Voor de kwantitatieve en kwalitatieve analyses van de leerlingprestaties is in dit hoofdstuk gebruik gemaakt van een schrijfvaardigheidsschaal bestaande uit 176 opgaven.¹²¹ Na heroverweging zijn er ten opzichte van de in hoofdstuk 5 beschreven schaal, bestaande uit 186 items, uiteindelijk tien

¹²¹ Daarvan zijn er 110 dichotoom (twee score categorieën, maximum score 1). De overige 65 opgaven zijn polytoom, waarvan 21 samengestelde som-opgaven.

items uitgezet bij de uitvoering van de analyses.¹²² Controles op de uitkomsten aan de hand van de oude schaal leverden geen noemenswaardige verschillen op.

8.1 Behaalde percentages referentieniveaus 1F en 1S/2F

Voor het beantwoorden van de vragen over de percentages leerlingen die referentieniveau 1F en 2F beheersen en de onderliggende vaardigheidsverdeling, is eerst de vaardigheidsverdeling bepaald. Zoals in hoofdstuk 7 is beschreven, zijn de grenswaarden voor 1F en 2F tijdens een tweedaagse standaardbepaling door experts bepaald. Aangezien voor iedere leerling de vaardigheid geschat is, is ook te bepalen welk schrijfniveau, in termen van deze twee referentieniveaus, deze leerling het meest waarschijnlijk heeft. Dit is gebeurd door de vastgestelde grenswaarden (cut scores) toe te passen op de vaardigheidsschaal. Deze vaardigheidsschaal bestaat uit WML-schattingen op de onderliggende variabele, die gebaseerd zijn op de antwoorden van de leerling op alle items die hij/zij heeft beantwoord. De grenswaarden op die vaardigheidsschaal geven daarbij de WML-scores aan die net voldoende zijn om indicatief te zijn voor de betreffende referentieniveaus, oftewel de scores die de grens tussen wel of geen beheersing van het referentieniveau weergeven. In Tabel 8.1 wordt de verdeling van de referentieniveaus voor de uniek vastgestelde leerlingen weergegeven.¹²³ Bij de referentieniveaus is een driedeling gemaakt: de leerlingen onder 1F (<1F), de leerlingen op 1F maar niet op een hoger niveau (1F) en de leerlingen op minimaal 1S/2F. Cumulatieve percentages voor leerlingen met minstens niveau 1F kunnen worden berekend door 100% te verminderen met de percentages van <1F. Dat levert voor bo 73.4% van de leerlingen op en voor sbo 33.4%.

In totaal zijn er vaardigheidsscores en referentieniveaus voor 3319 leerlingen bepaald. Uit Tabel 8.1 kan worden afgeleid dat 67% van de leerlingen in het sbo onder het niveau 1F presteert, terwijl dit in het bo voor 27% van de leerlingen geldt. De overige bijna driekwart van de bo-leerlingen presteert op niveau 1F of 2F. De inschatting dat 75% van de leerlingen in het po het fundamentele niveau 1F wordt daarmee in het bo bijna gehaald, terwijl dit voor sbo-leerlingen niet het geval is. Het percentage leerlingen dat het streefniveau 1S/2F behaalt, namelijk 28% in het bo en 9% in het sbo, is beduidend lager dan de inschatting dat 50% van de po-leerlingen het 1S/2F niveau behaalt. De door de Commissie Meijerink in 2008 geformuleerde *ambitie* dat respectievelijk minimaal 85% en 65% van de po-leerlingen de niveaus 1F en 1S/2F zou halen wordt niet behaald.

¹²² De vervallen items betreffen zes combinatie-items en vier losse items. Van deze losse items was één item zeer moeilijk. Dit was opgave 4 in Taak 09 met een p-waarde van 0,02 (datum en tijd, inhoudelijk hetzelfde item als bij Taak 04, die reeds verwijderd was vanwege een dalende IRF - zie Tabel 5.11). Ook het losse item dat gericht was op het bevragen van de hoeveelheid details die de tekst volgens de beoordelaars aantrekkelijk maken, is verwijderd (T05 item 20). De overige twee losse items bevroegen lay out-technische aspecten (T06 item 9 en T07 item 38). De verwijderde combinatie-items (somscores) zijn met name gericht op een afsluitende zin en enkele aspecten van die afsluitende zin (combinatie-items gebaseerd op losse opgaven: T04 item 22--26, T09 item 20-25, T10 item 16-21 en T11 item 14-19), Daarnaast bevatte de oorspronkelijke schaal een combinatie-item gericht op zowel een inhouds-, opbouw- en twee formuleringsitems (T09 item 13-16). Dit item is verwijderd. Tenslotte is het combinatie-item gericht op stap 5 bij het pannenkoeken bakken verwijderd (T07 item 28-31).

¹²³ Deze verdeling wijkt enigszins af van de in hoofdstuk 7 gegeven percentages. Reden hiervoor is dat de gebruikte data voor het standaardbepalingsonderzoek tevens schrijfoverdrachten bevatte die niet eenduidig aan deelnemende leerlingen konden worden toegewezen. Tabel 8.1 bevat enkel de uniek geïdentificeerde, deelnemende leerlingen.

Tabel 8.1: Verdeling van de referentieniveaus in regulier (bo) en speciaal (sbo) basisonderwijs

WML Referentieniveau	bo		sbo	
	n	%	n	%
<1F	647	26.6	591	66.6
1F	1115	45.8	217	24.5
>1F (1S/2F)	670	27.5	79	8.9
Totaal	2432		887	

8.1.1 Referentieniveaus en achtergrondkenmerken van de leerlingen

De Tabellen 8.2 en 8.3 laten relaties zien tussen achtergrondkenmerken van leerlingen en het referentieniveau dat zij halen. Tabel 8.2 relateert de niveaus aan geslacht en thuistaal. De percentages in de drie categorieën lopen op tot 100. Wanneer we kijken naar het verschil tussen meisjes en jongens in het bo, dan laat Tabel 8.2 zien dat jongens vaker presteren op niveau <1F dan meisjes: 40% van de jongens haalt 1F niet (en 60% wel) tegenover 14% van de meisjes dat dit niveau niet haalt (en 86% wel). Ook is het percentage meisjes dat niveau 2F haalt hoger dan het percentage jongens; 41% van de meisjes haalt niveau 2F (en 59% niet), terwijl 15% van de jongens niveau 2F haalt (en 85% niet). Er blijkt een significant verschil tussen de percentages behaalde referentieniveaus door jongens en meisjes in het bo en het sbo (bo: $\chi^2(df=2)=275.41$, $p=.00$; sbo: $\chi^2(df=2)=24.19$, $p=.00$). Bij moedertaal/thuistaal worden drie groepen leerlingen onderscheiden: 1) leerlingen die in het Nederlands hebben leren praten, 2) leerlingen die in het Nederlands en een andere taal/dialect hebben leren praten en 3) leerlingen die in een andere taal hebben leren praten. Wanneer we kijken naar de verdeling van de behaalde percentages referentieniveaus over de groepen leerlingen, dan zien we dat deze significant verschillen in het bo ($\chi^2(df=4)=22.04$, $p=.00$), terwijl dit in het sbo niet het geval is ($\chi^2(df=4)=8.72$, $p=.07$). In het bo is het percentage leerlingen met niveau 2F dat in het Nederlands heeft leren praten hoger dan in de andere twee groepen en het percentage leerlingen onder niveau 1F lager.

Tabel 8.2: Behaald percentage referentieniveaus (1F en 2F) voor geslacht en thuistaal

	bo			sbo		
	Lager dan 1F	1F	2F	Lager dan 1F	1F	2F
Jongen	39.9%	45.6%	14.5%	72.4%	21.3%	6.3%
Meisje	14.2%	44.5%	41.3%	56.3%	30.8%	12.9%
Thuistaal is NL	25.1%	45.4%	29.5%	64.1%	26.7%	9.2%
Thuistaal is NL en andere taal	32.8%	47.2%	20.0%	71.2%	20.7%	8.2%
Thuistaal is niet NL	31.5%	48.1%	20.4%	83.8%	13.5%	2.7%

Daarnaast zijn de behaalde referentieniveaus gerelateerd aan het schooladvies van de leerlingen (Tabel 8.3). De tabel laat zien dat leerlingen die het advies praktijkonderwijs (PRO) krijgen in de regel onder 1F presteren (80% in het bo, 82% in het sbo). Van de bo-leerlingen met een advies vmbo-b/k presteert 44% onder 1F (en 56% op 1F, waarvan 13% op 2F). In het sbo presteert 58% van de leerlingen met vmbo b/k advies onder 1F en 42% op 1F, waarvan 12% zelfs op 2F. Zeventig procent van de bo-leerlingen met het advies vmbo-g/t presteert op niveau 1F of hoger, terwijl dit in het sbo 43% betreft. Meer dan 82% van de bo-leerlingen met havo/vwo-advies presteert op niveau 1F of hoger (waarvan 36% op niveau 2F). In het sbo is dit 78%, maar betreft dit vanzelfsprekend weinig leerlingen. De

percentages behaalde niveaus verschillen significant van elkaar over de verschillende schooladviezen, zowel in het bo als in het sbo (bo: $\chi^2(df=6)=214.03$, $p=.00$; sbo: $\chi^2(df=6)=70.64$, $p=.00$).

Tabel 8.3: Behaald percentage referentieniveaus (1F en 2F) naar schooladvies

	bo						sbo					
	Lager dan 1F		1F		2F		Lager dan 1F		1F		2F	
	n	%	n	%	n	%	n	%	n	%	n	%
PRO	20	80	4	16	1	4	258	82	46	15	10	3
Vmbo b/k	193	44	187	43	57	13	202	58	106	30	42	12
Vmbo g/t	172	30	288	50	112	20	33	57	17	29	8	14
Havo/vwo	247	18	615	45	493	36	5	23	12	55	2	23

8.1.2 Relatie tussen referentieniveaus, schrijfniveaus en OFI-aspecten

De behaalde referentieniveaus zijn gebaseerd op de schrijfvaardigheidsniveaus van de leerlingen. Tabel 8.4 toont de gemiddelde prestaties van leerlingen van de drie onderscheiden referentiegroepen, weer uitgesplitst naar bo en sbo. Separate Anova-analyses voor bo en sbo laten zien dat de vaardigheidsscores van de drie groepen leerlingen, conform verwachting, significant van elkaar verschillen (bo: $F(df=2)=4502.95$, $p=.00$; sbo ($F(df=2)=773.24$, $p=.00$): hoe hoger het behaalde referentieniveau, hoe hoger de vaardigheidsscore. Post hoc analyses bevestigen dit beeld.

Tabel 8.4: Gemiddelde vaardigheid voor de drie referentieniveaus, naar type school

		n	M	SD	Min	Max
bo	Lager dan 1F	647	-.24	.11	-1.04	-.13
	1F	1115	-.02	.06	-.13	.09
	2F	670	.19	.09	.09	.72
sbo	Lager dan 1F	591	-.33	.16	-1.10	-.13
	1F	217	-.04	.06	-.13	.09
	2F	79	.22	.13	.09	.78

Zoals in hoofdstuk 1 en 5 is beschreven, is bij de beoordeling van de leerlingwerken onderscheid gemaakt tussen Opbouw-, Formulerings- en Inhoudsaspecten, die gezamenlijk de communicatieve effectiviteit van de tekst representeren. De *Opbouw*-aspecten gaan in op elementen aan het begin of eind van de tekst, op relaties en overgangen en op volgorde – oftewel de plaatsing van vaste tekstelementen. Onder *Formulerings*aspecten vallen zaken als de formaliteit van de tekst(elementen), beleefdheidselementen en details ter verlevendiging of ter bevordering van de uitvoerbaarheid van de taak die in de tekst centraal staat (zoals smakelijke pannenkoeken bakken, voldoen aan een verzoek, etc.). Bij *Inhoudsaspecten* bevragen de modellen de aanwezigheid van elementen (het gebruik van informatie uit de taakomschrijving en de aanwezigheid van voor het genre verplichte elementen). Nagegaan is of de drie onderscheiden groepen leerlingen (lager dan 1F, op 1F en op 2F) ook van elkaar verschillen in hun prestaties op de drie componenten Opbouw, Formulering en Inhoud. In zowel bo als sbo verschillen de prestaties van de leerlingen op O, F en I significant van elkaar al naar gelang hun

behaalde referentieniveau. Dit laten Anova-analyses zien.¹²⁴ Post hoc analyses (Tukey) zijn uitgevoerd om na te gaan of dit onderscheid in gemiddelde prestatie voor alle drie de OFI-aspecten op alle drie de niveaus geldt. Dit is het geval: leerlingen met niveau onder 1F presteren op zowel O, F als I significant lager dan leerlingen op 1F en 2F, in bo en sbo. Voor beide schooltypen geldt ook dat leerlingen die op 1F presteren significant lager presteren op O, F en I dan leerlingen die op 2F presteren.

Tabel 8.5: Gemiddelde vaardigheid OFI-aspecten, uitgesplitst naar referentieniveaus, per type school

		Opbouw			Formulering		Inhoud	
		N	M	SD	M	SD	M	SD
bo	Lager dan 1F	647	-0.27	0.21	-0.26	0.24	-0.24	0.15
	1F	1115	-0.02	0.15	-0.02	0.18	-0.02	0.13
	2F	670	0.20	0.17	0.18	0.18	0.18	0.16
sbo	Lager dan 1F	584	-0.38	0.34	-0.31	0.31	-0.33	0.22
	1F	215	-0.04	0.22	-0.06	0.22	-0.01	0.21
	2F	76	0.26	0.30	0.18	0.28	0.20	0.21

8.2 Verdeling van de vaardigheid

Naast een beschrijving van de behaalde percentages referentieniveaus, ook in relatie tot achtergrondkenmerken, is er gekeken naar de verdeling van de vaardigheid, zoals weergegeven in Tabel 8.6. Conform het beeld bij de referentieniveaus is het ook hier evident dat de bo-leerlingen beter presteren dan de sbo-leerlingen. Het effect van het type onderwijs is ongeveer Cohen's $d=0,93$, hetgeen een groot effect is. Uit Tabel 8.6 blijkt bovendien dat de geobserveerde prestaties niet volledig normaal verdeeld zijn; scheefheid en kurtosis wijken significant af van nul.

Tabel 8.6: Verdelingseigenschappen vaardigheid WML per type basis onderwijs (bo en sbo)

WML	bo	sbo
Mediaan	-0.022	-0.207
Gemiddelde	-0.024	-0.206
Standaardafwijking	0.182	0.230
Scheefheid	-0.231	0.056
SE(scheefheid)	0.049	0.081
Gepiektheid	0.813	1.300
SE(gepiektheid)	0.097	0.163

8.3 Typering van de prestaties van leerlingen

De tweede set onderzoeksvragen is gericht op de typering van de prestaties van een leerling als het gaat om de communicatieve effectiviteit van de leerlingtekst. De prestatie van de leerling kan hier het

¹²⁴ Voor het bo geldt: Opbouw: $F(df=2)=1242.64$, $p=.00$; formulering: $F(df=2)=841.51$, $p=.00$, Inhoud: $F(df=2)=1378.08$, $p=.00$. Voor het sbo geldt: Opbouw: $F(df=2)=209.62$, $p=.00$; formulering: $F(df=2)=134.42$, $p=.00$, Inhoud: $F(df=2)=319.00$, $p=.00$.

best gedefinieerd worden door de vaardigheid die de leerling laat zien op de taken die de leerling heeft gemaakt op basis van de OFI-score. Al met al houdt dit in dat we de leerlingen en de opgaven afzetten tegen de vaardigheid, waarbij we dankzij de standaardbepaling ook de grenzen van de referentieniveaus kennen op deze vaardigheidsschaal. Hierbij is speciale aandacht gevraagd voor relatief laag-vaardige leerlingen met een vaardigheid rond het 10^e percentiel, voor leerlingen met een gemiddelde vaardigheid en voor leerlingen met een hoge vaardigheid rond het 90^e percentiel. Tevens dienen de prestaties behorend bij de grenzen onder/op 1F en 1S/2F te worden getypeerd. Bij de typering van de vaardigheid van de verschillende percentielen is er, net als in de rest van dit hoofdstuk, voor gekozen de bo- en de sbo-populatie als verschillende populaties te beschouwen.¹²⁵

Binnen de peiling is gebruik gemaakt van analytische beoordelingsmodellen. Deze modellen bestonden uit zo concreet mogelijk geformuleerde vragen waarmee de aanwezigheid van kenmerken op het gebied van inhoud, opbouw en formulering werd bevraagd. In dit hoofdstuk worden zij zowel met de term opgaven als items aangeduid. De opgaven op de schaal zijn allen van het soort OFI; ze zijn ingedeeld als opbouw-, formulerings- of inhoudsaspect. We kunnen daarom onderscheid maken tussen O-opgaven, F-opgaven en I-opgaven om de opgaven verder te typeren. In Tabel 8.7 zijn de aantallen opgaven/items gegeven, evenals de maximale score, zowel gewogen als ongewogen.

Tabel 8.7: Aantal items en maximumscores per schaal

	N items	Max Ruwe score	Max Gewogen score
OFI	176	273	849
O	60	94	256
F	33	41	151
I	83	138	442

Om gevoel voor de resultaten te krijgen zijn in Figuur 8.2 de verwachte scores van de O-, F- en I-items weergegeven als percentage van de maximale ruwe score, zowel bij de gehele schaal (OFI – de grijze lijn) als de subschalen O, F en I.¹²⁶ De figuur laat zien dat leerlingen met een lage vaardigheid (de x-as) een laag percentage van de maximumscore halen (ca 10% op alle vier schalen) en dat dit percentage relatief hoog ligt bij leerlingen met een hoge vaardigheid. Merk op dat het verwachte percentage goed ook bij de zeer hoog presterende leerlingen het maximum niet bereikt. Conform verwachting¹²⁷ lopen de lijnen van de schaal en de subschalen tamelijk parallel aan elkaar. Het percentage van de maximumscores is bij inhoud voor alle leerlingen hoger dan dat voor de andere subschalen. Ook laat de figuur zien dat voor leerlingen met de hoogste vaardigheid de beheersing van formulering hoger lijkt te liggen dan voor opbouw. Daarnaast bevat Figuur 8.2 informatie over de verdeling: de puntjescurves geven grafieken van de normale verdeling met gegeven gemiddelde en SD voor bo en sbo weer, terwijl de curves met de doorgetrokken lijn de geobserveerde verdeling laten zien. De afwijking tussen de verdelingen is redelijk klein, maar de figuur laat nogmaals zien dat de vaardigheid van bo-leerlingen hoger ligt dan die van de sbo-leerlingen. Wat betreft de verdelingen dient overigens te worden opgemerkt dat de verticale as in de figuur niet kan worden gebruikt om de gegeven verdelingen te interpreteren. Tenslotte zijn ook de referentieniveaus als verticale lijnen (gestreept) weergegeven. We

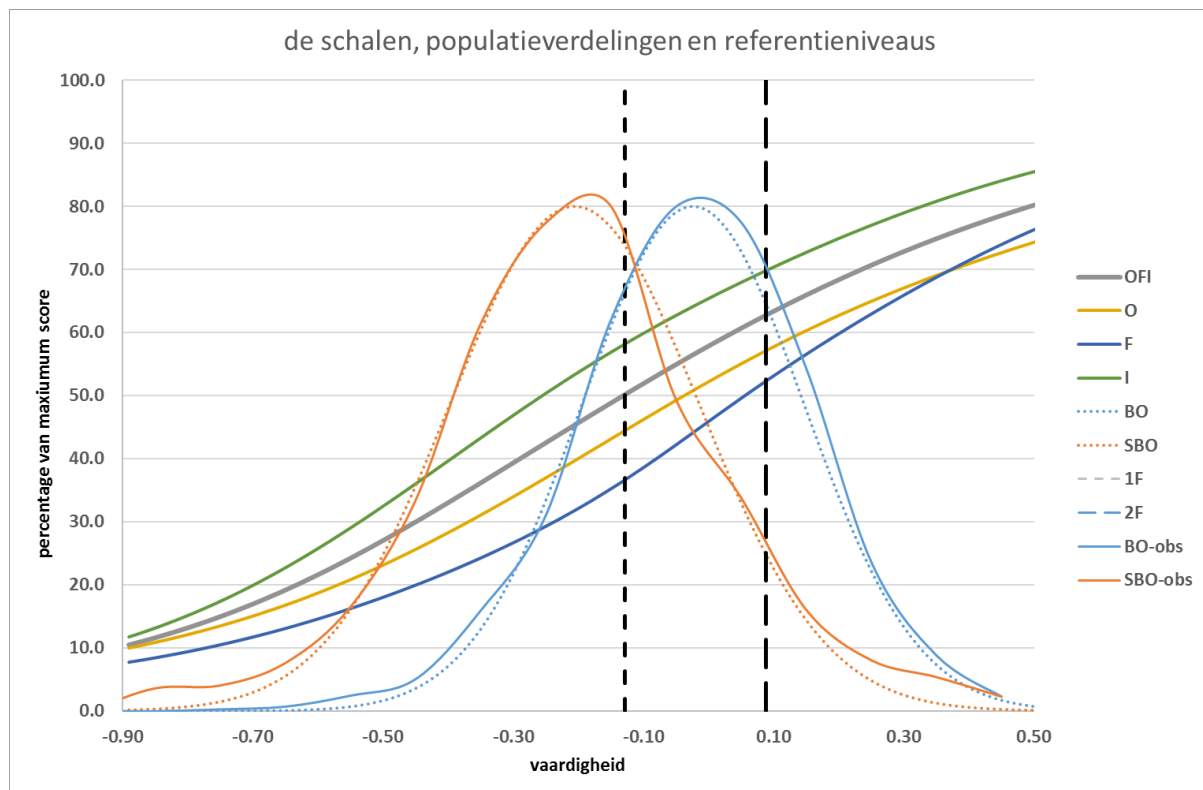
¹²⁵ Dit geldt niet voor de typering van de referentieniveaus 1F en 1S/2F, omdat deze absolute referentieniveaus voor beide typen onderwijs vastliggen.

¹²⁶ In dit geval zijn de Q en V items (samengestelde items) ook bij de O-items gerekend; dat scheelt 8 opgaven, 22 ongewogen en 54 gewogen scorepunten. Een enorm effect in de vorm van de curve is niet verwacht.

¹²⁷ Zie hoofdstuk 5 voor informatie over de schaling.

zien dat het vastgestelde niveau 1F hoger ligt dan de gemiddelde vaardigheid van de sbo-leerlingen en lager dan de gemiddelde vaardigheid van de bo-leerlingen, zoals al bleek uit paragraaf 8.1.

Figuur 8.2: de populatieverdelingen, referentieniveaus en verwacht percentage goed per schaal



We kunnen ook meer gedetailleerd op de kenmerken van de opgaven ingaan. Het aantal leerlingen exact op percentiel 10 of 90 is beperkt. Om die reden zijn de vaardigheidsgebieden rond percentiel 10 en percentiel 90 bekeken, en wel het vaardigheidsgebied tussen percentiel 5 en percentiel 15, en het gebied tussen percentiel 85 en 95. Voor de volledigheid is ook het vaardigheidsgebied rond de mediaan meegenomen (het gebied tussen percentiel 45 en 55). De ranges zijn gegeven in Tabel 8.8. Voor de referentieniveaus kunnen drie ranges gegeven worden, onafhankelijk van de populatie, namelijk onder 1F (alle waarden onder vaardigheidsniveau -0,128), op 1F (alle waarden vanaf vaardigheidsniveau -0,128 tot 0,088), en boven 1S/2F (alle waarden vanaf vaardigheidsniveau 0,088). Kijkend naar de specifieke onderzoeksvraag wordt tevens een overzicht gegeven van de vaardigheidsscores op p10, p50 en p90, alsmede van de grenswaarden behorend bij de referentieniveaus 1F en 2F (Tabel 8.9). Opvallend is dat er sprake is van een zekere overlap tussen de typen leerlingen in het bo en sbo. Een p10-leerling in het bo blijkt ongeveer een p50-leerling in het sbo en een p90-leerling in het sbo ligt tussen een p50- en p90-leerling in het bo in. Sterker nog, deze p90-leerling in het sbo presteert ongeveer op 2F-niveau. In het bo presteren de p28-leerlingen op 1F en de p73-leerlingen op 2F.¹²⁸

¹²⁸ Deze percentielen wijken enigszins af van de in Tabel 8.1 gepresenteerde percentages, zoals daar ook werd beschreven. De standaardbepaling is uitgevoerd op basis van alle aanwezige prestatie-informatie per gemaakte taak en de percentielen zijn daarop gebaseerd. De percentages in Tabel 8.1 zijn de percentages unieke leerlingen die de referentieniveaus halen. Verschillen zijn het gevolg van administratieve fouten.

Tabel 8.8: Typering van groepen leerlingen

			vaardigheidsrange			
			bo		sbo	
groep	percentielrange	typering	vanaf	t/m	vanaf	t/m
1	< p05	zeer laag	..	-0.330	..	-0.578
2	p05-p15	laag	-0.329	-0.205	-0.577	-0.414
3	p15-p45	onder gemiddeld	-0.204	-0.043	-0.413	-0.230
4	p45-p55	gemiddeld	-0.042	0.003	-0.229	-0.186
5	p55-p85	bovengemiddeld	0.004	0.153	-0.185	0.017
6	p85-p95	hoog	0.154	0.262	0.018	0.171
7	> p95	zeer hoog	0.263	..	0.172	..

Tabel 8.9: Vaardigheidsscores voor leerlingen op p10, p50, p90 en 1F en 2F

	P10	1F	P50	2F	P90
	WML-score	WML-score	WML-score	WML-score	WML-score
bo	-0.255	-0.128	-0.021	0,088	0.197
sbo	-0.478	-0.128	-0.206	0,088	0.089

8.3.1 Beheersing opgaven

Naast een beschrijving van de vaardigheid van verschillende groepen leerlingen is er ook gekeken naar de beheersing van de opgaven binnen het bo en sbo (zie Tabel 8.10). Deze beheersing is berekend aan de hand van de p-waarde van een opgave¹²⁹. Daarbij is een onderscheid gemaakt tussen drie typen opgaven:

- onvoldoende beheerste opgaven (<p50): de kans om de opgave goed te maken is kleiner dan 50%¹³⁰
- matig beheerste opgaven (p50-80): de kans om de opgave goed te maken ligt tussen 50% en 80%
- goed beheerste opgaven (>p80): de kans om de opgave goed te maken is groter dan 80%

Van alle opgaven bij elkaar werd in het bo 48% en in het sbo 60% niet beheerst door de leerlingen, ruim een kwart van de opgaven werd door beide groepen leerlingen matig beheerst (resp. 27% in het bo en 28% in het sbo) en 25% van de opgaven werd in het bo goed beheerst en 12.5% in het sbo.

¹²⁹ Bij de verdeling van de 12 versies over de scholen is rekening gehouden met de expliciete stratificatiecriteria binnen de steekproeftrekking. Voor het bo betekende dit dat de proportionele verdeling van scholen over de vier strata (gebaseerd op % gewichtenleerlingen op school) per versie zoveel mogelijk is aangehouden en dat er rekening is gehouden met school-/klasgrootte, zodat het aantal waarnemingen per versie vrijwel even groot zou zijn. Bij de verdeling van de versies over de sbo-scholen is met schoolgrootte (expliciet stratificatiecriterium) en met klasgrootte rekening gehouden. Op deze manier is per versie gestreefd naar een representatief beeld voor de populatie.

¹³⁰ Bij dichotome opgaven komt dit neer op de kans dat het betreffende beoordelingsaspect in het leerlingwerk aanwezig is. Bij onvoldoende beheerste opgaven is de kans kleiner dan 50% dat het aspect in de tekst aanwezig is, bij matig beheerste opgaven ligt deze kans tussen 50% en 80% en bij goed beheerste opgaven is de kans 80% of meer. Bij polytome opgaven gaat het om de verwachte waarde van 50% danwel 80% van de maximumscore.

Tabel 8.10: Beheerste opgaven (alle taken)

	bo		sbo	
	n	%	n	%
Niet beheerste opgaven	84	47.73 %	105	59.66 %
Matig beheerste opgaven	48	27.27 %	49	27.84 %
Beheerste opgaven	44	25.0 %	22	12.50 %
Totaal aantal opgaven	176		176	

Leerlingen kregen twaalf schrijftaken voorgelegd. Tabel 8.11 toont voor zowel bo als sbo de moeilijkheid van de taken door middel van de gemiddelde p-waarde over alle gemeten items van de taak. De taken zijn in Tabel 8.11 van boven naar onder in oplopende moeilijkheid weergegeven. Een hogere gemiddelde p-waarde betekent dat de taak (relatief) makkelijk is. De schrijftaken 06 (Vlinder - uitleg metamorfose), 07 (Pannenkoekenrecept) en 02 (Verboden boek – verhaal) zijn de drie makkelijkste taken, terwijl de taken 11 (Zomervakantie – verzoek aan de directeur), 05 (Nepnieuws - informatieve tekst) en 08 (Poekie - instructie m.b.t. verloren kat) het moeilijkst zijn.

Tabel 8.11: Gemiddelde p-waarde per schrijftaak, uitgesplitst naar bo en sbo

bo			sbo	
Taak	Gemiddelde p-waarde		Taak	Gemiddelde p-waarde
T06	0.67		T06	0.56
T07	0.63		T02	0.51
T02	0.61		T07	0.50
T12	0.55		T12	0.47
T01	0.54		T01	0.44
T09	0.54		T09	0.42
T04	0.54		T04	0.42
T03	0.49		T10	0.40
T10	0.48		T03	0.38
T11	0.46		T11	0.37
T05	0.44		T08	0.35
T08	0.44		T05	0.34

Zoals eerder gezegd, is de beoordeling van de leerlingwerken uitgevoerd aan de hand van opbouw-, formulerings- en inhoudsaspecten. Zoals in hoofdstuk 5 is beschreven, bleken bij de schaling enkele opgaven niet eenduidig aan één van deze drie aspecten toe te kennen, maar aan een combinatie. Het gaat daarbij om combinaties van O en F en van O, F en I (resp. V en Q¹³¹). In Tabel 8.12 staan de gemiddelde p-waarden van de opgaven die vallen onder de gehanteerde OFI-aspecten weergegeven, opgesplitst naar type onderwijs. Zowel in het bo als in het sbo blijken de inhoudsopgaven gemiddeld de hoogste p-waarde te hebben en daarmee over het algemeen het gemakkelijkst te zijn. De formulerings- en opbouwaspecten zijn moeilijker, en hebben gemiddeld een vergelijkbare moeilijkheid in beide schooltypen. De hoogte van de mediaan ten opzichte van de gemiddelde p-waarde is bij de

¹³¹ Voor meer informatie over de V en Q item types zie paragraaf 5.5 en Tabel 5.10.

O- en F-items lager, terwijl deze hoger is dan de gemiddelde p-waarde bij de I-items. Voor zowel O, F als I geldt dat er zowel heel moeilijke als heel makkelijke opgaven waren (resp. heel lage als heel hoge p-waarden). Een belangrijke opmerking is wel dat deze uitkomsten volledig afhankelijk zijn van de gehanteerde indeling in O, F en I. Uitkomsten dienen met enige voorzichtigheid te worden geïnterpreteerd, omdat er, in ieder geval voor een aantal items, discussie over de classificatie mogelijk is. Zo zijn items die de ‘inhoud van de inleiding’ betreffen geclassificeerd als inhoudsitems, maar er kan ook beargumenteerd worden dat ze als opbouw-items kunnen worden geïnterpreteerd vanuit de gedachtegang dat een goede inleiding een kwestie van opbouw is.

Tabel 8.12 laat tevens de p-waarden van opgaven met verschillende maximumscores zien. Op de meeste opgaven (152) kon de leerling 1 punt halen, maar er waren ook 21 opgaven waarop 2-4 punten konden worden gescoord. De gemiddelde en mediane p-waarden voor opgaven met 1 punt waren overal lager dan bij opgaven waarmee meerdere punten konden worden gescoord.

Tabel 8.12: p-waarden van de OFI-opgaven en type opgaven, opgesplitst naar type onderwijs

Type Item	Maximum-score	# items n	P-waarde in bo				P-waarde in sbo			
			Mdn	M	Min	Max	Mdn	M	Min	Max
F		33	0.41	0.45	0.02	0.97	0.27	0.34	0.01	0.88
I		83	0.68	0.62	0.07	0.97	0.55	0.51	0.05	0.91
O		49	0.37	0.42	0.02	0.97	0.26	0.33	0.01	0.93
Q		1	0.70	0.70	0.70	0.70	0.59	0.59	0.59	0.59
V		7	0.76	0.74	0.53	0.91	0.62	0.58	0.31	0.75
	1	152	0.51	0.52	0.02	0.97	0.37	0.43	0.01	0.93
	2 ¹³²	12	0.58	0.58	0.21	0.91	0.42	0.44	0.17	0.75
	3	6	0.74	0.64	0.22	0.87	0.57	0.50	0.16	0.75
	4	3	0.67	0.67	0.63	0.70	0.50	0.52	0.48	0.59
		173	0.53	0.53	0.02	0.97	0.40	0.43	0.01	0.93

8.3.2 Beheersing opgaven door verschillende typen leerlingen

In Tabel 8.8 zijn de berekende ranges in vaardigheidsscores weergegeven. De vraag was om voor laag- en hoogvaardige leerlingen en de gemiddelde presteerders alsmede voor leerlingen op 1F en 2F niveau de prestaties te typeren. Nu de bijbehorende ranges bekend zijn, kan voor iedere opgave aangegeven worden welke kans (50% dan wel 80%) de zeven in Tabel 8.8 gespecificeerde groepen leerlingen hebben om de opgave goed te maken. Voor de dichotome items betekent dit een kans van 50% of 80% op aanwezigheid van het betreffende beoordelingsaspect in de leerlingtekst. Bij de polytome opgaven gaat het om de verwachte waarde van 50% dan wel 80% van de maximumscore.

In Tabel 8.8 zijn de zeven percentielgroepen leerlingen onderscheiden. Op basis van de itemlocaties, dat wil zeggen de plek op de vaardigheidsschaal waar de verwachte waarde op dat item 50% of 80% is van de maximumscore - het AI-50 punt en het AI-80 punt - is nagegaan hoeveel items

¹³² Polytome opgaven zijn opgaven waarbij er meer dan twee onderscheidingen zijn in de waardering van het antwoord. De p-waarde wordt als volgt berekend: gemiddelde score op de opgave/maximaal haalbare score op die opgave.

deze percentielgroepen matig/goed beheersen (zie Tabel 8.13). Per percentielgroep (1 tot 7) is het aantal items weergegeven waarvoor deze groep tussen de 50 en 80% (linker kolom) of meer dan 80% (rechterkolom) kans heeft deze goed te maken, steeds afgezet tegen de percentielgroep eronder. Kijkend naar Tabel 8.13 kan bijvoorbeeld worden geconcludeerd dat de minst vaardige leerlingen (< p05) meerdere items matig of goed beheersen (respectievelijk 56 en 14 in bo en 27 en 0 in sbo). Percentielgroep 2 in het bo (tussen p5 en p15) heeft bij 16 extra items een kans van 50% deze goed te maken (en dus deze items matig te beheersen), terwijl de bo-leerlingen in percentielgroep 1 (< p05) deze 16 items niet met 50% kans goed maken (en dus onvoldoende beheersen). Datzelfde geldt voor de vergelijking tussen leerlingen in percentielgroep 1 en 2 in het sbo: sbo-p5-15 beheerst 15 extra items matig (ten opzichte van sbo < p05). Overigens beheerst, per onderwijstype, percentielgroep 2 de items die door percentielgroep 1 matig beheerst worden, minimaal ook matig.¹³³ Op basis van de onderste rij van Tabel 8.13 kan worden afgeleid dat 50 items pas in het bereik komen van de bo->p95 leerlingen uitgaande van het AI-50 punt van de betreffende items, en dit geldt voor 96 items uitgaande van het AI-80 punt. In het sbo ligt het aantal items dat pas binnen het bereik komt van de meest hoogvaardige leerlingen zoals verwacht hoger: dit geldt in het sbo voor 60 items uitgaande van het AI-50 punt van de betreffende items en voor 104 items uitgaande van het AI-80 punt. Er zijn dus relatief veel items (resp. 50 en 60 items) waarvoor geldt dat er alleen in de teksten van allerhoogste presteerders (> p95 leerlingen in resp. bo en sbo) 50% kans is dat die beoordelingsaspecten in de teksten aanwezig zijn. De bo-leerlingen beheersen dus minder items op onvoldoende niveau dan sbo-leerlingen. Gezien de gemiddeld hogere vaardigheid van bo-leerlingen ten opzichte van sbo-leerlingen is deze uitkomst conform de verwachting. Bijlage 11 bevat een typering van alle opgaven op basis van itemlocatie, naar referentieniveau en type leerling (p10, p50, p90, uitgesplitst naar bo-sbo). Uit Bijlage 11 kan worden afgeleid dat het beeld enigszins grillig is: hoewel opbouw- en formuleringsitems over het algemeen moeilijker zijn (zoals we eerder zagen), is het niet zo dat er in de hogere percentielen alleen nog maar opbouw- en formuleringsitems bij komen. Ook bepaalde inhoudsitems komen soms pas op hogere niveaus in het bereik.

¹³³ Per definitie wordt een item dat matig door percentielgroep 1 wordt beheerst ook door percentielgroep 2 of hoger minimaal matig beheerst en wordt een item dat goed door percentielgroep 1 wordt beheerst ook door percentielgroep 2 of hoger goed beheerst. Echter, opgaven die door een bepaalde percentielgroep matig worden beheerst, worden niet per definitie door de percentielgroep die één niveau hoger presteert (in het voorbeeld percentielgroep 2) goed beheerst. Het kan zo zijn dat een goede beheersing van een matig beheerste opgave door percentielgroep 1 pas in het bereik ligt van bijvoorbeeld leerlingen in percentielgroep 4 en hoger, zoals we in de gegeven voorbeelden ook zullen zien.

Tabel 8.13: aantal items dat in het bereik komt van de onderscheiden percentielgroepen, op het AI-50 punt en AI-80 punt

		Aantal Items			
		bo		sbo	
		AI-50	AI-80	AI-50	AI-80
1	< p05	56	14	27	0
2	p05-15	16	14	15	5
3	p15-p45	17	21	27	21
4	p45-p55	6	8	6	4
5	p55-p85	17	13	20	27
6	p85-p95	14	10	21	15
7	> p 95 *	50	96	60	104

* Deze items komen alleen in het bereik van de buitengewoon schrijfvaardige leerlingen (> p 95) en worden dus eigenlijk niet/amper beheerst. Voor sbo-leerlingen liggen meer items buiten het bereik dan bij bo-leerlingen.

Zoals uit Tabel 8.13 kan worden afgeleid, bevatten de beoordelingsmodellen ook opgaven die zelfs door de minst vaardige bo-leerlingen (percentielgroep 1) matig of goed worden beheerst. Opgemerkt moet worden dat sbo-leerlingen in de laagste percentielgroep (de <p5 sbo-leerlingen) geen enkele opgave goed beheersen.

Om meer gevoel te krijgen voor de aard van de opgaven in relatie tot de beheersing worden hieronder per percentielgroep tabellen (Tabel 8.14 – 8.21) gepresenteerd met daarin enkele voorbeeldopgaven. Zo beogen wij een beeld te geven van de schaalbaarheid van de opgaven. We proberen te laten zien dat veel soorten items pas op een bepaald punt op de vaardigheidsschaal worden beheerst en dat een hogere vaardigheid vooral bestaat uit het beheersen van steeds meer soorten items. Bij de kwalitatieve typering in paragraaf 8.4 zal daarnaast aan de hand van schrijfproducten op twee taken worden geïllustreerd hoe de verschillende schrijfniveaus kunnen worden getypeerd. De navolgende tabellen bevatten allemaal dezelfde informatie: de taak, het itemnummer, gevolgd door een korte inhoudelijke beschrijving van het item, alsmede twee itemlocaties AI-50 en AI-80. Deze itemlocaties geven de plaats op de vaardigheidsschaal weer, waar de verwachte waarde op dat item 50% respectievelijk 80% van de maximumscore is.¹³⁴ Omdat deze itemlocaties zijn gebaseerd op de schrijfvaardigheidsschaal, zijn deze itemlocaties voor beide typen onderwijs (bo en sbo) dezelfde. Het type leerling (in termen van percentielgroepen) dat deze kansen heeft, verschilt echter wel voor het type onderwijs.

¹³⁴ Bij dichotome items komt dit neer op een kans van 50% danwel 80% dat het item in de leerlingteksten voorkomt.

Duiding kleuraanduidingen Tabellen 8.14-8.20

In de tabellen is per item weergegeven bij welke itemlocatie het item in 50% van de teksten voorkomt* (matige beheersing) en bij welke itemlocatie het item in 80% van de teksten voorkomt* (goede beheersing). Om de interpretatie van de tabellen 8.14-8.20 te vergemakkelijken is *met kleuren aangegeven in welke mate de items in het bereik liggen van de beste leerlingen binnen de betreffende percentielgroep* (de bo- en sbo-grensléerlingen met de hoogste vaardigheid in die percentielgroep). Per locatie is weergegeven of het betreffende item in het bereik ligt van deze bo-leerlingen (groene arcering) danwel in het bereik van deze sbo-leerlingen (gele arcering). Geen arcering betekent dat het item niet in het bereik van bo- en/of sbo-leerlingen ligt.

* Voor alle items, dus ook voor de polytome items, betreft het de verwachte waarden van 50% respectievelijk 80% van de maximumscore op dat item.

Percentielgroep 1 (bo < p05)

Tabel 8.14 toont enkele van de eenvoudigste voorbeelditems. Zo wordt het eerste item (Taak 02, item 11, het chronologisch opbouwen van de kern van een verhaal) matig beheerst door leerlingen met een vaardigheidsscore vanaf -0.82 en goed beheerst door leerlingen met een vaardigheidsscore vanaf -0.54. De eerste drie opgaven in Tabel 8.15 worden goed beheerst door de < p05 leerlingen in het bo en de overige items matig. Sbo- < p05 leerlingen beheersen geen enkel item goed en, op het laatste item na, alle items matig (zie Bijlage 11); het laatste item beheersen deze sbo-leerlingen onvoldoende. Op basis van Bijlage 11 kunnen we ook nagaan door welke percentielgroepen de items goed beheerst worden. Het presenteren van informatie in een logische volgorde bij Taak 03 (samenvatting) wordt matig beheerst door bo-< p05-leerlingen in zowel bo als sbo, maar wordt goed beheerst door de bo-p5-15 percentielgroep en door sbo-p15-45 percentielgroep. Het formuleringsitem schrijven van een zakelijke tekst bij Taak 05 wordt matig beheerst door percentielgroep bo- < p05, maar de percentielgroepen bo-p15-45 en sbo-p45-55 beheersen deze opgave goed. Het laatste item in de tabel wordt matig beheerst door bo-< p05-leerlingen en sbo-p5-15 leerlingen, goede beheersing van deze opgave (passende toon) vindt pas plaats bij de bo-p45-55 percentielgroep en de sbo-p55-85 percentielgroep.

Tabel 8.14: Matig en goed beheerste opgaven door bo-< p05 leerlingen – percentielgroep 1 bo (bo: - -0.330 en sbo: - -0.578)¹³⁵

Taak	Item-nummer	Aspect	Korte inhoud van het item	50% ^a		80%	
				bo	sbo	bo	sbo
2	11	Opbouw	Kern van het verhaal is chronologisch opgebouwd	- 0.82		- 0.54	
8	14	Inhoud	Vermelding van telefoonnummer waaronder leerling te bereiken is	- 0.67		- 0.44	
1	5	Formulering	Het verhaal is onderhoudend	- 0.60		- 0.41	
4	13	Inhoud	Leerling beantwoordt de vraag wat Nederlanders graag eten	- 0.64		- 0.32	
3	17	Opbouw	Leerling presenteert de informatie in een logische volgorde	- 0.61		- 0.26	
5	19	Formulering	De tekst is zakelijk geschreven	- 0.90		- 0.20	
12	11	Formulering	Toon is passend voor een formele geschreven tekst	- 0.48		- 0.02	

^a: De groene arcering markeert de beheersing door de grensleerling bo en de gele arcering door de grensleerling sbo. Geen arcering geeft onvoldoende beheersing weer.

Percentielgroep 2 (bo p5-15))

Tabel 8.15 laat enkele voorbeeldopgaven zien, die goed en vooral matig beheerst worden door bo-p5-15, de percentielgroep die indicatief is voor p10-leerlingen. Het eerste item (Taak 4, item 13) is ook in Tabel 8.14 opgenomen. Waar dat item door de bo- < p05-leerlingen matig wordt beheerst, wordt het door (een groot deel van de) bo-p5-15 leerlingen goed beheerst, namelijk door de leerlingen met een vaardigheidsscore - .32 of hoger. Het tweede item (taak 6, item 14) wordt matig beheerst door leerlingen met een vaardigheidsscore vanaf -0.28 (percentielgroep bo-p5-15). Ook de daaropvolgende opgaven worden matig beheerst door de leerlingen in percentielgroep bo-p5-15. De leerlingen in percentielgroep sbo-p5-15 (met een vaardigheidsscore lopend van -0.577 t/m -0.414) beheersen alleen de eerste opgave (Taak 04, item 13) matig. De rest van de getoonde items beheersen deze sbo-leerlingen onvoldoende.

¹³⁵ Voor de Tabellen 8.14 – 8.20 geldt de volgende richtlijn: de labels in Bijlage 11 (b.v. T02i11_O) corresponderen met de gebruikte nummering van de items in de beoordelingsformulieren. Deze itemnummers zijn onder Taak en Itemnummer in de Tabellen 8.14-8.20 weergegeven. De in Bijlage 11 gehanteerde codes onder 'It|TK' betreffen de in de OPLM-analyses gehanteerde codes die niet direct corresponderen met de nummering in de beoordelingsformulieren. Zo is bij Taak 02, opgave 11 uit het beoordelingsformulier (gericht op de chronologische opbouw van de kern van het verhaal) gecodeerd als item 9 van Taak 02 in OPLM-nr 32.

Tabel 8.15: Matig en goed beheerste opgaven door bo-p5-15 leerlingen – percentielgroep 2 bo
(bo: -0.329 / -.205, sbo: -0.577 / -0.414)

Taak	Item-nummer	Aspect	Korte inhoud van het item	50% ^a		80%	
				bo	sbo	bo	sbo
04	13	Inhoud	Leerling beantwoordt de vraag wat Nederlanders graag eten	- 0.64		- 0.32	
06	14	Inhoud	De tekst beschrijft dat het stadium van eitje 2-6 dagen duurt	- 0.28		- 0.11	
05	17	Opbouw	Er worden causale relaties aangegeven met verbindingswoorden (<i>omdat, want, doordat, daarom, dus, etc.</i>)	- 0.25		0.21	
09	18	Formulering	Biedt de briefschrijver Naima duidelijk ruimte om nee te zeggen?	- 0.25		0.21	
03	5	Inhoud	Vermeldt de leerling het werkgebied van Unicef?	- 0.30		0.17	
02	13	Opbouw	Maakt de schrijver je nieuwsgierig over wat komen gaat?	- 0.22		0.12	

^a: De groene arcering markeert de beheersing door de grensleerling bo en de gele arcering door de grensleerling sbo. Geen arcering geeft onvoldoende beheersing weer.

Percentielgroep 3 (bo p15-45) – waaronder 1F

Tabel 8.16 toont vooral opgaven die matig worden beheerst door bo-leerlingen in percentielgroep 3 (p15-45, lopend van -0.204 / -0.043). De eerste opgave, item 8 in Taak 3 (het vermelden van schoon water [=inhoudsaspect] in een samenvatting) wordt matig beheerst door leerlingen met vaardigheidsscore vanaf -0.28 en goed beheerst door leerlingen met vaardigheidsscore vanaf -0.05. Dit betekent dat alle leerlingen in percentielgroep bo-p15-45 de opgave matig beheersen en alleen de allerbeste leerlingen uit deze groep de opgave goed beheersen (namelijk die met een score van minimaal -0.05).¹³⁶ De tweede opgave (taak 6, item 14) – die ook in Tabel 8.15 voorkwam - wordt goed beheerst door leerlingen met een vaardigheidsscore vanaf -0.11 (behorend tot de hier besproken percentielgroep 3). De overige opgaven in Tabel 8.16 worden matig beheerst door percentielgroep bo-p15-45, terwijl een goede beheersing voor deze groep leerlingen buiten het bereik ligt. In het sbo wordt de eerste opgave in de tabel (Taak 3, item 8) matig beheerst door een deel van de sbo-p15-45 leerlingen (die met een score van -0.28 t/m -0.23), maar de overige opgaven worden door deze groep onvoldoende beheerst.

Wanneer we kijken naar de grenswaarde voor 1F (grenswaarde=-0.128), dan zien we dat de bo-leerlingen die net op 1F functioneren (de p28-leerling in het bo) twee van de inhoudsitems (Taak 3, item 8 en Taak 6, item 14) en de twee formuleringitems in de tabel matig beheersen (Taak 11, item 22, duidelijk maken dat een brief namens de klas is geschreven; Taak 2, item 8, de schrijver laten

¹³⁶ Merk op dat de opgave ook matig wordt beheerst door een deel van de bo-p5-15 leerlingen (percentielgroep 2), namelijk die met een vaardigheidsscore van -.28 t/m -.205.

meeleven in een verhaal). Het opbouwitem (item 6 in Taak 7, toevoeging van een kop of inleiding voor het deel aangaande de ingrediënten) wordt nét niet matig door 1F leerlingen beheerst en ook het tweede inhoudsitem (T06, item 12, informatie over het ontpoppingsstadium) is te moeilijk voor 1F-leerlingen.

Tabel 8.16: Matig beheerste opgaven door bo-p15-45 leerlingen – percentielgroep 3 bo
(bo: -0.204 / -0.043, sbo: -0.413 / -0.230)

Taak	Item-nummer	Aspect	Korte inhoud van het item	50% ^a		80%	
				bo	sbo	bo	sbo
03	8	Inhoud	Vermeldt de leerling schoon water?	-0.28		-0.05	
06	14	Inhoud	De tekst beschrijft dat het stadium van eitje 2-6 dagen duurt	-0.28		-0.11	
07	6	Opbouw	De tekst bevat een kop of inleidende zin voor het deel over ingrediënten	-0.12		0.12	
11	22	Formulering	Het is duidelijk dat de brief is geschreven namens de hele klas	-0.17		0.11	
06	12	Inhoud	Er is informatie in de tekst over het zevende stadium 'ontpoppen'	-0.08		0.33	
02	8	Formulering	De schrijver laat je meeleven door intensiverders, uitroepen of vragen	-0.16		0.12	

^a: De groene arcering markeert de beheersing door de grensleerling bo en de gele arcering door de grensleerling sbo. Geen arcering geeft onvoldoende beheersing weer.

Percentielgroep 4 (bo p45-55) – gemiddelde bo-leerling

Tabel 8.17 laat drie items zien. Het eerste item, Taak 03 item 21 (de tekst maakt doel en werk van Unicef voldoende duidelijk in een samenvatting), wordt matig beheerst door leerlingen met een vaardigheidsscore vanaf -0.14 en goed beheerst door leerlingen met een vaardigheidsscore vanaf 0.00. Dit betekent dat dit opbouwitem door leerlingen in percentielgroep 4 in het bo (bo-p45-55 – indicatief voor p50, de gemiddeld presterende bo-leerling) matig tot goed wordt beheerst: bo-p45-55 die tussen -0.042 en -0.0001 scoren beheersen de opgave matig en de enkele bo-leerlingen met een vaardigheidsscore van 0-0.003 beheersen de opgave goed. De overige twee opgaven in Tabel 8.17 wordt door de hele percentielgroep bo-p45-55 matig beheerst. Overigens zou er bij formuleringssitem 20 uit Taak 11 (groetformule) een invloed kunnen zijn van instructie. Leerlingen waren geïnstrueerd hun eigen naam niet te gebruiken, omdat de leerlingwerken geen tot de leerling herleidbare informatie mochten bevatten. Hierdoor hebben sommige leerlingen mogelijk geheel afgezien van een groetformule.

Tabel 8.17: Matig tot goed beheerste opgaven door bo-p45-55 – percentielgroep 4 bo
(bo: -0.042 / 0.003, sbo: -0.229 / -0.186)

Taak	Item-nummer	Aspect	Korte inhoud van het item	50% ^a		80%	
				bo	sbo	bo	sbo
03	21	Opbouw	Maakt de tekst het doel en het werk van Unicef voldoende duidelijk?	- 0.14		0.00	
07	4	opbouw	De tekst bevat een titel over het doel van de instructie (pannenkoeken bakken)?	- 0.03		0.20	
11	20	Formulering	De brief eindigt met een groetformule	- 0.04		0.35	

^a: De groene arcering markeert de beheersing door de grensleerling bo en de gele arcering door de grensleerling sbo. Geen arcering geeft onvoldoende beheersing weer.

Percentielgroep 5 (bo p55-85) – waaronder 1S/2F

In Tabel 8.18 staan voorbeelden weergegeven van opgaven die door bo-p55-85 leerlingen (bo percentielgroep 5) matig worden beheerst. Opgave 6 uit Taak 07, een kopje of inleidend gedeelte voor het deel over de ingrediënten in een recept, hebben we eerder gezien. Dit item wordt matig beheerst door alle leerlingen in deze percentielgroep (bo-p55-85), maar goed beheerst door een deel van hen (namelijk de leerlingen met een vaardigheidsscore tussen 0.12 en 0.153). De overige getoonde opgaven worden door de hele percentielgroep bo-p55-85 matig beheerst. Opgave 6 in Taak 7 wordt door de sbo-leerlingen in deze percentielgroep (-0.185 / 0.017) matig beheerst, de overige opgaven onvoldoende.

In bo-percentielgroep 6 vallen ook de leerlingen die net op niveau 1S/2F presteren, namelijk de p73-leerlingen in het bo). Uitgaand van de grenswaarde 0.088 blijken de inhoudsitems 13 uit Taak 5 (het beschrijven dat jongeren moeten leren nepnieuws te herkennen in een artikel) en 12 in Taak 9 (uitleggen waarom de schrijver de geadresseerde benadert in een brief) door 2F-leerlingen matig te worden beheerst. Dit geldt ook voor opgave 6 in Taak 7. Het karakteristieke taalgebruik in een verhaal (T02, formuleringssitem 10) blijken zij *nét* matig te beheersen, maar een min of meer conventionele slotformule of afscheidsgroet in het bericht over de verloren kat (T08, item 17) niet meer. Ook het toevoegen van motiverende elementen of aanmoedigingen aan het pannenkoekenrecept beheersen zij onvoldoende (T07, item 43).

Tabel 8.18: Matig beheerste opgaven door bo-p55-85 leerlingen – percentielgroep 5 bo
(bo 0.004 / 0.153, in sbo -0.185 / 0.017)

Taak	Item-nummer	Aspect	Korte inhoud van het item	50% ^a		80%	
				bo	sbo	bo	sbo
07	6	Opbouw	Bevat de tekst een kop of inleidende zin voor het deel over ingrediënten?	- 0.12		0.12	
05	13	Inhoud	De tekst beschrijft dat jongeren moeten leren hoe ze nepnieuws kunnen herkennen	0.03		0.38	
08	17	Opbouw	Onder het briefje staat een min of meer conventionele slotformule en/of afscheidsgroet	0.11		0.45	
02	10	Formulering	De dialogen bevatten karakteristiek taalgebruik van personages of expressieve keuzes in spelling (<i>m-m-maar</i>) of interpunctie (...)	0.08		0.43	
07	43	Formulering	Er staan motiverende elementen of aanmoedigingen in de tekst	0.13		0.60	
09	12	Inhoud	De briefschrijver legt uit waarom hij/zij juist Naima benadert	0.04		0.39	

^a: De groene arcering markeert de beheersing door de grensleerling bo en de gele arcering door de grensleerling sbo. Geen arcering geeft onvoldoende beheersing weer.

Percentielgroep 6 (bo p85-95)

In Tabel 8.19 staan enkele voorbeelden van opgaven die matig worden beheerst door bo-leerlingen in percentielgroep 6 (bo-p85-95 – indicatief voor hoogwaardige leerlingen, p90). Formuleringssitem 23 uit Taak 08, het vergroten van de betrokkenheid van de lezer, wordt bijvoorbeeld matig beheerst door de leerlingen met een vaardigheidsscore vanaf 0.16, terwijl een goede beheersing in het bereik ligt van leerlingen die minimaal een vaardigheidsscore van 0.51 hebben. Deze formuleringsopgave wordt ook door percentielgroep sbo-p85-95 matig beheerst, maar dat geldt niet voor de andere twee items in de tabel. Merk op dat deze moeilijke opgaven ook weer een inhoudsitem bevatten (het vragen aan de directeur om zijn besluit te herzien of van mening te veranderen).

Tabel 8.19: Matig beheerste opgaven door bo-p85-95 leerlingen – percentielgroep 6 bo
(bo: 0.154 / 0.262, sbo: 0.018 / 0.171)

Taak	Item-nummer	Aspect	Korte inhoud van het item	50%		80%	
				bo	sbo	bo	sbo
11	12	Inhoud	De schrijver vraagt aan de directeur om zijn besluit te herzien of om van mening te veranderen	0.25		0.71	
12	10	Opbouw	De tekst bevat een of meer afsluitende zinnen	0.24		2.08	
08	23	Formulering	De leerling probeert de betrokkenheid van de lezer te vergroten	0.16		0.51	

Percentielgroep 7 (bo > p95)

De zevende percentielgroep bestaat uit de allerbeste schrijvers. De in Tabel 8.20 gepresenteerde opgaven zijn voorbeelden van opgaven die uitdagend zijn voor deze bo-leerlingen met een bijzonder hoge vaardigheid (> p95). Dit is een zeer kleine groep, waarmee gesteld kan worden dat de opgaven door bijna alle leerlingen onvoldoende beheerst worden. De in Tabel 8.20 genoemde items vragen dan ook gevorderde vaardigheden: retorische handigheid (eerste en tweede item in de tabel), aandacht voor grafische conventies (derde item in de tabel) en beleefdheidsaspecten (items 5-7 in de tabel). Dergelijke gevorderde vaardigheden zijn niet aan de orde bij item 5 in Taak 7 (het vierde item in de tabel), maar blijktbaar zijn leerlingen niet geneigd bij een recept het doel aan het begin van de tekst te benoemen: pas bij een vaardigheidsscore van 0.71 komt dit doel in de inleidende zin in de helft van de teksten voor. Wellicht vinden leerlingen dit doel te voor de hand liggend om te benoemen of krijgen zij de voordelen van een doelgerichte inleidende zin in een recept niet geïnstrueerd.

Tabel 8.20: Matig beheerste opgaven voor p>95 leerlingen – percentielgroep 7 bo
(bo: 0.263 - ..., sbo: 0.172 - ...)^a

Taak	Item-nummer	Aspect	Korte inhoud van het item	50%	80%
1	15	Formulering	De leerling probeert de beleving van de lezer te vergroten door hem bij het verhaal te betrekken	0.29	0.63
3	25	Formulering	Gebruikt de leerling stijffiguren om zijn tekst te verlevendigen ¹³⁷	0.81	1.51
5	18	Opbouw	Er zijn in de bodytekst (soms) alineascheidingen op relevante punten (bijvoorbeeld na inleiding, tussen deelthema's, voor slot	0.77	2.16
7	5	Opbouw	Bevat de tekst een inleidende zin over het doel van de instructie (pannenkoeken bakken)?	0.71	1.40
9	19	Formulering	Toont de briefschrijver bij voorbaat dankbaarheid	0.46	0.92
10	13	Inhoud	Stelt de briefschrijver de gevraagde concessie voor als minimaal	0.41	0.75
11	11	Inhoud	Toont de schrijver begrip voor het voornemen van de directeur	0.76	1.46

^a: bij deze tabel zijn geen kleuren weergegeven, omdat er geen bovengrens kan worden gemarkeerd.

8.4 Kwalitatieve typering van leerlingteksten

8.4.1 Inleiding

Zoals beschreven waren er drie taakspecifieke beoordelingsaspecten: inhoud, opbouw en formulering. De taakspecifieke oordelen zijn uiteindelijk afgebeeld op een schrijfvaardigheidsschaal, zoals beschreven in hoofdstuk 5. Dat betekent dat uiteindelijk elk leerlingwerk een waarde krijgt op de schaal. Ten tweede zijn bij de standaardbepaling per taak teksten handmatig getypeerd in termen van de referentieniveaus 'net' 1F en 'net' 2F. In termen van de schrijfvaardigheidsschaal-percentielen voor het bo bleken de grenswaarden behorend bij deze niveaus terecht te komen op p28 en p73.

In deze paragraaf geven we een kwalitatief beeld van de resultaten van de twee beoordelingsmethoden. Dat doen we aan de hand van twee voorbeeldtaken. Voor die taken geven we aan de hand van voorbeeldteksten een kwalitatieve indruk van zes verschillende schrijfvaardigheidsniveaus. Naast de twee referentieniveaus 1F en 2F (met een grenswaarde van p28 bo en p73 bo) zijn dat de volgende percentielen op de schrijfvaardigheidsschaal: p10 sbo, p10 bo, p50 bo en p90 bo. Tabel 8.21 bevat voor alle zes niveaus de schaalpositie, het referentieniveau, het bo-

¹³⁷ Taak 8, opgave 24 is vergelijkbaar.

percentiel en het sbo-percentiel. De niveaus waarop respectievelijk de referentieniveaus 1F en 2F n t behaald zijn, zijn groen gearceerd.

Tabel 8.21: De zes niveaus in termen van WML-score, referentieniveau en percentielen

Rangorde	WML-score	Referentie-niveau	Percentiel bo	Percentiel sbo
1	-.478	< 1F	<p5*	p10
2	-.255	< 1F	p10	p40
3	-.128	1F	p28	p65
4	-.021	1F	p50	p80
5	.088	2F	p73	p90
6	.197	2F	p90	p96

* bij benadering

Het ging er nu om twee taken te kiezen uit de twaalf. Verschillende overwegingen hebben die keuze bepaald. Ten eerste wilden we graag twee taken kiezen die onderling zoveel mogelijk verschillen in communicatieve doelen en teksthandelingen. Ten tweede hadden we een voorkeur voor nieuwe taken, omdat we daarbij zelf verantwoordelijk waren geweest voor de beoordelingsitems en een deel van de analyse hieronder is gebaseerd op die items. Ten derde hanteerden we liefst taken met een OFI-beoordelaarsovereenstemming van minimaal 0.7. Ten vierde vielen taken af die wat minder betrouwbaar geschaald konden worden, bijvoorbeeld T12. Ten vijfde prefereerden we taken waarbij de insteek niet was om vooraf gegeven informatie te parafraseren (zoals T5) maar waarbij nieuwe informatie gegenereerd moest worden. Tenslotte, en zeer relevant, wilden we graag taken kiezen die bij de standaardbepaling gebruikt waren om zowel de grenswaarde voor 1F als 2F te bepalen. Zodoende zijn vergelijkingen tussen leerlingwerken lager dan 1F en op 1F en op 1F en op 2F mogelijk.

De combinatie van deze vijf criteria leidde ons tot T04 en T10. T04 vraagt om een brief aan een Zweeds meisje dat volgend jaar bij de schrijver in de klas komt. De brief beschrijft het leven op school. T10 betreft een brief aan de burgemeester over het voornemen om huizen te bouwen op een speelveldje dicht bij de school. De schrijver moet proberen de burgemeester dit plan te laten heroverwegen. In termen van doelen gaat het om respectievelijk informeren en overtuigen, en in termen van teksthandelingen om respectievelijk beschrijven en verzoeken plus argumenteren. Een nadeel is dat het in beide gevallen om brieven gaat, maar de hemelsbrede verschillen wat betreft tekstfunctie wegen voor ons zwaarder dan de overeenkomsten in conventies voor aanhef en afsluiting van de tekst.

Voor de selectie van de voorbeeldteksten zijn ook criteria gehanteerd. Omdat wij belang hechtten aan een betrouwbare inschatting van de tekstkwaliteit, hebben wij ervoor gekozen alleen voorbeeldteksten te selecteren uit de dubbel beoordeelde werken. Belangrijk was dat beide beoordelaars tot een vergelijkbaar oordeel kwamen en de scores op de O, F en I-schalen ook vergelijkbaar waren voor de taak. Met andere woorden, we hebben dezelfde selectiecriteria gehanteerd als voor de selectie van leerlingwerken bij de standaardbepaling. Daarom zijn we gestart

met alle reeds voor de standaardbepaling geselecteerde leerlingwerken. Aangezien deze teksten onvoldoende toereikend waren om de zes verschillende niveaus te beschrijven, zijn extra teksten geselecteerd op de genoemde psychometrische gronden. Het betrof hier met name teksten voor de niveaus p10 sbo, p10 bo en p90 bo.

8.4.2 Een analyse van de itemscores

Na de schaling van de leerlingwerken is het mogelijk de beoordelingsitems zelf als het ware weer te schalen: sommige items zijn eenvoudig en hebben al een hoge p-waarde voor leerlingen die laag scoren op de schrijfvaardigheidsschaal, terwijl andere items alleen succes opleveren voor leerlingen met een hogere schrijfvaardigheid. Dit hebben we ook gezien in paragraaf 8.3.2.

In onze 12 taken zijn uiteindelijk 176 beoordelingsitems gebruikt: 60 opbouwitems, 33 formuleringsitems en 83 inhoudsitems. Als indicator voor de moeilijkheid van de items gebruiken we hieronder de itemlocatie, dat wil zeggen de plek op de vaardigheidsschaal waar de verwachte waarde op een item 50% of 80% is van de maximumscore. In Tabel 8.23 en 8.24 geven we de itemlocaties voor de twee geselecteerde taken. Bedenk daarbij dat de gemiddelde schaalpositie van 1F ligt op $-.128$, en die voor 2F op $.088$.

Tabel 8.23: De itemlocaties voor T04, de brief aan Annika

Item-nummer	Aspect	Korte inhoud van het item	Max score	Itemlocatie	
				50%	80%
5	Opbouw en formulering	Aanhef aanwezig en informeel	2	-0.22	-0.04
7	Opbouw	Inhoud inleiding	2	0.08	0.27
8	Formulering	Enthousiasme, persoonlijk aanspreken en zelfreferentie	3	-0.01	0.16
11	Inhoud	Thema 1: hoe ziet Nederland eruit	2	-0.48	0.00
12	Inhoud	Thema 2: hoe dicht wonen Nederlanders op elkaar	2	-0.46	-0.09
13	Inhoud	Thema 3: wat eten Nederlanders graag?	2	-0.64	-0.32
14	Inhoud	Thema 4: van welke sporten houden Nederlanders	2	-0.51	-0.30
15	Inhoud	Thema 5: wat doen Nederlandse kinderen in vrije tijd	2	-0.63	-0.35
16	Inhoud	Verwijzing naar Zweedse situatie	1	0.28	0.98
18	Opbouw	Heldere thema-overgangen in woorden	1	0.76	1.22
19	Opbouw	Herhaling van vragen als kopjes	1	0.76	1.22
20	Opbouw	Verwijzing naar vragen met opsommingstekens	1	0.34	0.80
21	Opbouw	Antwoorden zonder indicatie van de vraag	1	0.69	1.39
27	Opbouw en formulering	Groet aanwezig en informeel	2	-0.05	0.14

In T04 blijkt de gevraagde informele aanhef goed beheerst (op 80%-niveau) te worden op schaalniveau -.04, dat wil zeggen ergens tussen 1F en 2F in. Een informele groet is wat minder aanwezig. Mogelijk zien we hier een onbedoeld effect van een opdrachtinstructie omwille van de privacy van de deelnemende leerlingen. Leerlingen kregen de instructie in hun teksten geen informatie te gebruiken die herleidbaar is naar henzelf als persoon, zoals hun naam of woonplaats. Hetzelfde patroon (veel passend formele aanheffen, maar minder vaak voorkomende passend formele briefafsluitingen) zien we bij T10.¹³⁸

Niet al te lastig blijken vijf van de zes inhoudsitems, waarin werd gevraagd om de vijf thema's uit de opdracht te behandelen (items 11-15). Landschap en bevolkingsdichtheid zijn duidelijk lastiger dan voedsel, sporten en vrijetijdsbestedingen. Het zesde inhoudsitem, verwijzen naar de Zweedse situatie (item 16), is blijkbaar veeleisender. Het wordt op 80%-niveau beheerst op schaalniveau .098, dat wil zeggen iets boven 2F-niveau.

¹³⁸ De getoonde voorbeeldteksten zijn ontdaan van iedere naar de leerling herleidbare informatie.

Op schaalniveau .16, dus boven 2F, bevindt zich de kans dat een leerling 80% scoort op het enige formuleringsaspect dat in deze taak werd gemeten: hanteren van een voor een informele brief aan een leeftijdsgenoot passende toon, gekenmerkt door enthousiasme, persoonlijk aanspreken en naar jezelf verwijzen met 'ik' (item 8).

De opbouwitems zijn in deze schrijftaak over het algemeen het lastigst. De kans dat een leerling een 50%-score haalde op de aanwezigheid van heldere thema-overgangen (item 18) bevindt zich (ver) boven het schaalniveau dat door minder dan 10% van de BO-leerlingen wordt beheerst. Ditzelfde geldt voor zichtbare samenhang tussen de informatie in de brief en de 'vragen van Annika' met kopjes (item 19) en opsommingstekens (item 20). Een inleiding schrijven (item 7) is iets minder lastig, ook al is een matige beheersing pas te verwachten bij niveau 2F.¹³⁹

Tabel 8.24: De itemlocaties voor T10, de brief over het speelveldje¹⁴⁰

Item-nummer	Aspect	Beschrijving	Max score	Itemlocatie	
				50%	80%
4	Opbouw	Datum/ plaats boven brief	1	0.94	1.40
5	Opbouw en	Aanhef aanwezig en formeel	2	-0.47	-0.14
7	Opbouw	Inhoud inleiding	3	-0.01	0.98
8	Inhoud	Verzoek aanwezig	1	-0.83	-0.36
9	Inhoud	Belang speelveld	1	-0.56	-0.21
10	Inhoud	Andere argumenten	1	0.66	1.35
11	Inhoud	Begrip voor voornemen gemeente	1	0.39	1.08
12	Formulering	Samenhang verzoek en motivatie	3	-0.17	0.43
13	Inhoud	Gevraagde concessie voorgesteld als minimaal	1	0.41	0.75
14	Inhoud	Ruimte om nee te zeggen	1	-0.12	0.34
15	Inhoud	Bij voorbaat dank getoond	1	0.78	1.47
22	Opbouw en formulering	Groet aanwezig en formeel	2	-0.16	0.19

De itemlocaties voor T10 liggen duidelijk hoger dan die bij T04. De passend formele aanhef (item 5) blijkt de meeste leerlingen weinig problemen op te leveren, en dat geldt ook voor het doen van een verzoek om af te zien van de bouwplannen (item 8) en het door de leerling benoemen van het belang van het speelveldje (item 9). Dat zijn twee items die in de opdracht zelf al genoemd worden (zie hieronder paragraaf 8.4.3).

Vergeleken met items 5 en 8 is het voor leerlingen lastiger het verband tussen het verzoek en de motivatie weer te geven (item 12, maar dat verband blijkt vaak al uit het direct laten volgen van het een op het ander) en ruimte te bieden om nee te zeggen (item 14), wat bijvoorbeeld kan door modale werkwoorden te gebruiken in de formulering van het verzoek. Rond 1F-niveau gebeurt dat in de helft

¹³⁹ We zijn achteraf niet erg tevreden over item 21. Met dat item bedoelden we de teksten te beschrijven die geen van de drie structureringsprincipes uit de items 18-19-20 vertonen; maar de itemlocatie van item 21 laat zien dat het anders is opgevat, zodat we afzien van een interpretatie ervan.

¹⁴⁰ N.B. de gemiddelde schaalpositie van 1F ligt op -.128, en die voor 2F op .088

van de gevallen. Iets dergelijks geldt voor een niet al te informele groet aan het einde (item 22). Zoals eerder gezegd, zou de privacy-instructie aan de leerlingen voor problemen hebben kunnen zorgen bij de groet, zodat we aan dit resultaat niet te veel waarde moeten hechten.

De andere items blijken erg hoog gegrepen. Daarbij gaat het bijvoorbeeld om het aanvoeren van andere argumenten (item 10) dan het belang van het speelveld (bijvoorbeeld: deze plek is sowieso geen goede plek voor huizen, zo direct bij de school) en om het tonen van begrip voor de behoefte aan woningbouw (item 11). Deze inhoudselementen zijn eerder een vorm van extra ondersteuning dan 'kerninhouden'. Verder geldt hetzelfde voor beleefdheidskenmerken als het minimaliseren van het verzoek (item 13, 'u kunt er minstens nog een keer over denken?') en het bedanken voor het in overweging nemen van de brief (item 15). Het geldt ook voor het plaatsen van een datum en plaats boven de brief in T10 (item 4); de extreem hoge itemlocaties laten zien dat deze conventie vrijwel onbekend is.

8.4.3 Kwalitatieve beschrijving van de vaardigheidsniveaus aan de hand van T04

De taak

T04 Brief aan Annika

Lees dit eerst!

Stel je voor....

Volgende maand komt een nieuw meisje bij jullie in de klas. Zij heet Annika en ze woont in Zweden. Daar heeft ze Nederlands geleerd. Annika is nog nooit in Nederland geweest en heeft wat vragen:

- Zweden heeft veel bossen. Hoe ziet Nederland eruit?
- In Zweden wonen weinig mensen. Je burens kunnen kilometers verderop wonen. Hoe dicht wonen de mensen in Nederland op elkaar?
- De Zweden eten graag haring in roomsaus. Wat eten Nederlanders graag?
- Zweden houden van ijshockey en skiën. Van welke sporten houden de Nederlanders?
- Wat doen Nederlandse kinderen in hun vrije tijd?

De juf vraagt jou om Annika een brief te schrijven met het antwoord op haar vragen.

Teksten en observaties niveau 1 (p10 SBO)

Hieronder bespreken we twee voorbeelden van teksten op het eerste niveau. De tekstvoorbeelden zijn letterlijk overgenomen, inclusief eventuele spelfouten, ontbrekende punten en hoofdletters, en onregelmatige regelsprongen.

4004119 (-0,453)

nederland heeft
niet veel bossen
hier is het verschillend
ze eten in nederland graag
stampot en patat

voetbal formule1

gamen en buiten
speelen en bromer
reiden

5059108 (-0.416)

hallo annika Wij zitten best klijn wij eten graag
haaring friet wij doen voetballen en de kinderen
spellen (onleesbaar) vrije je tijd

Deze teksten geven zich nauwelijks rekenschap van de kenmerken die onderscheidend zijn voor het genre 'brief'. De tweede tekst bevat een passend informele aanhef, maar bij beide teksten ontbreken de passend informele afsluiting en de bij een informele brief aan een leeftijdsgenoot passende toon. Wel vermelden de schrijvers beknopt een deel van de gevraagde inhoudselementen: informatie over hoe Nederland eruitziet, eetgewoonten en sportvoorkeuren van Nederlanders en er wordt iets gezegd over de vrijetijdsbesteding van Nederlandse kinderen. Het lastiger inhoudselement 'Nederland vergelijken met Zweden' ontbreekt. Aan passende opbouw heeft geen van deze briefschrijvers aandacht besteed.

4004107 (-0.313)

beste Annika,
nederlands heef ook best wel veel
bossen maar ik denk niet zo veel
bossen als wat er in
zweeden is. bij nederland wonen
mensen ook best wel dicht bij elkaar
maar het licht er aan welke huizen
het zijn. nederland houden
heel erg van stroopwafels wand
dat is tipis nederlands. nederland
houd van voetballen
daar kijken nederlanders heel
graag na

De tekst op niveau 2, (net) onder 1F, lijkt sterk op de teksten die we bespraken als illustratie bij niveau 1, maar er zijn twee belangrijke verschillen. In de eerste plaats is de informatie die deze briefschrijver verschaft over Nederland gedetailleerder. Teksten die aan de beschrijving van Nederland nieuwe informatie toevoegden aan wat er al in de opdracht gegeven was, kregen maximaal 2 punten per beschreven aspect. Teksten die geen nieuwe details toevoegden aan de beschrijvingen die al in opdrachttekst waren gegeven, kregen 1 punt. In tekst 4004107 ontbreekt een beschrijving van de vrijetijdsbesteding van Nederlandse kinderen. Een tweede verbetering ten opzichte van de teksten op niveau 1 betreft de vergelijking van de Nederlandse situatie met die in Zweden, die zichtbaar is aan het begin van brief 4004107.

Overigens moet worden opgemerkt dat deze tekst, brief 4004107, geschreven is door een sbo-leerling. Reden om deze tekst toch op te nemen is dat de spoeling van bruikbare, bij onze criteria aansluitende leerlingteksten op dit vaardigheidsniveau zeer dun was en deze brief op de genoemde beoordelingsaspecten duidelijk verschilde van de teksten op het eerste niveau. Inhoudelijk denken wij deze keuze te kunnen verantwoorden, omdat de leerlingwerken van zowel bo- als sbo-leerlingen allemaal op dezelfde schaal gerepresenteerd zijn. Er is daarom geen reden om aan te nemen dat sbo-leerlingen anders schrijven.

Teksten en observaties niveau 3 (1F)

1091121 (-0.143)

Hoi Annika,
Nederland heeft wel bossen maar niet zo veel als in Zweden. In Nederland wonen mensen best wel dicht op elkaar. Vooral in steden in dorpen is het wel minder. In Nederland houden veel mensen van haring en kibbeling. Nederlanders houden van voetbal en schaatsen. Nederlandse kinderen spelen buiten of gaan gewoon op zich telefoon of ipad.
Groetjes,

4058208 (-0.126)

Beste Annika

Nederland heeft ook best wel veel bossen maar ik denk dat Zweden meer bossen heeft dan Nederland. Misschien liggen je burens bij ons 1cm. Ikzelf eet het liefst: pizza, friet en pannenkoek. Het is eigenlijk per persoon, de ander vindt het lekker en de ander niet. De hobby's zijn eigen bij ons ook verschillend, de ander vindt paardrijden leuk en de ander zwemmen. Ik zou het liefst willen teken, lezen of computeren. Ik denk dat de meesten buiten spelen.

Bovenstaande teksten scoren rond niveau 1F. Ze delen met de voorbeelden van het vorige niveau de systematische en wat uitgebreidere bespreking van de kenmerken van Nederland en de aanzetten tot een vergelijking met Zweden. In beide teksten krijgen deze kenmerken nét wat uitgebreider en verzorgder vorm. In deze teksten zien we ook andere formuleringskenmerken van het genre opduiken. Beide teksten bevatten een passend informele aanhef; tekst 1091121 bevat daarnaast een passend informele afsluiting. In tekst 4058208 zien we een aanzet tot een persoonlijke stijl, die past bij een informele brief aan een leeftijdsgenoot: de brievenaar presenteert zichzelf in de brief met 'ik'.

1041112 (-0.032)

Nederland is een klein land met veel tulpen.
in Nederland woon je naast je burens en de huizen
zitten aan elkaar vast.
nederlanders eten graag: kaas, stroopwafels, haring, drop
nederlanders houden van voetbal, hockey, basketball, golf
kinderen in hun vrije tijd doen game, buitenspelen

dit zijn antwoorden op jouw vragen

doei

3061102 (0.027)

Beste Annika

Wat leuk dat je bij ons in de klas komt, ik heb er zin
in. Op de plek waar onze school ligt is een park in de
buurt maar geen bos 'jammer' (We hebben wel bossen alleen niet
bij ons in de buurt). Buren die kilometers verderop wonen klinkt
voor ons als een mop wij wonen in rijtjes huizen of heel dicht
bij elkaar.

Misschien woon je wel bij mij in de buurt? dan kun je
nederlands voetsel proeven zoals stroopwafels, stampot en kaas
echt nederlands 😊

IJshockey nog nooit gespeeld

Wij doen hier klompenlopen, zaklopen en koekhappen dat zijn vooral
spelletjes voor op koningsdag. In mijn vrije tijd ben ik buiten
met mijn vriendinnen of ik ga tekenen. het verschild met
wat je doet. Ieder zijn eigen hobby!

Ik hoop dat ik je vragen goed beantwoord hebt en dat
je nederland nu beter kent

groetjes groep 8

Tekst 1041112 deelt met de voorbeelden van niveau 3 de gedetailleerde bespreking van de kenmerken van Nederland. Vergelijking met Zweden ontbreekt, evenals de informele brieftaanhef. Wel bevat de tekst een passend informele afsluiting - waarvan je je weliswaar kunt afvragen of het de meest adequate is, maar hij voldoet aan het gestelde criterium. Interessante extra kenmerken in deze tekst zijn het persoonlijk aanspreken van Annika met 'je' en 'jouw' in de laatste regel van de bodytekst, en

ook de thema-aanduidingen telkens aan het begin van de zinnen die kenmerken van Nederland beschrijven (*'in nederland woon je naast je buren', 'nederlanders eten graag: kaas, stroopwafels'*). Deze thema-aanduidingen verwijzen expliciet terug naar de vragen van Annika. Door dat te doen helpt de briefschrijver Annika de opbouw, of de structuur van de tekst te doorzien. Zoals we hierboven bespraken is het aanbrengen van structuursignalen relatief moeilijk. Het verbaast dus niet dat we dit type kenmerken pas aantreffen bij de teksten die tussen niveaus 1F en 2F in zitten. Verder heeft deze tekst een recapitulerende slotzin, wat we eerder ook niet gezien hebben.

Tekst 3061102 scoort heel goed op de aspecten passend informele aanhef en passende toon: de briefschrijver toont enthousiasme over de komst van Annika, verwijst expliciet naar zichzelf met 'ik' en spreekt Annika regelmatig direct aan. Mooi is ook de aanwezigheid van een inleidende zin, waarmee de schrijver laat zien kennis te hebben van de conventies voor de opbouw van een brief. Veel andere brieven vallen met de deur in huis door meteen van wal te steken met kenmerkbeschrijvingen. Op een aantal andere structuurcriteria scoort de tekst echter minder goed. Anders dan bij het eerste voorbeeld van dit niveau ontbreken er hier heldere thema-overgangen. De beschrijvingen van Nederland worden wat terloops gepresenteerd als onderdeel van andere thema's, zoals in een beschrijving van de omgeving van de school en in een (overigens zeer enthousiasmerende en passende) uitnodiging om te komen eten. Het effect van deze presentatiewijze is dat de tekst minder goed gerelateerd kan worden aan de vragen die de fictieve Annika had; het is voorstelbaar dat een dergelijke manier van de gevraagde informatie presenteren tot wat verwarring zou leiden bij de fictieve ontvanger. De briefschrijver vergelijkt de Nederlandse situatie met die in Zweden ('Buren die kilometers verderop wonen klinkt voor ons als een mop'), maar ook die vergelijking blijft erg impliciet. Mooi is weer dat de tekst een afsluitende groet bevat. Maar minder sterk is het feit dat die afsluiting ('groetjes groep 8') niet goed aansluit bij de enkelvoudige zelfpresentatie ('ik') van de briefschrijver.

Teksten en observaties niveau 5 (2F)

1056125 (0.083)

Beste Annika,

Ik woon in Nederland en ik zal je er wat over vertellen.

Nou heb jij een paar vragen ik zal ze beantwoorden.

1. ja, in Nederland hebben we veel bos, maar minder dan in Zweden.
2. Bij ons staan de huizen dicht op elkaar je zit zelfs aan elkaar vast!
3. In Nederland is haring ook wel een specialist, maar zonder roomsaus
4. Nou bij ons skieën we niet wij voetballen.
5. gamen, voetballen, op hun telefoon, spelen

Groetjes,

2080224 (0.09)

Beste annika, ik ga je vragen nu beantwoorden dus luister goed. 1. nou nederland heeft niet veel bossen maar wel veel planten en allerlei mooie bloemetjes. 2. in nederland wonen je burens niet 10 kilometer verder maar gewoon naast je heel gezellig dus. 3 nederlanders eten graag stampot, en als ze nog zin hebben in een een toetje, dan nemen ze vla dat is heel lekker zoet. 4 nederlanders houden van fietsen dat is heel bekend hier en iedereen kan het eigenlijk wel in ons land. 5. de kinderen hier in nederland spelen in hun vrije tijd buiten of ze gaan gamen even als ze zin hebben. of ze gaan een stukje wandelen/fietsen groetjes van valerie uit nederland 😊

Tekst 1056125 bevindt zich op de niveauschaal een fractie onder referentieniveau 2F, tekst 2080224 een fractie erboven. Deze teksten delen de sterke punten van de hoogst scorende tekst die we bespraken voor niveau 4: er is een passend informele aanhef, de brief heeft een inleidende zin (of eigenlijk twee die elkaar inhoudelijk wat overlappen, wat minder goed is, maar dat is niet als criterium meegewogen in ons beoordelingsmodel) en de directe aansprekingen en zelfreferenties in de eerste regels van de brief zorgen voor een passende persoonlijke toon. Maar wat de kwaliteit van deze brieven nog wat sterker maakt zijn de heldere vergelijkingen met de situatie in Zweden en de gestructureerde verwijzing naar de afzonderlijke 'vragen van Annika' met opsommingstekens. Bijzonder fraai is deze strategie op zichzelf niet; lopend proza vinden we mooier in een brief. Maar in het kader van deze opdracht, waarin de te beschrijven aspecten van Nederland ook in een bolletjeslijst werden aangeboden, valt dit te zien als een vorm van efficiëntie. De beschrijvingen van Nederland in tekst 1056125 zijn overigens niet zeer uitgebreid. In tekst 2080224 is dat wel het geval.

Teksten en observaties niveau 6 (p90 BO)**3007128 (0.19)**

Hoi Annika.

Heel leuk dat je bij ons in de klas komt.
in Nederlands is het heel anders dan in Zweden.
Wij hebben niet super veel bossen in Nederland,
wow in nederland wonen mensen heel
dicht bij elkaar de meeste mensen hun huizen
zitten zelfs tegen elkaar!

3007128 (0.19) (vervolg)

in Nederland eten wij van alles: iets
typisch Nederlands is boerenkool en rookworst.

Nederlanders hebben niet een sport die iedereen
doet, maar de meest mensen kijken voetbal!

Wij Nederlanders doen in onze vrije tijd
van alles, de meeste mensen kijken filmpjes/youtube
Of gewoon op onze telefoon. Ik hoop dat je
hier wat aan heb! en tot snel!

Groetjes: A

1056111 (0.226)

Beste Annika,

ik had gehoord dat je bij ons in de
klas komt. en heb gehoord dat je al nederlands
kan. ik hoorde ook dat je een aantal vragen
had en die ga ik vandaag
beantwoorden. er is best veel bos
in Nederland en wij zijn ook een waterland.
Wij wonen in Nederland dichtbevolkt en dat
betekent dat we best dicht op
elkaar wonen. wat echt nederlands is is: boerenkool,
hutspot, zuurkol en rode kool maar patat
vinden we ook heel erg lekker. Wij houden heel
erg van voetbal. wij gamen en spelen graag in
onze vrije tijd. ik hoop dat je
voldoende antwoord hebt op je vragen en
tot snel

groetjes X

Teksten 3007128 en 1056111 bevinden zich rondom de grens van niveau 6 (p90 BO). Ze delen de sterke punten van de 2F-teksten: ze hebben een passend informele aanhef en afsluiting, de brieven worden netjes ingeleid, de inhoudelijke 'vragen van Annika' worden in detail beantwoord en de thema's worden in de meeste gevallen helder aangegeven. Met name de eerste brief maakt goed uitgewerkte vergelijkingen met de situatie in Zweden. Opmerkelijk is de spreektaalige uitdrukking van verbazing 'wow' aan het begin van de vierde regel. Vermoedelijk is dat een reactie op het element in de opdrachtbeschrijving dat stelt dat in Zweden 'je burens kilometers verderop kunnen wonen'. Zonder nadere context is deze terugverwijzing naar de Zweedse situatie niet heel geslaagd, maar de vorm

ervan past wel erg goed bij de formuleringseis dat het informele briefgenre een persoonlijke toon hanteert. Dit aspect dook bij de voorbeelden die we bespraken bij de lagere niveaus soms ook al wel op, maar in 3007128 en 1056111 is het veel consistentier doorgevoerd, en bovendien met een mooie variatie aan formuleringen gerealiseerd. De brieven onderscheiden zich ook door de aanwezigheid van een passende afsluiting. Deze zaken weerspiegelen de algemene observatie dat een verzorgde en heldere tekstopbouw en consistent doorgevoerde formuleringseisen relatief lastig worden gevonden; zij worden pas zichtbaar bij de hogere niveaus van taakgerichte schrijfvaardigheid.

8.4.4 Kwalitatieve beschrijving van de vaardigheidsniveaus aan de hand van T10

De taak

T10 Ons speelveld

Lees dit eerst!

Stel je voor... Naast jouw school ligt een speelveld. Er staan speeltoestellen, een net waar je in kunt klimmen en doelen voor voetbal of handbal. Vorige week hoorden jullie dat de gemeente daar huizen wil bouwen. Als dat doorgaat, raken jullie speelveld kwijt.

Alle leerlingen in jouw klas gaan een brief schrijven aan de burgemeester om te vragen jullie speelveld te beschermen en een andere plek voor de huizen te zoeken.

De opdracht

Schrijf een brief aan de burgemeester waarin je vraagt jullie speelveld te beschermen en een andere plek te zoeken voor de nieuwe huizen. Leg uit hoe belangrijk het speelveld voor jullie is.

Schrijf je brief op de volgende pagina.

Teksten en observaties niveau 1 (p10 SBO)

We laten nu twee voorbeelden volgen van teksten op het eerste niveau. De tekstvoorbeelden zijn weer letterlijk overgenomen, inclusief eventuele spelfouten, ontbrekende punten en hoofdletters, en onregelmatige regelsprongen.

4014201 (-.473) (hier en daar spaties toegevoegd)

lieve burgemeester.
kinder willen en moet kunnen gaan.
en dit kreen mooi plek om te spellen.
dus zoek maar een andere voor
dat

4010212 (-.473)

Wilt u het speelfeld laten staan
om dat het schoolplein saai is
en het speelfeld veel leuker en
verderop is een gras veld waar niks staat
daar kun je de huizen neer
zetten. van B

Onze observaties zijn de volgende. Allereerst bevatten de teksten op dit laagste niveau geen conventionele opening of slot. Ook inleidingen ontbreken. Wanneer we naar tekst 4014201 kijken, dan zie we dat deze incomplete en/of onbegrijpelijke zinnen bevat. Ook bevat deze tekst geen verzoek en geen argument. Een verzoek en argument is wel aanwezig in de tweede tekst 4010212. In deze leerlingtekst wordt tevens een alternatief aangedragen, zij het in een onsamenhangende, voortdurend doorlopende zin. Opvallend is dat beide teksten causale verbindingswoorden bevatten (*dus* resp. *omdat*), maar dat dit niet leidt tot een samenhangende tekst en evenmin tot een tekst die gepast is qua toon.

Teksten en observaties niveau 2 (p10 BO)

1030101 (-.0256)

Beste burgemeester.

Wij hebben een speelveld en jullie willen
Daar gaan bouwen. Dat vinden wij niet tof
dus vraag ik u om er over te Bespreken.

Mvr C.P.

1044128 (-.0256)

Geachte burgemeester
Ik ga altijd na school naar het speelveld
we spreken daar héél vaak af. Je kan
ook huizen bouwen buiten het dorp daar is een
heel groot gebied te koop.

1074122 (-.0256)

beste, burgemeester

Ik zou graag willen dat de speelvelden blijven, omdat het gezond is dat kinderen bewegen.

(doorhalingen)

Maar mijn vraag is waarom wilt u daar huizen bouwen?

kan het niet een andere plek zijn?

(doorhalingen)

met vriendelijke groet, ...

De teksten bevatten in tegenstelling tot die op niveau 1 begrijpelijke en helder gesegmenteerde zinnen. Verder bevatten alle teksten een aanhef, en twee teksten ook een afsluiting. In de eerste tekst verschijnt een 'inleidingszin', die de aanleiding voor de brief schetst.

De beperkingen springen echter ook in het oog. Het verzoek is op dit niveau nog niet ondubbelzinnig aanwezig. In de eerste tekst is dat een formuleringskwestie, in de tweede staat alleen een alternatieve suggestie, en de derde bevat een wens ('ik zou graag willen ...), gevolgd door een klacht. Geen enkele tekst bevat een vorm van begrip of van beleefdheid. De teksten tonen vooral ergernis; de eerste twee teksten zijn daarnaast ongepast informeel. De eerste teksten bevatten slechts elementaire argumentatie: het speelveld wordt gebruikt, en is daarom van belang. De laatste tekst voert een extra argument aan, namelijk gezondheid. Wederom blijken de causale verbindingswoorden (*dus, omdat*) geen garantie voor een gepaste formulering (eerste tekst).

Teksten en observaties niveau 3 (1F)

1021116 (-.135)

Y.Z

Beste burgenmeester.

Ik ben er niet meens dat er huissen worden
ge plaatst mijn kind speeld daar
met haar vriendinnen niemand is er
mee eens kunt u asjblieft die
huisen er niet zetten. Groetjes Y.Z

1005124 (-.104)

Geachte heer/mevrouw

er worden huizen gezet waar wij altijd
Spelen. we wilden u vriendelijk vragen
of u de huizen ook ergens anders wilt
neerleggen?

met vriendelijke groet de school

1120302 (-.104)

Beste burgemeester,

Ik zou u graag vragen om het speelveld
te houden.

Als er huizen gebouwt worden, denk ik niet
dat de bewoners helemaal tevreden zouden
zijn; er spelen kinderen buiten op het schoolplein
dat kan voor de bewoners niet fijn zijn,
omdat kinderen wel herrie kunnen maken,
bovendien hoeft u ook geen geld te besteden aan
de bouw van de huizen.

Bij deze vraag ik u om het speelveld te houden.

Wederom bevatten alle teksten een aanhef, en twee ervan ook een afsluiting. Een inleidende zin komt slechts één keer voor. Echter, alle teksten bevatten wel een verzoek. Bovendien bevatten alle verzoeken beleefdheidskenmerken: *we wilden u vriendelijk vragen ...*, *kunt u alsjeblieft .*, *ik zou u vragen om ...*

Daarentegen is de argumentatie is nog gebrekkig. In de tweede tekst ontbreekt argumentatie geheel, in de eerste is deze wederom elementair. De laatste tekst biedt een relatief uitgebreide argumentatie; maar die richt zich niet op het belang van het speelveld maar tegen het bouwen van huizen. Dat lijkt geen kansrijke keuze, nog afgezien van dat de opdracht vraagt om een andere argumentatierichting.

1074214 (-.041)

Beste burgemeester,

ik heb gehoord dat u op het speelveldje naast onze school huizen wil gaan bouwen. Wij spelen daar heel vaak, en ik wil u vragen om een anderen plek uit te kiezen voor de huizen, want wij zijn heel gehecht aan dat speelveld.

Groeten W.

2103111 (-.041)

Ons speelveld!

Beste burgemeester ik hoorde dat er plannen zijn om het speelveld af te breken en er huizen te bouwen. Maar dat is ons speelveld! We kunnen er nog jaren spelen zonder dat het saai word. Daarnaast ligt onze school en als jullie daar huizen willen bouwen gaan onze leuke speeltoestellenweg en dat willen we niet! Misschien kunnen jullie op een andere plek huizen bouwen? Dat zouden we erg waarderen.

3038114 (-.041)

aan de burgemeester

beste burgemeester.

Hierbij vragen wij u of u ons speelveld wil laten staan en de huizen op een andere plek te laten bouwen. We willen graag dat het speelveld blijft staan omdat er zulke mooie herinneringen zijn, en we spelen er heel vaak en het is ook zo mooi. Iedereen wil dat het blijft staan. want er zijn niet zoveel goede plekken voor een speelveld, maar er zijn wel veel goede plekken voor huizen.

De hierboven geselecteerde teksten liggen alle drie op -0.41 op de schrijfvaardigheidsschaal. Alle teksten bevatten een aanhef, één ervan een afsluiting. Er zijn nu twee teksten met een inleidende zin. Wederom bevatten alle teksten een verzoek, veelal ook met beleefdheidskenmerken zoals modale werkwoorden (*wil u vragen*) en woorden die ruimte bieden (*misschien kunnen jullie ...*).

Maar vooral valt op dat het belang van het speelveld krachtiger positief wordt verwoord: de schrijvers zijn *gehecht* aan het speelveld, zij kunnen er nog “jaren spelen voordat het saai wordt”, of zij hebben er *mooie herinneringen* aan.

Teksten en observaties niveau 5 (2F)

1057107 (.096)

Geachtte Burgermeester we hoorden dat jullie
nieuwe huizen wilden bouwen op ons speel
terrein, maar alle kinderen van onze school wil
dat niet omdat we dan niet lekker kunnen spelen
Zou u daar iets aan kunnen doen bijv.
Het ergens anders kunnen doen
Vriendelijke groet X. Y.

1074118 (.096)

Geachte burgemeester

Wij hebben gehoord dat de gemeente
huizen gaat bouwen op het speel
veldje naast onze school.
Maar wij spelen daar zo vaak.
het is zo'n leuk speeltuintje wand
wij kunnen daar voetballen
klimmen en nog meer anderen dingen.
wij zouden het daarom niet
leuk vinden als het speeltuintje
weg gaat. Dus wij wouden u
vragen of u hier iets aan
kan doen.

We houden contact.

met vriendelijke groeten
de XXX school

Ook hier hebben beide teksten weer dezelfde score op de schrijfvaardigheidsschaal (.096, iets boven de grenswaarde voor 2F). De teksten bevatten nu een aanhef en een afsluiting. Verder is het verschil met het vorige niveau bescheiden. De teksten bevatten een inleidende zin, een beleefd verzoek, en daarnaast positief geformuleerde argumentatie voor het belang van het speelveld.

1063405 (.173)

Geachtte burgemeester.

We hoorden dat u huizen wilt gaan bouwen op ons speelveld.

We hoopten dat u uw plannen ergens anders kunt toepassen omdat: dat speelveld betekent veel voor ons.

Het is de enige plek waar we echt leuk kunnen spelen dus haal het a.u.b. niet weg.

Vriendelijke groet X.

1078221 (.173)

Beste meneer de burgenmeester.

Het is zonde van het speelveld. Veel kinderen spelen hier en U doet deze kinderen er pijn mee. Kunt u de huizen misschien op een andere plek bouwen? Denk aan het oude terrein Buiten de stad of de verlaten kaasfabriek.

Dankuwel anoniempje07

mens@watdanook.nl

[mobiel nummer]

2119107 (.173)

Beste burgemeester,

Wij hebben gehoord dat naast onze school het speelveld word vervangen door huizen. toen ik en mijn klas

Dat hoorde waren we heel verdrietig want we willen graag het speelveld houden omdat

we na school daar altijd snel kunnen afspreken omdat het heel dichtbij is en het is een plek

waar iedereen mag komen en spelen en daarom willen we het speelveld houden dus wij hopen dat

u dat begrijpt en misschien ergens anders

huizen bouwd. bedankt voor het begrip en met vriendelijke groeten van X.

2042114 (.214)

Beste burgemeester,

Naast onze school staat een superleuk speelweijtje

Je kan er echt super leuk dingen doen, klimmen, voetballen en nog veel meer.

Elke keer als ik klaar ben met school ga ik er spelen ook gaan we er met de klas af en toe daar ook spelen

Maar vorige week hoorde ik dat het speelveldje

weg moest omdat er huizen worden gebouwd. We vinden het heel erg omdat ons speelveldje dan weg moet. Burgemeester wilt u alstublieft ons speelveldje Beschermen en zorgen dat het veldje blijft staan. Want als hier huizen moeten komen kunnen wij er nooit meer spelen. Wilt u zorgen dat er ergens anders huizen worden gebouwd. het veldje is zo belangrijk voor ons. Zoek alstublieft een andere plek voor huizen!

Groetjes van: mij

Tel [mobielnummer]

Email: xyz@gmail.com

Plaats: Plantenwijk

De teksten bevatten weer een aanhef en een afsluiting, een beleefd verzoek en positieve argumentatie. Drie van de vier teksten bevatten een zin over de aanleiding voor de brief. Twee daarvan gebruiken die zin om er een teleurgestelde reactie aan te koppelen. In de laatste tekst is de constructie het meest vernuftig: daar verschijnt de aanleiding, na een schets van onbekommerd spelplezier, als een wrede streep door de rekening. Verder laat het hoogste niveau een meer gepolijst en correcter taalgebruik zien. Ten slotte bevatten de teksten meer details die het mogelijk maken om contact op te nemen over de brief.

8.4.5 Tot slot

In deze paragraaf hebben we een demonstratie gegeven van de werkwijze bij de beoordeling van leerlingteksten op taakspecifieke eisen. Het ging hierbij telkens om eisen op het gebied van inhoud, formulering en tekstopbouw die belangrijk waren voor het genre 'informele brief' waarin het communicatieve doel 'beschrijven' centraal stond (Taak 04), en voor het genre 'formele brief' waarin de communicatieve doelen 'verzoeken' en 'argumenteren' centraal stonden (taak 10). Spelling, interpunctie en grammaticale correctheid bleven hierbij buiten beschouwing. Over deze correctheidsaspecten wordt in de volgende paragraaf gerapporteerd.

We hebben aan de hand van voorbeelden van leerlingteksten gedemonstreerd hoe de beoordelingscriteria voor beide taken zijn toegepast in deze peiling. Ook hebben we laten zien wat de beste en de zwakste leerlingen kunnen op het gebied van taakspecifieke schrijfvaardigheid. Uit zowel de kwantitatieve gegevens die we rapporteerden als uit de kwalitatieve beschrijvingen van niveauverschillen valt op te maken dat de taakspecifieke vaardigheden zelf ook weer geschaald kunnen

worden. Over het algemeen zijn de eisen op het gebied van ‘inhoud genereren’ het gemakkelijkst, terwijl de formuleringsaspecten en een adequate aspecten van tekstopbouw lastiger te realiseren zijn.

8.5 Algemene indruk, correctheid en aantal woorden

8.5.1 Resultaten algemene indruk, correctheid en aantal woorden

De beoordelingen van de leerlingteksten vonden plaats aan de hand van het beantwoorden van concrete vragen over de aanwezigheid van opbouw-, formulerings- en inhoudsaspecten. Op deze manier werd de communicatieve effectiviteit van de leerlingteksten in kaart gebracht. Zoals in paragraaf 1.4.2 en ook in hoofdstuk 5 is beschreven, vormden correctheidsaspecten geen onderdeel van de vaardigheidsmeting. Leerkrachten laten deze aspecten echter vaak wel een rol spelen bij de beoordeling van leerlingteksten. Ook de communicatieve effectiviteit van een tekst kan beïnvloed worden door veel correctheidsfouten, wanneer zij de herkenning van de bedoelde woorden en constructies belemmeren. In deze paragraaf gaan we in op de aspecten die aan het eind van de beoordelingsformulieren zijn bevraagd: de algemene indruk van het geschrevene, de lengte van de tekst en de aanwezigheid van grammaticale, interpunctie- en spelfouten. De vraag naar de algemene indruk betrof een holistisch oordeel over de tekst. Daarmee is dit oordeel principieel anders tot stand gekomen dan de analytische oordelen die eerder in dit hoofdstuk zijn besproken. Net als in de rest van dit hoofdstuk is er bij de beschrijving van de resultaten een onderscheid gemaakt tussen de ankeropgaven en de nieuwe opgaven en de resultaten in het bo en sbo.

Resultaten bij de Ankertaken

Bij de ankertaken is gevraagd naar de algemene indruk, schrijffouten en het aantal woorden. Bij de algemene indruk van de leerlingtekst konden drie oordelen worden gegeven: onvoldoende, voldoende en goed, gecodeerd als 0, 1 en 2. Ook bij schrijffouten kon de beoordelaar kiezen uit drie niveaus: a) veel schrijffouten, b) niet veel, maar ook niet weinig schrijffouten (‘midden’) en c) weinig of geen schrijffouten – gescoord als ‘0’, ‘1’ en ‘2’. De resultaten voor algemene indruk en correctheid/schrijffouten worden in Tabel 8.25a en 8.25b gegeven.

Tabel 8.25a en 8.25b: Resultaten algemene indruk en schrijffouten bij de Ankeropgaven, bo en sbo

BO								
ANKERS	Algemene indruk				Schrijffouten			
	T01	T03	T08		T01	T03	T08	
N	879	889	688		N	879	886	692
Gem	1.11	0.98	0.81		Gem	1.22	1.33	1.38
SD	0.654	0.673	0.66		SD	0.727	0.722	0.719

	T01	T03	T08	
	%	%	%	overall
Onvoldoende	16.5	23.8	33.3	23.67
Voldoende	56.1	54.7	52.8	54.68
Goed.	27.4	21.5	14.0	21.65
Voldoende/Goed	83.5	76.2	66.7	76.33

	T01	T03	T08	
	%	%	%	overall
Veel	17.7	14.9	14.0	15.5
Midden	42.2	36.9	34.0	37.5
Weinig	40.0	48.2	52.0	47.0
Veel/midden	60.0	51.8	48.0	53.0

SBO								
ANKERS	Algemene indruk				Schrijffouten			
	T01	T03	T08		T01	T03	T08	
N	142	154	123		N	142	154	133
Gem	0.62	0.57	0.51		Gem	0.55	0.64	0.85
SD	0.604	0.625	0.632		SD	0.602	0.73	0.793

	T01	T03	T08	
	%	%	%	overall
Onvoldoende	44.4	50.0	56.1	49.55
Voldoende	49.3	42.9	36.6	43.57
Goed.	6.3	7.1	7.3	6.88
Voldoende/Goed	55.6	50.0	43.9	50.45

	T01	T03	T08	
	%	%	%	overall
Veel	50.7	51.3	39.8	47.1
Midden	43.7	33.8	35.3	37.4
Weinig	5.6	14.9	24.8	15.5
Veel/midden	94.4	85.1	75.2	84.5

De algemene indruk over de volledige teksten verschilt per taak. Voor taak 1 is de algemene indruk het hoogst en voor taak 8 het laagst. Dit resultaat wordt niet gevonden bij het aantal schrijffouten. In het bo zijn de gemiddelden voor schrijffouten redelijk vergelijkbaar voor de drie verschillende taken.

Evenals bij de eerder in dit hoofdstuk gepresenteerde resultaten op basis van de schrijfvaardigheidsschaal is het duidelijk dat de algemene, holistische indruk van de tekst bij de bo-leerlingen beduidend hoger is dan bij de sbo-leerlingen. Dit zien we in de gemiddelden, maar ook in de percentages onvoldoende. In het bo wordt bij de ankertaken in totaal iets minder dan een kwart van de schrijfproducten als onvoldoende beoordeeld, terwijl dit in het sbo bijna de helft betreft. Het percentage algeheel goed bevonden schrijfproducten verschilt ook: 21.7% in het bo en 6.9% in het sbo. Bij schrijffouten is bovendien duidelijk te zien dat in het sbo veel meer schrijffouten worden gemaakt dan in het bo. In paragraaf 8.5.2 wordt ingegaan op de samenhang tussen scores op de algemene indruk en op het aantal schrijffouten en de (taak)vaardigheidsscores.

Resultaten bij de nieuwe taken

Vergelijkbare patronen worden ook gevonden bij de nieuwe taken (zie Tabellen 8.26a en 8.26b). Hieronder worden achtereenvolgens de resultaten voor algemene indruk, de hoeveelheid tekst, interpunctiefouten, niet lopende constructies en spelfouten bij de nieuwe taken gegeven. Per vraag worden de resultaten apart voor bo en sbo gegeven. Bij algemene indruk worden naast de percentages voor slecht, onvoldoende, voldoende en goed beoordeelde werken (gecodeerd als 0, 1, 2 en 3), ook de percentages van de tweedeling "slecht/onvoldoende" en "voldoende/goed" gegeven. De gemiddelden geven de gemiddelde scores weer, waarbij een hoger gemiddelde staat voor een betere algemene indruk.

Tabel 8.26a en 8.26b: Resultaten algemene indruk bij de nieuwe opgaven, bo en sbo

BO										
NIEUW	Algemene indruk									overall
	T02	T04	T05	T06	T07	T09	T10	T11	T12	
N	734	726	845	725	829	754	781	755	773	
Gem	2.04	1.79	1.55	1.78	1.76	1.72	1.74	1.73	1.7	1.75
SD	0.706	0.83	0.941	0.776	0.919	0.807	0.748	0.732	0.737	0.80

Algemene indruk										
Indruk	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
Slecht	2.2	8.0	16.0	6.6	10.7	8.4	6.1	6.6	6.5	8.0
Onvoldoende	16.2	23.3	28.8	23.4	24.6	25.1	26.0	24.0	27.3	24.4
Voldoende	56.5	50.7	39.4	55.0	42.2	52.4	55.7	59.1	56.1	51.6
Goed.	25.1	18.0	15.9	14.9	22.4	14.2	12.2	10.3	10.1	15.9
slecht/onv	18.4	31.3	44.7	30.1	35.3	33.4	32.1	30.6	33.8	32.4
vold/goed	81.6	68.7	55.3	69.9	64.7	66.6	67.9	69.4	66.2	67.6

SBO										
NIEUW	Algemene indruk									overall
	T02	T04	T05	T06	T07	T09	T10	T11	T12	
N	120	114	158	134	158	145	140	138	133	
Gem	1.74	1.23	0.83	1.51	1.14	1.28	1.21	1.43	1.42	1.29
SD	0.845	0.893	0.784	0.773	0.885	0.752	0.88	0.735	0.771	0.81

Algemene indruk										
Indruk	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
Slecht	8.3	22.8	39.9	9.7	29.7	15.2	23.6	9.4	11.3	19.5
Onvoldoende	26.7	39.5	38.0	36.6	29.7	44.1	37.9	42.8	41.4	37.4
Voldoende	47.5	29.8	21.5	46.3	37.3	37.9	32.1	42.8	41.4	37.1
Goed.	17.5	7.9	0.6	7.5	3.2	2.8	6.4	5.1	6.0	6.0
slecht/onv	35.0	62.3	77.8	46.3	59.5	59.3	61.4	52.2	52.6	56.9
vold/goed	65.0	37.7	22.2	53.7	40.5	40.7	38.6	47.8	47.4	43.1

De algemene, holistische indruk van de leerlingteksten bij de nieuwe taken is wederom duidelijk hoger in het bo dan in het sbo. Verder zien we aardig wat verschillen over de taken heen. Taak 2 is een taak waarvoor de algemene indruk van de beoordelaars redelijk hoog is (ruim 80% in bo en 65% in het sbo met voldoende/goed), terwijl Taak 5 een omgekeerd beeld laat zien, met een beoordeling van voldoende/goed voor slechts 55% in het bo en 22% in het sbo. Bij de andere taken worden minder grote afwijkingen van het gemiddelde gevonden.

Kijken we naar de hoeveelheid tekst dan zien we ook duidelijke verschillen over taken en tussen bo en sbo (Tabellen 8.27a en b). De hoeveelheid tekst is vastgesteld aan de hand van het aantal geschreven regels, waarbij een indeling is gemaakt in een tekst bestaande uit 1-5 regels, 6-10 regels, 11-15 regels en meer dan 15 regels (gescoord als resp. 0, 1, 2 en 3). Tabellen 8.27a en 8.27b geven respectievelijk voor bo en sbo de gemiddelde score per taak weer en de percentages voor zeer korte tot lange teksten (resp. 0 – 3). Uit de Tabellen 8.27a en 8.27b kan worden afgeleid dat met name bij taak 2 en 4 de leerlingen tamelijk veel tekst produceren. De teksten bij taak 6 zijn in het bo relatief kort ten opzichte

van de teksten bij de andere taken. Over het algemeen zijn de teksten van bo-leerlingen wat langer dan teksten van sbo-leerlingen.

Tabel 8.27a en 8.27b: Resultaten hoeveelheid tekst, bo en sbo

BO										
	Hoeveelheid									
	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
N	734	725	845	724	831	754	780	756	774	
Gem	2,41	1,78	1,47	1,04	1,66	1,11	1,17	1,36	1,33	1,48
SD	0,85	0,94	0,89	0,68	0,84	0,77	0,80	0,84	0,88	0,83
	Hoeveelheid									
# zinnen	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
1-5 regels	3,7	7,7	12,7	18,4	6,5	20,4	18,1	12,8	16,1	12,9
6-10 regels	13,1	34,2	41,3	62,4	38,7	53,2	54,1	49,2	46,0	43,6
11-15 regels	22,2	30,5	32,0	16,3	37,1	21,8	20,9	26,9	26,1	26,2
>15 regels	61,0	27,6	14,1	2,9	17,7	4,6	6,9	11,1	11,8	17,3

SBO										
	Hoeveelheid									
	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
N	120	115	158	134	158	145	140	138	135	
Gem	1,83	1,05	0,85	0,97	1,22	0,68	0,98	0,91	1,15	1,06
SD	1,07	0,86	0,92	0,74	0,86	0,82	0,89	0,87	0,92	0,88
	Hoeveelheid									
# zinnen	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
1-5 regels	13,3	25,2	43,0	24,6	17,1	49,7	32,1	35,5	26,7	30,2
6-10 regels	27,5	53,0	36,1	57,5	55,1	37,2	46,4	45,7	40,7	44,4
11-15 regels	22,5	13,0	13,3	14,2	16,5	8,3	12,9	11,6	23,7	15,0
>15 regels	36,7	8,7	7,6	3,7	11,4	4,8	8,6	7,2	8,9	10,5

In Tabellen 8.28a en 8.28b en 8.29a en 8.29b worden de resultaten voor de correctheidsaspecten als interpunctie en niet-lopende zinsconstructies getoond. Bij de niet-lopende constructies, dat wil zeggen bij grammaticaal incorrecte zinnen of bij incorrect gebruik van vaste uitdrukkingen, zijn drie categorieën gebruikt: in meer dan de helft van de zinnen (=‘0’), in minder dan de helft van de zinnen (=‘1’) en in geen enkele zin (=‘2’). Ook bij de interpunctiefouten, in deze peiling opgevat als het ontbreken van hoofdletters en punten, is deze scoring gebruikt, lopend van ‘0’ (in meer dan de helft van de zinnen), ‘1’ (in minder dan de helft van de zinnen) en ‘2’ (in geen enkele zin). De gemiddelden in Tabellen 8.28a en 8.29a zijn gebaseerd op deze indeling. Dat wil zeggen dat het gemiddelde van 0.77 bij Taak 2 (bo) betekent dat leerlingen gemiddeld in iets minder dan de helft van de zinnen interpunctiefouten maken. Het gemiddelde geeft dus niet het absolute gemiddelde aantal fouten weer.

In de Tabellen 8.28b en 8.29b worden, respectievelijk voor bo en sbo, de percentages weergegeven voor teksten 1) waarin in meer dan de helft van de zinnen interpunctiefouten/niet lopende constructies voorkomen, 2) waarin in minder dan de helft van de zinnen interpunctiefouten/niet lopende constructies voorkomen en 3) waarin deze fouten of niet lopende teksten niet voorkomen. Kijken we nu naar de vier tabellen (Tabellen 8.28a en b en 8.29a en 8.29b), dan zien we bij de interpunctiefouten verschillen tussen bo en sbo. Leerlingen in het sbo maken over de taken heen in 72-91% van de teksten in meer dan de helft van de zinnen interpunctiefouten (zie Tabel 8.28b), terwijl dit percentage bij het bo varieert van 34-56% (zie Tabel 8.28a). Verschillen tussen bo en sbo zien we in mindere mate terug bij niet-lopende constructies (Tabellen 8.29a en 8.29b).

Tabel 8.28a en 8.28b: Resultaten interpunctiefouten, bo en sbo^a

BO										
Interpunctie fouten										
	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
N	733	725	844	725	830	755	781	756	774	
Gem	0,77	0,84	0,88	0,77	0,63	0,91	1,00	0,86	0,93	0,84
SD	0,77	0,80	0,84	0,83	0,78	0,82	0,83	0,83	0,85	0,82
Interpunctie fouten										
# zinnen	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
>helft	43,8	41,2	41,6	49,0	55,7	38,8	34,4	42,6	39,7	43,0
<helft	35,1	33,4	28,7	25,2	25,3	31,8	31,1	28,7	27,6	29,6
geen enkele	21,1	25,4	29,7	25,8	19,0	29,4	34,4	28,7	32,7	27,4

SBO										
Interpunctie fouten										
	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
N	120	114	158	134	158	145	140	138	135	
Gem	0,20	0,40	0,20	0,25	0,14	0,31	0,24	0,26	1,04	0,33
SD	0,51	0,70	0,52	0,59	0,46	0,67	0,59	0,61	0,68	0,59
Interpunctie fouten										
# zinnen	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
>helft	85,0	71,9	86,1	83,6	90,5	80,7	83,6	82,6	80,0	83,0
<helft	10,0	15,8	8,2	8,2	5,1	7,6	8,6	8,7	12,6	9,2
geen enkele	5,0	12,3	5,7	8,2	4,4	11,7	7,9	8,7	7,4	7,8

^a: Het gemiddelde geeft niet het absolute gemiddelde aantal fouten weer (zie bovenstaande tekst).

Tabel 8.29a en 8.29b: Resultaten niet lopende constructies, bo en sbo^a

BO										
Niet lopende constructies										
	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
N	732	725	844	724	829	755	781	756	774	
Gem	1,38	1,26	1,23	1,26	1,39	1,41	1,40	1,19	1,24	1,31
SD	0,55	0,61	0,63	0,63	0,61	0,59	0,61	0,65	0,63	0,61
Niet lopende constructies										
# zinnen	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
>helft	3,1	8,8	11,4	10,2	6,5	5,4	6,8	13,4	11,0	8,5
<helft	55,9	56,8	54,7	53,3	48,0	48,5	46,1	54,4	54,3	52,4
geen enkele	41,0	34,3	33,9	36,5	45,5	46,1	47,1	32,3	34,8	39,1

SBO										
Niet lopende constructies										
	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
N	120	115	158	134	158	145	139	138	135	
Gem	1,23	1,03	0,83	1,24	1,13	1,19	0,99	0,94	0,27	0,98
SD	0,61	0,70	0,73	0,64	0,70	0,66	0,73	0,71	0,59	0,68
Niet lopende constructies										
# zinnen	T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
>helft	10,0	22,6	36,7	11,2	19,0	13,8	26,6	28,3	20,7	21,3
<helft	57,5	51,3	43,7	53,7	49,4	53,1	47,5	49,3	54,1	50,8
geen enkele	32,5	26,1	19,6	35,1	31,6	33,1	25,9	22,5	25,2	27,9

^a: Het gemiddelde geeft niet het absolute gemiddelde aantal fouten weer (zie bovenstaande tekst).

Voor spelfouten is een vergelijkbare scoring als bij interpunctiefouten en niet lopende constructies gehanteerd. Er zijn weer drie categorieën gebruikt: in meer dan de helft van de zinnen (=‘0’), in minder dan de helft van de zinnen (=‘1’) en in geen enkele zin (=‘2’). De resultaten in Tabellen 8.30a en 8.30b laten zien dat leerlingen in het bo gemiddeld minder spelfouten maken in hun teksten dan sbo-leerlingen: een lagere gemiddelde waarde in de tabel staat namelijk voor meer fouten. Wanneer we kijken naar het aantal teksten met veel spelfouten (in meer dan de helft van de zinnen) en geen spelfouten, dan wordt dit beeld bevestigd. Het percentage teksten waarin in meer dan de helft van de zinnen spelfouten staan, is gemiddeld beduidend hoger bij sbo-leerlingen (37%) dan bij bo-leerlingen (11%). Bovendien is het gemiddelde percentage teksten zonder spelfouten in het sbo (15%) aanzienlijk lager dan in het bo (30%).

Tabel 8.30a en 8.30b: Resultaten spelfouten, bo en sbo^a

BO											
		Spelfouten									
		T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
N		733	723	844	725	832	755	781	755	774	
Gem		1,04	1,08	1,16	1,30	1,16	1,29	1,26	1,18	1,26	1,19
SD		0,52	0,58	0,59	0,59	0,56	0,64	0,67	0,63	0,62	0,60
		Spelfouten									
# zinnen		T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
>helft		11,2	13,4	10,9	6,8	8,8	10,2	12,9	12,1	9,9	10,7
<helft		73,3	65,6	62,2	56,3	66,6	50,9	48,3	57,6	54,5	59,5
geen enkele		15,6	21,0	26,9	37,0	24,6	38,9	38,8	30,3	35,5	29,9

SBO											
		Spelfouten									
		T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
N		120	115	158	134	157	145	140	138	135	
Gem		0,65	0,70	0,72	0,96	0,73	1,03	0,66	0,66	0,85	0,78
SD		0,62	0,73	0,72	0,61	0,63	0,73	0,68	0,66	0,66	0,67
		Spelfouten									
# zinnen		T02	T04	T05	T06	T07	T09	T10	T11	T12	overall
>helft		42,5	46,1	43,7	20,9	36,9	24,8	45,7	44,2	30,4	37,1
<helft		50,0	38,3	41,1	62,7	53,5	46,9	42,9	45,7	54,1	48,4
geen enkele		7,5	15,7	15,2	16,4	9,6	28,3	11,4	10,1	15,6	14,5

^a: Het gemiddelde geeft niet het absolute gemiddelde aantal fouten weer (zie bovenstaande tekst).

8.5.2 Relatie algemene indruk en correctheidskenmerken met vaardigheid

De vraag is of de hierboven beschreven algemene, holistische indruk over de tekst en de correctheidskenmerken van de geschreven teksten samenhangen met de schrijfvaardigheid van de leerlingen, die is gemeten in termen van opbouw-, formulerings- en inhoudsaspecten. Hiertoe zijn per taak per algemeen en correctheidsaspect effectgroottes berekend van deze algemene en correctheidsscores op de overall vaardigheidsscore en op de taakvaardigheid. Dit is gedaan door te kijken naar de gestandaardiseerde verschillen in (taak)vaardigheidsscores tussen groepen leerlingen met verschillende scores op algemene indruk, tekstlengte en de drie correctheidsaspecten. Meer

concreet betekende dit het volgende voor bijvoorbeeld algemene indruk.¹⁴¹ Bij algemene indruk diende de beoordelaar zijn holistische oordeel te geven ('wat is de kwaliteit van de tekst'): slecht, onvoldoende, voldoende of goed, respectievelijk gescoord als 0, 1, 2 en 3. De effectgrootte op de overall vaardigheidsscore is voor algemene indruk het verschil in gemiddelde gewogen schrijfvaardigheidsscore (OFI-WML) tussen de groep leerlingen met score X op algemene indruk en die van leerlingen met score Y op algemene indruk. De effectgrootte op taakvaardigheid is op eenzelfde wijze berekend: $((\text{gewogen taak_WML}_{\text{score}_x} - \text{gewogen taak_WML}_{\text{score}_y}) / \text{gepoolde s.d.})$. In beide gevallen gaat het om een vergelijking in (taak)vaardigheidsscores bij groepen leerlingen met score 0 versus 1, met score 1 versus 2 en met score 2 versus 3. Er is sprake van een impact bij een effectgrootte (ES) groter dan 0.2. Benadrukt moet worden dat de effectgroottes *geen causale verbanden* aanduiden. Sterker nog, de beoordelaars hebben deze vragen pas ingevuld, nadat de OFI-vragen waren ingevuld. Dit om te voorkomen dat het holistische oordeel en het oordeel op correctheidsaspecten de beoordeling van de opbouw-, formulerings- en inhoudsaspecten zouden beïnvloeden. In Bijlage 12 wordt een overzicht gegeven van de gefixeerde effecten voor alle taken. De resultaten worden hieronder samengevat.

Algemene indruk

Bij de meeste nieuwe taken blijkt het onderscheid tussen *goed en voldoende* niet gerelateerd aan de algemene schrijfvaardigheid (de overall vaardigheid) of de taakvaardigheid. Alleen bij Taak 07 (pannenkoek) is er sprake van een klein effect m.b.t. zowel taak- als overall vaardigheid (resp. ES=0.23 en ES=0.27). Bij Taak 02 (Verboden boek) en Taak 12 (Frisdrank en Snoep) is er alleen sprake van een klein effect m.b.t. de taakvaardigheid (ES in beide gevallen 0.23) maar niet m.b.t. de overall vaardigheid. Over het algemeen is de overall vaardigheid en taakvaardigheid van leerlingen met teksten die de beoordelaar in zijn algemeenheid als *slecht* beoordeelt, lager dan die van leerlingen met teksten waarvan de algemene indruk als *onvoldoende* is beoordeeld, met effectgroottes tussen 1.10 en 1.42 op overall vaardigheid en 0.94 en 1.42 op taakvaardigheid. Ditzelfde geldt voor het onderscheid tussen *onvoldoende* en *voldoende* beoordeelde leerlingteksten: over het algemeen is de overall en taakvaardigheid bij de laatste groep leerlingen hoger. De effectgroottes zijn hier echter wat lager (tussen 0.52 en 0.90 voor overall vaardigheid en 0.53 en 0.86 voor taakvaardigheid).

Bij de ankers waren drie categorieën gedefinieerd: onvoldoende (0), voldoende (1) en goed (2). Ook voor de drie ankertaken is de relatie tussen *onvoldoende-voldoende* wat betreft de algemene indruk duidelijk aan de overall en taakvaardigheid gerelateerd, met middelgrote tot grote effectgroottes tussen 0.75 en 1.16 op overall vaardigheid en 0.71 en 1.11 op de taakvaardigheid. Het verschil tussen een *voldoende-goede* algemene indruk is beduidend minder sterk gerelateerd aan de vaardigheid, met effectgroottes tussen 0.14 (geen impact) en 0.31 (kleine impact) op de taakvaardigheid en effectgrootte van 0.22 (kleine impact) op de overall vaardigheid bij alle drie de ankertaken.

Hoeveelheid tekst

Bij de nieuwe taken diende de beoordelaar aan het eind van de beoordeling in te schatten hoeveel regels bodytekst de tekst bevatte. Dit gebeurde in vier categorieën: 1) 1-5 regels, 2) 6-10 regels, 3) 11-15 regels en 4) meer dan 15 regels, gescoord als 0, 1, 2 en 3.

¹⁴¹ De werkwijze was dezelfde bij de andere aspecten: tekstlengte, spelfouten, interpunctiefouten en niet-lopende constructies.

Het onderscheid tussen veel en iets minder regels (3 en 4) blijkt niet gerelateerd aan de vaardigheid overall of per taak. Alleen bij Taak 02 (Verboden boek - verhaal) is er sprake van een klein effect m.b.t. zowel taak- als overall vaardigheid (resp. $ES=0.30$ en $ES=0.39$). De hoeveelheid geschreven tekst blijkt met name bij het onderscheid tussen zeer kort en korte teksten (1-5 regels en 6-10 regels) gerelateerd aan schrijfvaardigheid (met grote effecten: lopend van 0.87 tot 1.51 bij overall vaardigheid en van 0.60 tot 1.42 bij taakvaardigheid – grootste effect bij T07, pannenkoekenrecept). De effectgroottes bij het onderscheid tussen kort en iets langer (6-10 regels versus 11-15 regels) variëren tussen 0.16 en 0.72 voor de overall vaardigheid en tussen 0.11 en 0.62 voor de taakvaardigheid. Bij Taak 09 (Naima) en 10 (speelveldje) is er geen sprake van impact: het feit of de tekst uit 6-10 regels of 11-15 regels bestaat hangt hier niet samen met de vastgestelde overall en taakvaardigheid. Bij Taak 06 (vlinder) en taak 11 (korte zomervakantie) is er geen sprake van impact wanneer gekeken wordt naar de relatie tussen korte en iets langere teksten (2 en 3) en de taakvaardigheid en van een kleine impact op de overall vaardigheid (resp. 0.24 en .27).

Correctheid

Interpunctie

Bij de nieuwe taken diende de beoordelaar aan het eind van de beoordeling in te schatten bij hoeveel van de zinnen een hoofdletter ontbrak en/of er geen punt aan het eind van de zin stond. Er waren drie niveaus: 1) bij meer dan de helft van de zinnen, 2) bij minder dan de helft van de zinnen en 3) bij geen enkele zin, gescoord als 0, 1 en 2. De effectgroottes bij het onderscheid tussen 2 en 3 (bij minder dan de helft van de zinnen en geen enkele zin) waren laag en er is geen impact m.b.t. overall en taakvaardigheid. Dit is wel het geval bij het onderscheid tussen meer/minder dan de helft van de zinnen. Middelgrote effectgroottes (tussen 0.50 en 0.74) worden gevonden m.b.t. de overall vaardigheid, terwijl er sprake is van kleine tot middelgrote effectgroottes (tussen 0.32 en 0.61) m.b.t. de taakvaardigheid.

Grammaticale correctheid

Bij de nieuwe taken diende de beoordelaar aan het eind van de beoordeling het aantal niet-lopende zinsconstructies in te schatten. Er waren drie niveaus: 1) bij meer dan de helft van de zinnen, 2) bij minder dan de helft van de zinnen en 3) bij geen enkele zin, gescoord als 0, 1 en 2. Een vergelijkbaar beeld ontstaat als bij interpunctie: met name het onderscheid meer/minder dan de helft van de zinnen is gerelateerd aan de vaardigheid. Effectgroottes lopen van 0.33 tot 1.05 bij overall vaardigheid en van 0.47 (T12 – frisdrank en snoep) tot 0.87 (T2 – verboden boek) bij taakvaardigheid. Het onderscheid tussen minder dan de helft en geen enkele zin met niet-lopende zinsconstructies heeft grotendeels geen impact. Alleen bij taak 7 (pannenkoekenrecept) is er een kleine impact op overall en taakvaardigheid en bij taak 5 (nepnieuws) en 6 (vlinder) op de overall vaardigheid.

Spelfouten

Bij de nieuwe taken diende de beoordelaar aan het eind van de beoordeling het aantal spelfouten in te schatten. Er waren wederom drie niveaus: 1) bij meer dan de helft van de zinnen, 2) bij minder dan de helft van de zinnen en 3) bij geen enkele zin, gescoord als 0, 1 en 2. Hetzelfde beeld komt weer naar voren: met name het onderscheid tussen meer/minder dan de helft van de zinnen hangt samen met de vaardigheid. Er worden middelgrote effectgroottes gevonden m.b.t. overall vaardigheid (lopend van ES 0.62 tot 0.90) en m.b.t. de taakvaardigheid (lopend van 0.38 tot 0.64). Het onderscheid tussen

minder dan de helft en geen enkele zin met niet-lopende zinsconstructies heeft grotendeels geen impact. Alleen bij Taak 07 (pannenkoekenrecept) is er een kleine impact t.o.v. overall vaardigheid.

Bij de drie ankertaken is aan het eind van de taak gevraagd naar een inschatting van het aantal schrijffouten, onderverdeeld in drie niveaus: 1) veel schrijffouten, 2) niet veel, maar ook niet weinig schrijffouten ('matig') en 3) weinig of geen schrijffouten, gescoord als 0, 1 en 2. Zowel het onderscheid tussen veel en 'matig' als tussen 'matig' en weinig/geen is gerelateerd aan de overall en taakvaardigheid. Bij het onderscheid tussen veel en matig lopen de effectgroottes van 0.48 tot 0.69 t.o.v. overall vaardigheid en van 0.32 tot 0.55 t.o.v. de taakvaardigheid. De effectgroottes voor het verschil tussen respectievelijk 'veel en matig' en 'matig en veel' schrijffouten zijn lager, lopend van 0.29 tot 0.37 t.o.v. overall vaardigheid en van 0.19 tot 0.28 t.o.v. taakvaardigheid, bij T03 is geen spraken van een impact op taakvaardigheid (ES=0.19).

Benadrukt moet worden dat de in deze paragraaf besproken effectgroottes niet impliceren dat de taalverzorgingsaspecten de taakspecifieke oordelen hebben beïnvloed. Immers, de taakspecifieke oordelen gingen over andere elementen dan taalverzorging, namelijk over inhouds-, opbouw- en formuleringsaspecten. Bovendien werden de taalverzorgingsaspecten pas ná de beoordeling op taakspecifieke OFI-aspecten gegeven. De meest aannemelijke interpretatie is dat taalverzorging en communicatieve kwaliteit daadwerkelijk gerelateerd zijn.

8.6 Trendvergelijking

Om na te gaan hoe de schrijfvaardigheid van de leerlingen aan het eind van het basisonderwijs zich door de tijd ontwikkelt, is een trendvergelijking uitgevoerd waarin de schrijffprestaties van bo- en sbo-leerlingen in 2009 zijn vergeleken met die van bo- en sbo-leerlingen in 2019. Met dit doel waren drie van de twaalf taken dezelfde als in 2009 en zijn zij op dezelfde wijze beoordeeld. Dit zijn de zogenaamde ankertaken.

Ankertaken en data

Voor de jaarvergelijking bij de schrijftaken kon gebruik gemaakt worden van een drietal ankertaken. Dat betreft de volgende taken (taaknummers zijn de taaknummers zoals gebruikt in 2019):

- Opdracht Wim (taak 1): Vrij schrijven
- Opdracht Unicef (taak 3): Verslagen, werkstukken, samenvattingen, artikelen
- Opdracht Poes (taak 8): Formulieren invullen, berichten, advertenties en aantekeningen

Deze taken zijn in de eerdere peiling ook afgenomen in groep 8. Ze vertegenwoordigen een spreiding over de verschillende taaktypen zoals gedefinieerd in het Referentiekader Taal en Rekenen, zoals hierboven is weergegeven. Bij de keuze van deze ankertaken is er ook op gelet dat de taken niet duidelijk gedateerd waren. Alle beoordelingsaspecten (I, F en O) worden bij deze taken meegenomen. Bij de selectie van de taken is ook rekening gehouden met de beoordelingsovereenstemming. Van de mogelijk te kiezen taken hadden deze taken in de vorige peiling een zo hoog mogelijke overeenstemming voor het totaal van de beoordelingsaspecten.

Wat betreft de betrouwbaarheid van de ankertaken, in vergelijking met de nieuwe taken die afgenomen zijn in 2019, is te zien dat de ankertaken en de nieuwe taken ongeveer even betrouwbaar

zijn. Dat is het geval bij de inter-item-betrouwbaarheid (λ -2) van de OFI-taken met gepolytomiseerde scores en bij de interbeoordelaarovereenstemming.

De ankertaken zijn beide jaren op dezelfde manier beoordeeld. Hierdoor verschillen bij de ankertaken de meer globale beoordelingsaspecten enigszins van die in de nieuwe taken. Bij de subschalen O, F en I heeft het geen impact. Bij het omzetten van de observaties - zoals ingevuld op de beoordelingsformulieren - naar scores is voor de nieuwe en de oude data de procedure van 2019 toegepast. Dit betekent dat de hier gepresenteerde resultaten iets van de in Kuhlemeijer e.a. (2013) gerapporteerde resultaten kunnen verschillen, maar dat de resultaten over de jaren goed vergelijkbaar zijn. Hiervoor is gekozen omdat de gekozen rapportage-eenheden (schalen) in 2019 toch al anders zijn dan die van de oude peiling en daarmee hoe dan ook enigszins afwijken van het vorige peilingsonderzoek.

De scoreregels zijn toegepast op de datasets van het vorige peilingsonderzoek¹⁴²:

- Taak 01 Taak 16 Poes.xls
- Taak 04 Taak 27 Wim.xls
- Taak 18 Unicef.xls

Ter controle zijn de scoreregels ook opnieuw op de dataset van 2019 toegepast. Daar komen exact dezelfde scores uit als waarover in de analyses van 2019 gerapporteerd is. De oude en de nieuwe datasets zijn per taak samengevoegd. Dit levert een drietal datasets op:

- PS.JaarVgl.TK01.sav (Wim)
- PS.JaarVgl.TK03.sav (Unicef)
- PS.JaarVgl.TK08.sav (Poes/Poekie)

In deze datasets staan alle gegevens die verzameld zijn over deze taken in beide jaren. Zowel de observaties als de scores op de opgaven zijn hier te vinden. Naast deze scores zijn de volgende zaken in deze dataset gegeven:

- JAAR 2 niveaus: 1=2019 (Nieuwe peiling); 2=2006 (Oude peiling (2009))
- Taak Per dataset een waarde (1 bij Wim, 3 bij Unicef en 8 bij Poes)
- Recordtk Recordnummer binnen taak (2019; 1- (extra identificatie)
- Schoolnummer School ID (2019: 4 cijfers; 2009: 5 cijfers)
- klas_leerling Leerling ID in school (3 cijfers, tenzij het eerste een 0 is); 1^e cijfer is een klas ID waarbij deze in 2009 altijd een 8 is om leerjaar 8 aan te geven
- LLnr Uniek leerlingnummer (schoolnummer *1000 + klas_leerling)
- BeoNr Beoordelaars ID
- Nbeo Aantal beoordelingen bij deze leerling
- RandomNbeo Bij meer dan 1 beoordelaar, beoordelaar random aangemerkt als 1^e of 2^e beoordelaar (bij Unicef in 2009 ook 3^e t/m 6^e beoordelaar mogelijk)
- Random Randomnummer om te bepalen welk RandomNbeo de beoordelaar krijgt
- Nvragen Aantal vragen in de beoordelingslijst (25 bij T01; 30 bij T03 en 29 bij T08)
- Nscores Aantal opgaven met een beoordeling (aantal vragen verminderd met het aantal overgeslagen beoordelingen)
- Pct_Scores Percentage beoordeelde items

¹⁴² Het betreft hier de beschikbaar gestelde bestandsnamen. In 2009 is de peiling in groep 5, groep 8 en bij schoolverlaters afgenomen. De taken Poes en Wim zijn in groep 5 en 8 afgenomen en hadden daar een andere taaknummering. Dit verklaart de dubbele taaknummering in de bestandsnaam. In een later stadium zijn SBO-data toegevoegd. Dat betreft de datasets (cito1004, cito1018 en cito 1001).

- BO_vs_SBO Leerling is BO groep 8 leerling (in 2009: leerlingnummer tussen 800-999; in 2019, als school nummer tussen 1000 en 3999)
- LJq Leerjaar: 0=SBO, 8= BO
- SEL_LJ Selectie op basis van leerjaar: 1=Ja (laatste leerjaar, BO8, of SBO); 0=Nee
- ONOFF Verdere selectie

In de dataset ontbreken de gegevens voor andere achtergrondvariabelen. Dat betreft bijvoorbeeld schoolvariabelen zoals percentage gewichtenleerlingen of regio, of leerlingvariabelen zoals geslacht en leeftijd of leertijd (regulier, zittenblijver, vervroegd). Als deze variabelen binnen een jaar niet volgens de voor dat jaar geldende populatieverdeling verdeeld zijn, kan daar niet meer voor gewogen worden. Dit kan de gevonden resultaten enigszins vertekenen.

Er zijn drie datasets gemaakt, met daarin zowel de bo als sbo data:

- PS.JaarVgl.BO8.TK01.sav (Wim)
- PS.JaarVgl.BO8.TK03.sav (Unicef)
- PS.JaarVgl.BO8.TK01.sav (Poekie)

In de datasets PS.JaarVgl.TK\$.sav zitten een paar records die als BO8 in 2009 aangegeven zijn, maar dat zijn allen leerlingen met een leerlingnummer 999 waarvan niet op zich vast te stellen was dat deze BO8-leerlingen betroffen¹⁴³. Deze zijn verwijderd.

8.6.1 Vergelijking data van de oude en de nieuwe peiling

Vergelijking aantal records, leerlingen en beoordelaars per leerling

Per taak is onderzocht hoeveel observaties er zijn in ieder jaar. Een ingevuld beoordelingsformulier wordt beschouwd als een observatie. Doordat de schrijfproducten van veel leerlingen door meer dan een beoordelaar zijn beoordeeld zijn is het aantal observaties aanzienlijk groter dan het aantal leerlingen. De aantallen observaties en aantallen leerlingen per taak zijn gegeven in Tabellen 8.31a (bo) en 8.31b (sbo).

Tabel 8.31a: aantal observaties en aantal leerlingen per taak per jaar, bo

taak	Aantal beoordelingen			taak	Aantal leerlingen		
	2019	2009	Totaal		2019	2009	Totaal
T01 Wim	1008	313	1321	T01 Wim	890	162	1052
T03 Unicef	1012	717	1729	T03 Unicef	904	333	1237
T08 Poekie	805	383	1188	T0 8 Poekie	718	202	920

¹⁴³ Bij de weggevallen leerlingen met leerlingnummer 999 was het casenummer 99999 of van een enkele school waar geen andere leerlingen van te vinden waren. Deze scholen zijn niet te controleren als werkelijk scholen met BO8 leerlingen. Dit betrof drie records bij T01, drie records bij T03 en twee leerlingen bij T08. Er was een leerling die op een sbo school zat die eerder onterecht als BO8 leerling gezien was (T01).

Tabel 8.31b: aantal observaties en aantal leerlingen per taak per jaar, sbo

taak	Aantal beoordelingen			taak	Aantal leerlingen		
	2019	2009	Totaal		2019	2009	Totaal
T01 Wim	163	178	341	T01 Wim	145	92	237
T03 Unicef	183	166	349	Tk03 Unicef	162	85	247
T08 Poekie	157	178	335	T0 Poekie	137	90	227

Uit Tabel 8.31 blijkt dat het aantal beoordelingen per leerling in 2009 groter was dan in 2019. In 2019 is ruim 10% van de leerlingen dubbel beoordeeld (gemiddeld 12,9%, weinig verschil tussen groep 8 en SBO), terwijl er in 2009 gemiddeld iedere leerling ongeveer twee beoordelingen had (bij groep 8 iets meer en bij SBO iets minder). In 2009 had gemiddeld één op de 14 leerlingen (7%) slechts een beoordeling, terwijl dat in 2019 ongeveer zeven van de acht leerlingen betrof.

Het aantal unieke leerlingen per taak was in 2019 duidelijk groter dan in 2009. Dat is met name in het BO het geval met over de drie taken heen 3,6 keer zo veel unieke leerlingen in 2019 dan in 2009. Bij het SBO is die verhouding lager, met gemiddeld 1,7 keer zo veel unieke leerlingen.

Tot slot is er gekeken naar het aantal unieke beoordelaars per taak per jaar. Dit aantal lag in 2009 hoger dan in 2019. Doordat het aantal beoordelingen in 2019 wel hoger lag, is zeker ook het gemiddelde aantal schrijfproducten per beoordelaar voor de ankertaken in 2019 hoger dan in 2009. Daarbij valt nog wel op te merken dat in 2009 niet alleen groep 8 en sbo gemeten waren, maar ook leerlingen in groep 5, 6 en 7.

Vergelijking ontbrekende waarden en de basisvragen (vraag 1 tot en met 3)

Niet alle beoordelingspunten zijn bij alle beoordelaars gebruikt. Een deel hiervan is door voorwaardelijke beoordelingen: of een beoordeling gegeven wordt, is afhankelijk van het antwoord op de vraag ervoor. Een voorbeeld hiervan is Taak 08, item 11, die hier is gepresenteerd in combinatie met item 10:

10. Vermeldt de leerling behalve de kenmerken waarnaar hierboven reeds gevraagd is, nog andere uiterlijke kenmerken, dan wel karakter- of gedragseigenschappen van Poekie?

Ja Nee

Als u deze vraag met Nee heeft beantwoord, kunt u vraag 11 overslaan.

11. Maken deze extra kenmerken en eigenschappen het makkelijker om Poekie te herkennen?

Ja

Nee, ze maken het juist moeilijker

Nee, ze zijn in dit opzicht niet relevant

Ja en nee

Dit betreft een viertal vragen in deze ankertaken, te weten vraag 7 uit Taak 01 (T01_v07) waar 30% van de mogelijke waarnemingen ontbreekt, vraag 15 uit Taak 03 (T03_v15) met 57% ontbrekend, en vraag 11 en 20 uit Taak 08 met 19% en 11% ontbrekend. Om die reden is bij veel van de beoordelingen

niet 100% van de beoordelingspunten beoordeeld. Deze vragen zijn omgevormd tot polytome items. Ook voor 2009 is dezelfde werkwijze gebruikt.

Daarnaast zijn er schrijfproducten die verder niet beoordeeld zijn door de beoordelaars op basis van de initiële vragen over de leesbaarheid (V1), of de tekst af is (V2) en of de tekst over het beoogde onderwerp gaat (V3). In Tabel 8.32 zijn de percentages positieve antwoorden per vraag gegeven voor BO en SBO en per jaar.

Tabel 8.32: Percentages leerlingen per taak per jaar voor de drie basisvragen

Lij	JAAR	V01: % leesbaar			V02: % tekst af			V03: % beoogd onderwerp		
		Taak 1	Taak 3	Taak 8	Taak 1	Taak 3	Taak 8	Taak 1	Taak 3	Taak 8
sbo	2019	99%	99%	99%	96%	96%	96%	93%	94%	95%
	2009	99%	98%	94%	93%	91%	94%	93%	81%	90%
bo	2019	99%	99%	99%	97%	97%	98%	96%	96%	92%
	2009	100%	99%	99%	92%	95%	93%	89%	91%	92%
Totaal	2019	99%	99%	99%	97%	97%	98%	95%	95%	92%
	2009	99%	99%	98%	91%	95%	94%	86%	89%	92%
Totaal		99%	99%	99%	94%	96%	96%	91%	92%	92%

Bij vraag 1 is het antwoord veruit in de meeste gevallen positief. Dat is in mindere mate het geval bij vraag 2 en nog minder bij vraag 3, waarbij de verschillen over taken niet heel groot zijn. In 2019 waren de antwoorden duidelijk vaker positief dan in 2009 bij vraag 2 en vraag 3, al was er bij vraag 3 bij taak 8 geen verschil tussen de beide jaren. Wanneer we de antwoorden bij bo en sbo nu met elkaar vergelijken, dan zien we bij deze vragen (2 en 3) dat het percentage positieve antwoorden weinig van elkaar verschilde in 2019, ook al waren er iets meer positieve antwoorden bij de bo-teksten. In 2009 waren de verschillen in positieve antwoorden tussen bo en sbo bij vraag 2 en 3 grilliger. Afhankelijk van vraag en taak was het percentage positieve antwoorden hoger bij de bo dan wel de sbo-teksten.

Als we kijken hoe het zit met combinaties van vragen, willen we weten hoeveel procent van de werken op alle drie de vragen met een "ja" beoordeeld zijn, omdat dit schrijfproducten zijn die zeker te beoordelen zouden moeten zijn. Er is ook te beargumenteren dat eventueel een tekst niet af hoeft te zijn om toch te kunnen worden beoordeeld. Om die reden is ook gekeken naar de leerlingwerken die alleen op vraag 1 en 3 beiden positief beoordeeld zijn. De resultaten staan in Tabel 8.33.

Tabel 8.33: Percentages schrijfproducten die zeker beoordeelbaar zouden zijn

% records	Taak 1				Taak 3				Taak 8			
	sbo		bo		sbo		bo		sbo		bo	
Ja op alles	2019	2009	2019	2009	2019	2009	2019	2009	2019	2009	2019	2009
v1 t/m v3	90.8%	89.3%	93.5%	84.4%	91.8%	77.7%	94.6%	88.2%	92.4%	83.2%	91.7%	88.8%
v1 & v3	92.6%	92.7%	95.7%	88.5%	94.0%	80.1%	95.6%	90.8%	94.9%	86.5%	91.8%	91.9%

Zoals te verwachten, ook op basis van de resultaten zoals gegeven in Tabel 8.32, is te zien dat meer schrijfproducten in 2019 beoordeelbaar werden geacht dan in 2009. Dit verschil geldt zowel voor bo als sbo. Deze verschillen worden altijd gevonden als er naar vraag 1 tot en met vraag 3 gekeken wordt. Als alleen vraag 1 en 3 bekeken worden, dan valt het verschil tussen 2019 en 2009 weg bij Taak 01 sbo en Taak 08 bo. Of het percentage hoger is bij sbo of bo hangt af van de taak en van het jaar. Alleen bij Taak 03 zien we dat meer records op alle vragen (V1 & V3 en/of V3) een 1 scoren bij bo dan bij sbo.

Als een taak niet te beoordelen is heeft dat impact op de percentages te beoordelen punten. Beoordelaars die een tekst niet kunnen lezen kunnen deze ook niet goed beoordelen. Ook bij teksten die over een ander onderwerp gaan dan de bedoeling is, zijn moeilijk vergelijkbaar te beoordelen gelijk aan opgaven die wel de bedoelde inhoud hebben. In die gevallen kunnen de beoordelaars meer beoordelingspunten overslaan.

We kunnen ook direct naar de percentages ontbrekende scores bij de verschillende taken kijken in 2009 en 2019. Deze zijn gegeven in Tabel 8.34 waarbij er vier categorieën onderscheiden worden, te weten:

- (Vrijwel) geen beoordeling gegeven (0 tot minder dan 10% beoordeeld).
- Zeer weinig beoordeeld (10% tot minder dan de helft – *in de data tot en met 38%*).
- Beperkte beoordeling (de helft –*in de data vanaf 62%*– tot 90% beoordeeld).¹⁴⁴
- (Vrijwel) alles beoordeeld (90 tot en met 100% beoordeeld).

Tabel 8.34: Percentage ontbrekende waarden per taak per jaar

Percentage records			Percentage gescoorde items				Cumulatief		
Taak	LJ	JAAR	0 - <10	10 - <50	50 - <90	90 - 100	<10	<50	<90
1	sbo	2019	0.6	0.6	0.6	98.2	0.6	1.2	1.8
		2009	0.6	0.0	0.0	99.4	0.6	0.6	0.6
	bo	2019	0.6	0.0	0.3	99.1	0.6	0.6	0.9
		2009	0.3	0.3	1.3	98.1	0.3	0.6	1.9
3	sbo	2019	2.2	1.1	0.5	96.2	2.2	3.3	3.8
		2009	3.0	0.0	0.0	97.0	3.0	3.0	3.0
	bo	2019	1.3	0.5	0.2	98.0	1.3	1.8	2.0
		2009	1.7	0.0	0.3	98.0	1.7	1.7	2.0
8	sbo	2019	1.3	1.9	0.0	96.8	1.3	3.2	3.2
		2009	1.7	1.1	0.0	97.2	1.7	2.8	2.8
	bo	2019	1.2	0.4	2.2	96.1	1.2	1.6	3.9
		2009	1.6	1.6	0.8	96.1	1.6	3.1	3.9

¹⁴⁴ Er komen geen waarden tussen 38% en 62% voor in de data.

Het valt hierbij op dat het vooral van de taak afhangt of er meer of minder beoordelingen overgeslagen zijn. De verschillen tussen sbo en bo 8 en tussen 2019 en 2009 zijn niet groot.

Vergelijking op opgave niveau (bo)

Om een indruk te krijgen hoe de opgaven over de jaren functioneren, zijn voor het bo per taak de verdelingen van de scores over de twee jaren vergeleken. In de omschrijving is gewerkt met beoordelingen in plaats van leerlingen als eenheid. Dat betekent wel dat bij de data van 2009 de meeste leerlingen dubbel geteld worden (vanwege de grote hoeveelheid dubbele beoordelingen), en dat dit bij 2019 voor een deel van de leerlingen geldt. Bij de vergelijking van de resultaten in 2009 en 2019 op somscore niveau (O, F, I afzonderlijk en de totale OFI-schaal) zal echter de vergelijking op het niveau van de leerling plaatsvinden (zie later in deze paragraaf). De functie van de analyses in deze paragraaf is echter om een indruk te krijgen per item om te zien of hier opmerkelijke resultaten zijn die tot aanpassingen van de somscores zouden kunnen leiden. Hierbij willen we zoveel mogelijk data gebruiken, en de originele scorecategorieën van de beoordelingspunten behouden.

Tabel 8.35: O, F en I items bij taak 1 (bo)

			Gemiddelde scores						P'-waarden		
			Alle beoordelingen			Mits 90% of meer ¹⁴⁵			Mits 90% of meer		
JAAR	schaal	MaxSc	2019	2009	V	2019	2009	V	2019	2009	V
T01_i04_O	o	1	.86	.88	-.02	.87	.89	-.02	87	89	-2
T01_i05_C	o	1	.95	.85	.10	.96	.86	.10	96	86	10
T01_i06_O	o	1	.22	.04	.18	.22	.04	.18	22	4	18
T01_i07_O	o	1	.19	.02	.17	.19	.02	.17	19	2	17
T01_i08_O	o	1	.21	.12	.09	.21	.13	.08	21	13	8
T01_i09_O	o	1	.85	.71	.14	.85	.72	.13	85	72	13
T01_i10_I	i	1	.44	.44	.00	.44	.44	.00	44	44	0
T01_i11_I	i	1	.89	.83	.06	.90	.84	.06	90	84	6
T01_i12_I	i	1	.90	.82	.08	.90	.82	.08	90	82	8
T01_i13_I	i	2	1.80	1.64	.16	1.81	1.65	.16	91	83	8
T01_i14_F	f	1	.50	.32	.18	.50	.32	.18	50	32	18
T01_i15_F	f	1	.24	.10	.14	.24	.10	.14	24	10	14
T01_i16_F	f	1	.52	.36	.16	.52	.36	.16	52	36	16
T01_i17_I	i	1	.50	.27	.23	.50	.27	.23	50	27	23
T01_i18_F	f	2	.43	.32	.11	.43	.32	.11	22	16	6
T01_i19_O	o	2	1.40	1.09	.31	1.41	1.11	.30	71	56	15
T01_i20_O	o	1	.20	.22	-.02	.21	.22	-.01	21	22	-1
T01_i21_F	f	1	.32	.21	.11	.33	.22	.11	33	22	11
T01_i22_F	f	1	.21	.27	-.06	.21	.27	-.06	21	27	-6

Bij Taak 1 is te zien dat slechts bij drie opgaven de leerlingen in 2019 minder goed presteren dan in 2009 (zie rode markeringen). De meeste opgaven worden beter gedaan in 2019. Als we kijken naar de score dan is het grootste verschil bij item T01_i19_O, waarbij de leerlingen 0.31 punt beter scoren in

¹⁴⁵ Het gaat hier om de taken waarbij 90% of meer is beoordeeld.

2019. Bij deze opgave zijn in totaal twee punten te verkrijgen. Als we naar de p-waarde kijken is het verschil het grootst bij T01_i17_I met een p-waarde verschil van 23 (p=50 in 2019, en p=27 in 2009).

Tabel 8.36: O, F en I items bij taak 3 (bo)

			Gemiddelde scores						P'-waarden		
			Alle beoordelingen			Mits 90% of meer			Mits 90% of meer		
JAAR	schaal	MaxSc	2019	2009	V	2019	2009	V	2019	2009	V
T03_i04_I	i	2	1.68	1.75	-.07	1.71	1.78	-.07	86	89	-4
T03_i05_I	i	1	.67	.64	.03	.69	.65	.04	69	65	4
T03_i06_I	i	2	1.65	1.64	.01	1.68	1.67	.01	84	84	1
T03_i07_I	i	1	.86	.8	.06	.88	.81	.07	88	81	7
T03_i08_I	i	1	.76	.8	-.04	.78	.82	-.04	78	82	-4
T03_i09_I	i	1	.71	.72	-.01	.72	.73	-.01	72	73	-1
T03_i10_I	i	1	.75	.75	.00	.76	.76	.00	76	76	0
T03_i11_I	i	3	.73	.74	-.01	.74	.75	-.01	25	25	0
T03_i12_I	i	1	.61	.69	-.08	.62	.7	-.08	62	70	-8
T03_i13_I	i	2	.75	.81	-.06	.76	.82	-.06	38	41	-3
T03_i14_I	i	1	.19	.12	.07	.19	.12	.07	19	12	7
T03_i15_I	i	2	.22	.13	.09	.22	.13	.09	11	7	5
T03_i16_O	o	1	.7	.64	.06	.72	.66	.06	72	66	6
T03_i17_O	o	1	.89	.85	.04	.9	.87	.03	90	87	3
T03_i18_O	o	3	1.68	1.99	-.31	1.71	2.03	-.32	57	68	-11
T03_i19_O	o	1	.04	.01	.03	.04	.01	.03	4	1	3
T03_i20_O	o	1	.06	.08	-.02	.06	.08	-.02	6	8	-2
T03_i21_O	o	1	.68	.71	-.03	.69	.72	-.03	69	72	-3
T03_i22_F	f	1	.28	.21	.07	.29	.22	.07	29	22	7
T03_i23_O	o	1	.17	.17	.00	.18	.18	.00	18	18	0
T03_i24_F	f	2	1.45	1.62	-.17	1.48	1.65	-.17	74	83	-9
T03_i25_F	f	1	.16	.11	.05	.16	.11	.05	16	11	5
T03_i26_F	f	1	.01	0	.01	.01	0	.01	1	0	1
T03_i27_F	f	1	.19	.18	.01	.19	.19	.00	19	19	0

Ook bij Taak 3 wordt het merendeel van de items in 2019 beter gescoord dan in 2009. Bij de opgaven die echter slechter gescoord worden, is de daling wel erg sterk. De meest opvallende opgaven zijn hier bij drie-punts-item T03_i18_O en twee-punts-item T03_i24_F. De daling is -.31 (p-waarde daling van 11 punten) en -.17 (p-waarde daling 9 punten). De grootste stijging in p-waarden is 7 p-waarde punten (item T03_v14 en item T03_v22). De grootste stijging in punten is te vinden bij item T03_v15 (0.09 punt hoger in 2019).

Bij taak 8 (Tabel 8.37) zien we een daling bij iets meer dan de helft van de opgaven (12 van de 23). De grootste daling in punten is te zien bij drie-punts-item T08_i04_I (daling van .45), en de grootste daling in p-waarden bij item T08_i13_I (39 punten: van 73 in 2009 naar 34 in 2019 wat echt wel een erg sterke daling is). De sterkste stijging is te vinden bij één-punts-items T08_i12_I, T08_i14_I en

T08_i21_I (ongeveer 0.15). Opvallend is dat zowel de grootste positieve als de grootste negatieve impact beiden bij inhoudsitems is.

Deze verschillen in p-waarden bij deze taken zijn geen aanleiding om opgaven te verwijderen uit de schaal. Wel zijn de verschillen tussen de taken interessant. Taak 1 wordt in 2019 beter gedaan (ongeveer 2.5 punt op 22 mogelijke punten van 19 opgaven), taak 3 wordt vergelijkbaar gemaakt (daling van 0.4 punt op 33 mogelijke punten van 24 opgaven) en taak 8 wordt slechter gemaakt (2.2 punt slechter met op 33 mogelijke punten 23 opgaven). Deze verschillen per taak per schaalonderdeel zijn ook weergegeven in Tabel 8.38.

Tabel 8.37: O, F en I items bij taak 8 (bo)

			Gemiddelde scores						P'-waarden		
			Alle beoordelingen			Mits 90% of meer			Mits 90% of meer		
JAAR	schaal	MaxSc	2019	2009	V	2019	2009	V	2019	2009	V
T08_i04_I	i	2	.65	1.10	-0.45	.67	1.13	-0.46	34	57	-23
T08_i05_I	i	2	1.05	1.39	-0.34	1.07	1.44	-0.37	54	72	-19
T08_i06_I	i	1	.56	.70	-0.14	.57	.72	-0.15	57	72	-15
T08_i07_I	i	1	.29	.28	.01	.30	.29	.01	30	29	1
T08_i08_I	i	1	.06	.14	-0.08	.06	.14	-0.08	6	14	-8
T08_i09_I	i	2	.31	.30	.01	.32	.31	.01	16	16	1
T08_i10_I	i	1	.56	.51	.05	.58	.53	.05	58	53	5
T08_i11_I	i	3	1.17	1.03	.14	1.20	1.07	.13	40	36	4
T08_i12_I	i	1	.70	.54	.16	.71	.56	.15	71	56	15
T08_i13_I	i	1	.33	.71	-0.38	.34	.73	-0.39	34	73	-39
T08_i14_I	i	1	.95	.79	.16	.97	.82	.15	97	82	15
T08_i15_O	o	1	.89	.89	-0.00	.91	.92	-0.01	91	92	-1
T08_i16_O	o	1	.28	.53	-0.25	.28	.55	-0.27	28	55	-27
T08_i17_O	o	1	.38	.32	.06	.39	.34	.05	39	34	5
T08_i18_O	o	1	.26	.41	-0.15	.27	.42	-0.15	27	42	-15
T08_i19_I	i	1	.46	.60	-0.14	.47	.63	-0.16	47	63	-16
T08_i20_I	i	2	.38	.77	-0.39	.39	.80	-0.41	20	40	-21
T08_i21_I	i	1	.96	.80	.16	.97	.84	.13	97	84	13
T08_i22_I	i	1	.78	.72	.06	.79	.75	.04	79	75	4
T08_i23_F	f	1	.32	.31	.01	.32	.32	.00	32	32	0
T08_i24_F	f	1	.08	.27	-0.19	.09	.27	-0.18	9	27	-18
T08_i25_O	o	1	.23	.27	-0.04	.23	.28	-0.05	23	28	-5
T08_i26_F	f	1	.18	.16	.02	.18	.16	.02	18	16	2

In totaal zou dat neerkomen op een resultaat dat aan zou geven dat er geen echte verandering over de jaren te vinden is: over de drie taken is een totaalscore-daling van 0.07 te zien, wat een minieme daling van ongeveer 0.001 per item is. Als we het per taak bekijken dan zien we per item bij Taak 01 een stijging van 0.112 punt en bij taak 3 en 8 dalingen van 0.011 punt per item (T03) en 0.074 (T08). Gemiddeld zou dat - ongewogen voor het aantal items per taak - juist een zeer kleine stijging

betekenen van 0.026 per punt. Al met al zou het resultaat dan zijn dat er geen verschil tussen 2009 en 2019 is.

De verschillen zoals die over de taken gezien worden, zien we niet zo over de schalen O, F en I (zie Tabel 8.38). Het stijgen, ongeveer gelijk blijven en dalen is binnen de afzonderlijke taken vergelijkbaar. Per taak zijn er echter interessante verschillen te vinden. Om die reden worden de jaarverschillen verder ook per taak onderzocht en niet over het geheel. Als we dat laatste zouden doen, weten we dat het gevonden effect niet echt van 0 zou verschillen.

Tabel 8.38: Verschillen per taak per schaal (bo)

		Gemiddelde scores						P ¹ -waarden		
		Alle beoordelingen			Mits 90% of meer			Mits 90% of meer		
Schaal	Taak	2019	2009	V	2019	2009	V	2019	2009	V
O	1	0.56	0.44	0.12	0.57	0.45	0.12	46.5	36.8	9.7
F	1	0.45	0.35	0.11	0.46	0.35	0.11	42.5	32.7	9.8
I	1	0.91	0.80	0.11	0.91	0.80	0.11	72.9	63.9	9.0
OFI	1	0.61	0.50	0.11	0.62	0.51	0.11	52.0	42.4	9.6
O	3	0.42	0.42	-0.01	0.43	0.43	-0.01	45.1	45.7	-0.5
F	3	0.80	0.80	0.00	0.81	0.81	0.00	27.8	26.9	0.9
I	3	0.60	0.64	-0.03	0.61	0.65	-0.04	58.9	58.7	0.3
OFI	3	0.66	0.67	-0.01	0.67	0.69	-0.01	48.4	48.3	0.2
O	8	0.19	0.25	-0.05	0.20	0.25	-0.05	41.6	50.2	-8.6
F	8	0.61	0.69	-0.08	0.63	0.72	-0.09	19.7	25.0	-5.3
I	8	0.41	0.48	-0.08	0.42	0.50	-0.09	49.2	54.7	-5.5
OFI	8	0.51	0.59	-0.07	0.53	0.61	-0.08	43.7	49.9	-6.1
O	1	0.56	0.44	0.12	0.57	0.45	0.12	46.5	36.8	9.7
O	3	0.42	0.42	-0.01	0.43	0.43	-0.01	45.1	45.7	-0.5
O	8	0.19	0.25	-0.05	0.20	0.25	-0.05	41.6	50.2	-8.6
F	1	0.45	0.35	0.11	0.46	0.35	0.11	42.5	32.7	9.8
F	3	0.80	0.80	0.00	0.81	0.81	0.00	27.8	26.9	0.9
F	8	0.61	0.69	-0.08	0.63	0.72	-0.09	19.7	25.0	-5.3
I	1	0.91	0.80	0.11	0.91	0.80	0.11	72.9	63.9	9.0
I	3	0.60	0.64	-0.03	0.61	0.65	-0.04	58.9	58.7	0.3
I	8	0.41	0.48	-0.08	0.42	0.50	-0.09	49.2	54.7	-5.5
OFI	1	0.61	0.50	0.11	0.62	0.51	0.11	52.0	42.4	9.6
OFI	3	0.66	0.67	-0.01	0.67	0.69	-0.01	48.4	48.3	0.2
OFI	8	0.51	0.59	-0.07	0.53	0.61	-0.08	43.7	49.9	-6.1

8.6.1 Vergelijking van scores per taak (bo en sbo)

Om een indruk te krijgen hoe de opgaven over de jaren functioneren zijn per taak de verdelingen van de scores over de twee jaren vergeleken. In de omschrijving is hiervoor gewerkt met beoordelingen in plaats van leerlingen als eenheid.

Bij de definitieve vergelijking is het belangrijk dat leerlingen slechts eenmaal meegenomen worden in de analyses. Wanneer de analyses uitgevoerd worden op beoordelingsniveau dan heeft dat impact op de spreiding van de score en de veronderstelde meetprecisie. De scores zijn dan namelijk niet onafhankelijk van elkaar, wat met name bij de gegevens van 2009 een grote rol speelt: in 2019 heeft gemiddeld 1 op de 8 twee beoordelingen, terwijl in 2009 het gemiddeld aantal beoordelingen per leerling op 2 lag.

Er zijn verschillende mogelijkheden om een enkele beoordeling per leerling te verkrijgen. De eerste is om willekeurig een van de (meestal) twee beoordelaars als eerste beoordelaar aan te wijzen, en de analyses met die beoordelaar uit te voeren. Als de originele scores gehandhaafd moeten worden is dat een goede manier. Het nadeel is wel dat in dat geval de toevallige selectie van willekeurige eerste beoordelaars impact kan hebben, zeker als er bij vrijwel alle leerlingen meer dan 1 beoordelaar is, zoals bij de 2009 data. Om dat op te lossen zouden een groot aantal trekkingen gedaan moeten worden, en de resultaten over de trekkingen gemiddeld moeten worden.

In de analyses hier hoeven de originele scores niet bewaard te worden. We kunnen dus met gemiddelden werken. Hierbij moet een persoon met meer beoordelaars niet zwaarder wegen. De score van een persoon kan bepaald worden door het gemiddelde van de beoordelaars te gebruiken. Hoe dit de gemiddelde score oplevert is gegeven in Bijlage 13, en kan toegepast worden voor iedere taak-schooltype-jaar-combinatie. De verdelingseigenschappen zijn gegeven in Tabel 8.39.

Tabel 8.39: verdelingseigenschappen vaardigheidsscores (alle beoordelingen)

			Gemiddelde				Standaardafwijking				SE(gemiddelde)			
			OFI	O	F	I	OFI	O	F	I	OFI	O	F	I
T01	sbo	2019	10,0	3,4	2,7	3,9	4,1	2,0	1,7	1,5	0,40	0,19	0,17	0,14
		2009	10,0	3,3	2,5	4,1	3,7	1,6	1,8	1,4	0,35	0,15	0,17	0,12
	bo	2019	11,7	3,9	3,2	4,5	3,3	1,6	1,8	1,2	0,13	0,07	0,08	0,04
		2009	9,5	3,1	2,4	4,0	3,9	1,6	1,9	1,4	0,29	0,11	0,14	0,10
T03	sbo	2019	14,0	3,7	2,3	8,0	5,3	2,1	1,1	3,2	0,48	0,19	0,10	0,30
		2009	13,6	3,7	2,2	7,7	5,3	2,0	1,1	3,3	0,51	0,18	0,10	0,32
	bo	2019	15,9	4,2	2,1	9,6	4,6	1,9	1,2	3,0	0,18	0,08	0,04	0,12
		2009	16,1	4,4	2,1	9,6	5,0	2,0	1,1	3,2	0,31	0,13	0,06	0,19
T08	sbo	2019	8,9	1,4	0,4	7,1	4,6	1,0	0,7	3,7	0,46	0,10	0,07	0,37
		2009	11,7	2,1	0,5	9,1	4,9	1,3	0,8	3,9	0,47	0,12	0,06	0,37
	bo	2019	11,8	2,0	0,6	9,2	4,4	1,1	0,8	3,6	0,19	0,05	0,03	0,16
		2009	13,5	2,4	0,7	10,3	5,4	1,4	0,9	4,2	0,37	0,10	0,06	0,30

Deze verdelingseigenschappen zijn te vertalen naar de effectgroottes voor de jaarverschillen, die zijn berekend door het verschil te delen door de gepoolde standaardafwijking van de twee jaren. Deze effectgroottes worden gegeven in Tabel 8.40. Hierbij is ook aangegeven of het verschil al dan niet significant is.

Tabel 8.40: Effectgroottes van de gemeten vaardigheden (alle beoordelingen)

		OFI		O		F		I	
Taak	Onderwijstype	effect		effect		effect		Effect	
T01	sbo	0,00		0,03		0,09		-0,15	
	bo	0,63	*	0,52	*	0,43	*	0,44	*
T03	sbo	0,08		-0,01		0,07		0,11	
	bo	-0,05		-0,10		-0,03		0,00	
T08	sbo	-0,58	*	-0,58	*	-0,12		-0,54	*
	bo	-0,36	*	-0,34	*	-0,15	*	-0,30	*

*: $p < ,01$

Effectgroottes die in absolute waarde kleiner dan 0.20 zijn kunnen als weinig betekenisvol gezien worden. Alle betekenisvolle effecten zijn significant. Duidelijk is dat het gevonden effect sterk

taakafhankelijk is. Bij het bo zien we een vooruitgang bij Taak 01 en een achteruitgang bij Taak 08. Er is daarbij weinig verschil tussen O, F en I. Als we het gemiddelde over de drie taken zouden nemen komt daar een niet betekenisvol effect uit.

De sbo-resultaten zijn enigszins vergelijkbaar met die bij bo, zij het dat er bij Taak 01 geen vooruitgang gevonden wordt. Over het algemeen wordt daardoor in het sbo een negatiever jaareffect gevonden dan in het bo. Wel moet hierbij opgemerkt worden dat de aantallen waarnemingen in het sbo beperkt waren.

Er kan ook rekening worden gehouden met het feit dat het ontbreken van veel waarden bij leerlingen bij een beoordelaar ook het gevolg kan zijn van de beoordelaars en dat daardoor de beoordeling van deze leerlingen te streng is. Als we nu alle beoordelingen eruit halen waarbij meer dan 10% ontbreekt dan krijgen we een meer milde beoordeling. Door die leerlingen niet mee te nemen verandert ook het aantal beoordelingen per leerling voor een aantal leerlingen. Deze verschuiving is gegeven in Tabel 8.41.

Tabel 8.41: Aantal leerlingen met N beoordelingen per taak als alle beoordelingen meegerekend worden, of alleen de beoordelingen waarvan maximaal 10% van de aspecten niet beoordeeld is.

		Aantal beoordelaars voor deze persoon bij deze taak					Totaal
		1	2	3	4	6	
Aantal beoordelaars met 90% of meer beoordeeld per persoon	0	67	6				73
	1	2587	39				2626
	2		1184		1		1185
	3			5			5
	4				27		27
	5					1	1
	6					3	3
Total		2654	1229	5	28	4	3920

In Tabel 8.41 is te zien dat er bij de milde beoordeling 73 taakcores (persoon taakcombinaties) niet meer meegenomen worden en dat er bij 97.1% van de 3920 taakcores geen enkele beoordeling wegvalt (de vetgedrukte taakcores). Bij 7 taakcores (0.2%) vallen er twee beoordelingen weg. Allereerst vallen er zes taakcores weg met twee beoordelaars voor de persoon bij de taak, terwijl geen van beide 90% of meer had beoordeeld. Verder was er één taakcore waarbij er vier beoordelaars voor de persoon per taak waren, waarvan er twee meer dan 90% hebben beoordeeld. De resultaten voor de scores zoals in Tabel 8.39 is gegeven voor de hele steekproef, zijn hieronder gegeven voor de steekproef met de taakcores die gebaseerd zijn op beoordelingen op (vrijwel) alle aspecten (n=3847).

Tabel 8.42: Verdelingseigenschappen vaardigheidsscores (bij de beoordeling ontbreekt maximaal 10% van de beoordeelde aspecten)

			Gemiddelde				Standaardafwijking				SE(gemiddelde)			
			OFI	O	F	I	OFI	O	F	I	OFI	O	F	I
T01	sbo	2019	10,1	3,4	2,7	4,0	3,9	2,0	1,7	1,4	0,38	0,19	0,17	0,14
		2009	10,0	3,3	2,5	4,2	3,7	1,6	1,8	1,4	0,39	0,17	0,18	0,14
	bo	2019	11,7	4,0	3,2	4,6	3,2	1,6	1,8	1,1	0,13	0,07	0,08	0,04
		2009	9,6	3,1	2,5	4,0	3,9	1,6	1,9	1,4	0,30	0,12	0,14	0,11
T03	sbo	2019	14,5	3,8	2,3	8,3	4,6	2,0	1,0	2,9	0,43	0,19	0,10	0,27
		2009	13,9	3,8	2,2	7,9	4,9	2,0	1,0	3,0	0,54	0,22	0,09	0,34
	bo	2019	16,2	4,3	2,1	9,7	4,1	1,8	1,2	2,7	0,16	0,07	0,04	0,11
		2009	16,4	4,5	2,2	9,7	4,6	1,9	1,0	3,0	0,31	0,13	0,07	0,20
T08	sbo	2019	9,2	1,5	0,4	7,3	4,4	0,9	0,7	3,5	0,45	0,10	0,07	0,36
		2009	11,9	2,1	0,5	9,3	4,6	1,3	0,8	3,6	0,48	0,13	0,07	0,36
	bo	2019	12,1	2,1	0,6	9,5	4,1	1,1	0,8	3,4	0,18	0,05	0,03	0,16
		2009	13,8	2,5	0,7	10,6	4,8	1,3	0,9	3,8	0,37	0,10	0,06	0,30

Tabel 8.43: Effectgroottes van de gemeten vaardigheden (bij de beoordeling ontbreekt maximaal 10% beoordeelde aspecten)

		OFI		O		F		I	
Taak	Onderwijs	effect		effect		effect		effect	
T01	sbo	0,03		0,05		0,11		-0,13	
	bo	0,66	*	0,52	*	0,42	*	0,49	*
T03	sbo	0,13		0,02		0,11		0,16	
	bo	-0,04		-0,09		-0,03		0,02	
T08	sbo	-0,61	*	-0,59	*	-0,11		-0,56	*
	bo	-0,40	*	-0,36	*	-0,16	*	-0,32	*

*: p<,01

Voor de gevonden effecten maakt het weinig uit of er nu wel of niet geselecteerd wordt op het al dan niet beoordelen van (vrijwel) alle aspecten. De gevonden effecten zijn vergelijkbaar met wat eerder gerapporteerd was voor bo.

Op basis van de gegevens is ook het verschil tussen sbo en bo te bepalen. Bij alle taken en alle jaren zien we dat het bo beter scoort dan het sbo. De enige uitzondering is Taak 1 in 2009 waarbij het bo iets slechter presteert dan het sbo.

8.6.2 Vergelijking scores additionele aspecten – bo

Voor het bo is nagegaan in hoeverre er verschillen zijn tussen 2009 en 2019 wat betreft de scores van bo-leerlingen op de additionele aspecten, zoals holistische, algemene indruk en correctheidsfouten, die ook in 2009 geen onderdeel uitmaakten van de schrijfvaardigheidsschaal. In de ankertaken werden drie aspecten onderscheiden: Spelfouten (S), aantal woorden (NW) en algemene indruk van de tekst (A). In Tabel 8.44 en 8.45 zijn de gemiddelden en de standaardafwijking gegeven voor de verschillende taken, per schaal, voor de data met alle leerlingen en de selectie van leerlingen waarbij de leerlingen die meer dan 10% ontbrekende scores hadden, verwijderd zijn (de milde variant).

Tabel 8.44: gemiddelde scores per taak, per jaar, per schaal (alle bo-leerlingen en selectie van 90% beoordeeld)

JAAR	Taak 01		Taak 03		Taak 08	
	2019	2009	2019	2009	2019	2009
S_ALL	1,23	1,65	1,33	1,70	1,38	1,64
NW_ALL	140,43	101,63	81,70	63,59	40,41	41,05
A_ALL	1,12	1,36	0,97	1,21	0,80	1,13
Taak	1		3		8	
JAAR	2019	2009	2019	2009	2019	2009
S_SEL	1,22	1,66	1,33	1,70	1,38	1,64
NW_SEL	141,57	101,63	83,13	63,40	41,22	41,11
A_SEL	1,12	1,37	0,97	1,21	0,80	1,13

Tabel 8.45: Standaardafwijking scores per taak, per jaar, per schaal (alle bo-leerlingen en selectie)

JAAR	Taak 01		Taak 03		Taak 08	
	2019	2009	2019	2009	2019	2009
S_ALL	0,72	0,51	0,70	0,47	0,71	0,51
NW_ALL	63,04	49,66	41,71	26,45	25,20	18,65
A_ALL	0,64	0,61	0,66	0,59	0,65	0,57
Taak	1		3		8	
JAAR	2019	2009	2019	2009	2019	2009
S_SEL	0,72	0,51	0,70	0,47	0,71	0,51
NW_SEL	62,10	49,69	40,69	26,37	24,86	18,74
A_SEL	0,64	0,61	0,66	0,59	0,65	0,58

Het verschil van de gemiddelden is gegeven in Tabel 8.46. Om deze te interpreteren zijn ook de effectgroottes berekend door het verschil te delen door de gepoolde standaardafwijking van de twee jaren.

Tabel 8.46: Verschil in scores en effectgroottes

	Verschil score (Gemiddelde 2019 – 2009)			Effectgrootte (V / pooled sd)		
Schaal	Taak 1	Taak 3	Taak 8	Taak 1	Taak 3	Taak 8
S_ALL	-0,43	-0,37	-0,26	-0,62	-0,57	-0,39
NW_ALL	38,80	18,11	-0,64	0,63	0,47	-0,03
A_ALL	-0,24	-0,24	-0,33	-0,39	-0,37	-0,51
	Verschil score			Effectgrootte		
Schaal	Taak 1	Taak 3	Taak 8	Taak 1	Taak 3	Taak 8
S_SEL	-0,44	-0,37	-0,26	-0,64	-0,57	-0,40
NW_SEL	39,94	19,73	0,10	0,66	0,53	0,00
A_SEL	-0,25	-0,24	-0,32	-0,40	-0,37	-0,51

Als we kijken naar S (Indicatie van het aantal schrijffouten) dan zijn de prestaties over alle taken heen beter in 2009 dan in 2019. Er is een verslechtering geobserveerd (matig effect). Dat is ook het geval bij de algemene indruk (A) die in 2019 iets negatiever is dan in 2009. Aangezien deze twee variabelen niet heel strak gedefinieerd zijn (veel/weinig; goed/voldoende/onvoldoende) is het overigens goed mogelijk dat wat in 2009 en 2019 als veel en weinig, dan wel goed, voldoende, of onvoldoende beschouwd werd in tien jaar is veranderd. Dat is moeilijk te achterhalen. Dat betekent dat het moeilijk is om te achterhalen of de geconstateerde effecten het gevolg zijn van verschillen in prestaties of in beoordelaars. Een objectiever gedefinieerde variabele is het aantal woorden. Het is opvallend dat dit aantal bij Taak 01 en Taak 03 in 2019 aanzienlijk hoger ligt dan in 2009. Bij Taak 08 is echter geen verschil gevonden.

Voor het sbo zijn de analyses aangaande additionele aspecten niet uitgevoerd vanwege de kleine aantallen in het sbo en de niet strakke definiëring. Dit bemoeilijkt een goede interpretatie van de uitkomsten.

Hoofdstuk 9. Meerniveau-analyses

In hoofdstuk 8 zijn de percentages leerlingen die op 1F en op 1S/2F niveau presteren gepresenteerd en de prestaties van de leerlingen kwantitatief getypeerd. Daarnaast zijn voor twee schrijftaken de leerlingteksten kwalitatief getypeerd en is een trendvergelijking gemaakt, waarbij is nagegaan in welke mate de schrijfprestaties anno 2019 verschillen van die bij de vorige peiling in 2009. In dit hoofdstuk wordt beschreven in hoeverre de schrijfprestaties van de leerlingen samenhangen met kenmerken van hun school, hun leerkracht, het onderwijsleerproces in de klas en van henzelf. Hiertoe zijn meerniveau analyses uitgevoerd, waarmee de volgende onderzoeksvragen zijn beantwoord:

- Hoeveel variabiliteit in de leerlingprestaties is toe te schrijven aan scholen en hoeveel aan leerlingen? En: kan worden bepaald welk deel van de variantie is toe te schrijven aan de school en welk deel aan de klas binnen de school?
- In welke mate verschillen scholen in het onderwijsleerproces en in hoeverre hangen deze verschillen samen met verschillen in de leerlingprestaties?
- Welke kenmerken van leerlingen, leerkrachten en onderwijsleerproces zijn in onderlinge samenhang mogelijke verklarende variabelen voor verschillen in leerlingprestaties en schoolprestaties?

In paragraaf 9.1 wordt de strategie beschreven die is gevolgd voor de meerniveau-analyses. Daarin is beschreven hoe is omgegaan met ontbrekende waarden en welke modellen zijn opgebouwd. Tevens wordt in die paragraaf informatie gegeven over de manier waarop effecten zijn berekend en hoe deze moeten worden geïnterpreteerd. In paragraaf 9.2 worden de resultaten beschreven. Na de bespreking van de modellen waarin telkens groepen predictoren (algemene en domeinspecifieke kenmerken op leerling-, klas-, leerkracht- en schoolniveau) zijn opgenomen, presenteren we een compact totaalmodel.

9.1 Gehanteerde strategie voor de meerniveau-analyses

Omdat niet alle informatie bekend was voor elke leerling in elke klas en op elke school zijn ontbrekende waarden geïmputeerd. Dit heeft als voordeel dat daarmee alle leerlingen betrokken zijn in de analyses. Het alternatief, *listwise deletion* van leerlingen met ontbrekende gegevens, zou zorgen voor een kleinere dataset en een verlies aan gegevens die wel bekend zijn. Onderstaand wordt eerst in detail beschreven hoe ontbrekende gegevens zijn geïmputeerd en vervolgens wordt ingegaan op de aanpak van de meerniveau-analyses.

9.1.1 Multipale imputatie van ontbrekende gegevens

De gegevens over leerlingen, leerkrachten, het onderwijsleerproces en algemene kenmerken van de scholen waren grotendeels compleet (rond 90-95%), maar de domeinspecifieke gegevens uit de schoolvragenlijst waren minder volledig (van ongeveer twee derde van de scholen was dit

bekend¹⁴⁶).¹⁴⁷ Op leerlingniveau viel alleen een lager percentage valide waarden op voor de variabele leerlinggewicht. Doordat leerlinggewicht niet op schoolniveau was opgevraagd, hebben testleiders per klas de leerkracht/contactpersoon laten invullen welke leerlingen een leerlinggewicht hadden (en zo ja, welk gewicht). Achteraf was het twijfelachtig of dit correct en volledig is gebeurd, omdat er ook voor de S3- en S4-scholen soms geen leerlingen met een gewicht waren genoteerd. Na afloop is daarom ter controle gebruik gemaakt van DUO-gegevens (per school op leeftijdscategorie). Er waren inderdaad scholen die volgens de gegevens van het DUO wel gewichtenleerlingen moesten hebben, terwijl op die scholen voor geen enkele leerling een leerlinggewicht van 0.3 of 1.2 was ingevuld en alle leerlingen een 0-gewicht hadden. Voor alle leerlingen op die scholen is het leerlinggewicht als ontbrekend gehercodeerd. Door deze hercodering liep het totale percentage ontbrekende waarden in de dataset voor de variabele leerlinggewicht op tot 25.5%.

Ontbrekende gegevens zijn geïmputeerd met behulp van het statistische programma R, pakket Mice (Van Buuren & Groothuis-Oudshoorn, 2011). Mice produceert meerdere imputaties voor multivariate ontbrekende data en gebruikt daarbij de informatie van andere variabelen in het gegevensbestand voor het voorspellen en imputeren van de ontbrekende gegevens. Met behulp van Mice zijn vijf geïmputeerde datasets aangemaakt. Hierbij diende rekening te worden gehouden met de opbouw van het gegevensbestand. Deze heeft een hiërarchische structuur: zowel op leerling- als op klas- en schoolniveau zijn gegevens verzameld. Hierdoor was het niet mogelijk om alle ontbrekende gegevens in één keer te imputeren. Per niveau is daarom apart geïmputeerd, waarbij de gegevens vooraf geaggregeerd werden op het betreffende niveau¹⁴⁸.

Geprobeerd is om zoveel mogelijk informatie te betrekken bij de imputaties, ook de informatie gemeten op een ander niveau dan het betreffende. Mice imputeerde echter niet goed als alle variabelen tegelijk werden meegenomen (te veel informatie). Daarom is steeds gekozen voor de meest relevante variabelen. Bij de imputatie op leerlingniveau is vanwege het verhoogde percentage ontbrekende gegevens over het leerlinggewicht ervoor gekozen om extra gegevens over het leerlinggewicht mee te nemen bij de imputatie. Naast alle variabelen gemeten op leerlingniveau is de informatie op schoolniveau meegenomen die (1) aangeeft of de leerling op een school voor regulier basisonderwijs of voor speciaal basisonderwijs zit, (2) het percentage leerlinggewichten op schoolniveau (schoolcategorieën S1 t/m S4), (3) het percentage 0.3-leerlingen en (4) het percentage 1.2-leerlingen. Voor punt 3 en 4 is gebruik gemaakt van gegevens uit de peiling (vóór imputatie), of, indien onbekend, gegevens van het DUO voor leeftijdscategorie 11-13 jaar. De twee variabelen met percentages 0.3- respectievelijk 1.2-gewichtenleerlingen zijn op dezelfde wijze ingedeeld in categorieën als punt 2, met een categorie 0% (S1), minder dan 10% (S2), 10-25% (S3) en meer dan 25% (S4). Voor de imputatie van alle ontbrekende waarden op klasniveau zijn als extra informatie meegenomen: het gemiddelde (voorlopige) schooladvies van de klas, de gemiddelde scores op schrijfvaardigheid, schrijfplezier en self efficacy, en alle variabelen gemeten op schoolniveau, behalve stedelijkheid en denominatie. Voor de imputatie van ontbrekende waarden op schoolniveau zijn dezelfde leerlingvariabelen opgenomen als bij de imputatie op klasniveau, en daarnaast ook de meeste variabelen gemeten op klasniveau. Alleen die variabelen die gemeten waren op klasniveau en niet op zinvolle wijze geaggregeerd konden worden naar schoolniveau, zijn buiten beschouwing gelaten. Het gaat dan om variabelen als vooropleiding en geslacht van de leerkracht.

¹⁴⁶ In hoofdstuk 6 is nadere informatie gegeven over de respons op de schoolvragenlijst.

¹⁴⁷ In Tabel 5.18 is een overzicht gegeven van de variabelen die in de meerniveau analyses zijn opgenomen.

¹⁴⁸ Meerniveau multiple imputatie was helaas niet mogelijk, omdat daarvoor teveel gegevens ontbraken.

Op elk meetniveau zijn steeds vijf geïmputeerde datasets aangemaakt. Voor elke imputatie zijn de gegevens op leerling- klas- en schoolniveau weer samengevoegd tot één bestand, met als resultaat in totaal vijf volledige, geïmputeerde datasets.

9.1.2 Meerniveau regressie-analyse

Vanwege de hiërarchische structuur van de data (leerlingen in klassen op scholen) is meerniveau regressie-analyse de aangewezen methode om de invloed van verschillende variabelen op de schrijfvaardigheid van leerlingen te onderzoeken. De meerniveau-analyses zijn uitgevoerd met het statistische programma MLwiN, versie 2.27 (Rasbash, Browne, Healy, Cameron & Charlton, 2013). De analyses zijn steeds uitgevoerd voor elk van de vijf geïmputeerde datasets, waarna de resultaten van deze vijf datasets zijn gecombineerd (gepooled) door middel van de formules van Rubin en Schenker (1986). De school vormde het derde niveau ($n = 133$), de klas ($n = 205$) het tweede en de leerling ($n = 3358$) het eerste niveau. Voor de bepaling of een effect significant is, is uitgegaan van een p -waarde kleiner dan 0.05, bij tweezijdige toetsing.

Ter beantwoording van de onderzoeksvragen zijn meerdere modellen getoetst met de score op schrijfvaardigheid als afhankelijke variabele. De score op schrijfvaardigheid is de WML-schatting op de onderliggende variabele en is gebaseerd op de antwoorden van de leerling op alle items die hij/zij heeft beantwoord. De schrijfvaardigheidsscores hadden oorspronkelijk een beperkte range ($M = -0.07$; $SD = 0.21$; minimum = -1.10 ; maximum = 0.78), waardoor verschillen in prestaties nauwelijks opvielen. Bovendien bevatte de schaal negatieve scores die lastig zijn te interpreteren. Om deze redenen zijn de WML-schattingen voor de start van de meerniveau-analyses getransformeerd. De transformatie heeft geen gevolgen voor de relatieve verschillen tussen leerlingen. Zou je de oorspronkelijke scores en getransformeerde scores omzetten in gestandaardiseerde Z-scores, dan krijg je exact dezelfde waarden voor dezelfde leerlingen. De volgende formule is gebruikt voor de transformatie:

$$\text{Schrijfvaardigheid} = (\text{WML-schatting} + 1.1) * 100.$$

Als eerste is een leeg model getoetst, zonder variabelen. Dit is Model 0. Model 1 bevat alleen de covariaat type onderwijs (regulier of speciaal basisonderwijs). De volgende modellen bevatten steeds het type onderwijs als covariaat en elk een aparte groep predictoren. In Model 2 zijn de algemene en domeinspecifieke leerlingkenmerken opgenomen als predictoren, in Model 3 de kenmerken van het onderwijsleerproces, Model 4 bevat de algemene en domeinspecifieke leerkrachtkenmerken en Model 5 de algemene en domeinspecifieke schoolkenmerken.

De hierboven genoemde modellen zijn daarnaast ook apart getoetst voor het regulier basisonderwijs en het speciaal basisonderwijs (zonder type onderwijs als covariaat). Deze modellen zijn gebruikt voor de selectie van variabelen van het onderwijsleerproces waarbij een interactie met type onderwijs getoetst zou gaan worden. Hiertoe zijn de analyses per type onderwijs met elkaar vergeleken, waarbij op basis van deze analyses én op inhoudelijke gronden een aantal interactie-effecten tussen variabelen van het onderwijsleerproces en het type onderwijs is geselecteerd. Dit is gedaan in onderlinge afstemming met de Inspectie van het Onderwijs. De onderwijsleerproces-variabelen, die zijn gekozen op basis van de analyses, zijn geselecteerd op basis van een tegengesteld effect¹⁴⁹ op de schrijfvaardigheid van leerlingen in het regulier basisonderwijs en het speciaal

¹⁴⁹ Inhoudelijk waren we het meest geïnteresseerd in de tegengestelde effecten. Er had ook kunnen worden gekeken naar interacties die significant zijn maar een effect in het sbo verkleinen of vergroten (zonder de richting aan te tasten). Hier

onderwijs. Hierbij was het in dat stadium het niet mogelijk om te toetsen of het verschil in effect in bo en sbo *significant* was. Dat is juist later gedaan met de interactie-effecten. De interactie-effecten zijn één voor één getoetst in het model met de kenmerken van het onderwijsleerproces (Model 3) waarin beide onderwijstypen tegelijk zijn geanalyseerd.

Op basis van de resultaten van de modellen 2 tot en met 5 en de getoetste interactie-effecten is een totaalmodel opgesteld met daarin alle variabelen van het onderwijsleerproces en de significante interactie-effecten met type onderwijs, de significante domeinspecifieke school-, leerkracht- en leerlingkenmerken en alle (ook de niet-significante) algemene leerling, leerkracht- en schoolkenmerken. Vervolgens is door middel van een backwards (stepwise) regressie een compact totaalmodel samengesteld, waarbij de verklaarde variantie van het model maximaal is met zo min mogelijk variabelen. De stapsgewijze backwards regressie-analyse is uitgevoerd op slechts één van de datasets, at random gekozen. Voorafgaand daaraan is gecontroleerd dat er geen dataset was met duidelijk afwijkende varianties. Er zijn vervolgens net zolang variabelen uit het model verwijderd tot het verwijderen van de variabele die op dat moment de minste variantie in schrijfvaardigheidsscores verklaarde, tot een significant slechtere model fit leidde. Het ontstane compacte totaalmodel is daarna getoetst bij elk van de vijf geïmputeerde datasets en de resultaten daarvan zijn weer gecombineerd (gepooled).

9.1.3 Berekening van effecten

Per model wordt aangegeven hoeveel procent van de variantie in schrijfvaardigheidsscores de opgenomen predictoren verklaren ten opzichte van het model met alleen de covariaat type onderwijs. Daarnaast berekenen we voor elke significante predictorvariabele de effectgrootte Cohen's f^2 . De f^2 geeft de proportie variantie in schrijfvaardigheid aan die een variabele extra verklaart ten opzichte van alle verklaarde variantie. Zou je de waarde met 100 vermenigvuldigen, dan krijg je het percentage verklaarde variantie. De f^2 wordt als volgt berekend: Je hebt twee modellen, één model met alle variabelen (model AB) en een model met alle variabelen, maar zonder de specifieke predictorvariabele waar je de effectgrootte voor wilt berekenen (model A). $f^2 = (\text{residuele variantie model A} - \text{residuele variantie model AB}) / \text{residuele variantie model A}$.¹⁵⁰

De interpretatie van Cohen's f^2 is anders dan die van de meer gebruikelijke Cohen's d , namelijk een f^2 van 0.02, van 0.15 en van 0.35 zijn respectievelijk een klein, matig en groot effect (Cohen, 1988). De effectgroottes van variabelen zijn steeds berekend op basis van dezelfde at random gekozen dataset als bij de backwards regressie-analyse.

hebben we van afgezien, omdat het lastig te bepalen zou zijn bij welk verschil in effect (in dezelfde richting) we een variabele wel of niet zouden moeten opnemen. Dit was lastig, omdat nog niet was vast te stellen of het verschil significant was. Of een verschil ertoe deed, was per variabele weer anders, afhankelijk van de SD's/SE's. Overigens blijkt uit de resultaten achteraf dat de getoetste interactie-effecten ten hoogste een heel beperkte invloed hebben. We verwachten daarom niet dat het eventueel gemist hebben van een interactie-effect, door het niet opnemen van interactie met een variabele die in het BO en SBO een effect in dezelfde richting had maar met een groter/kleiner effect, de resultaten merkbaar zou beïnvloeden.

¹⁵⁰ In plaats van de hier genoemde formule is f^2 ook de relatieve toename in R^2 door het toevoegen van de variabele B bij

$$f^2 = \frac{R_{AB}^2 - R_A^2}{1 - R_A^2}$$

alle andere regressoren (hier voorgesteld als A) : . Gebruik van deze formule levert dezelfde uitkomsten op.

9.2 Resultaten meerniveau-analyses

Model 0 en 1, leeg model en model met covariaat

Tabel 9.1 bevat de gecombineerde effecten op schrijfvaardigheid van de vijf datasets voor het lege model (Model 0) en het model met type onderwijs als covariaat (Model 1). Het intercept in Model 0 is de gemiddelde score op schrijfvaardigheid van alle leerlingen in de steekproef. Door imputatie van ontbrekende waarden en door rekening te houden met de geneste structuur van de data (leerlingen in klassen op scholen) kan dit gemiddelde iets afwijken van het gemiddelde in Tabel 8.8. Het intercept in Model 1 is de gemiddelde score op schrijfvaardigheid van de leerlingen in het regulier basisonderwijs (de referentiecategorie voor de covariaat type onderwijs). De regressiecoëfficiënt voor sbo geeft aan dat de gemiddelde score van de sbo-leerlingen -18.57 punten lager is dan die van leerlingen in het regulier basisonderwijs. Het sterretje (*) achter de standaardfout (*SE*) van de regressiecoëfficiënt laat zien dat dit verschil significant is bij $p < 0.05$.

Het random deel van de modellen geeft de variantie in scores weer, uitgesplitst naar niveau. Dit geeft inzicht in welk deel van de totale residuele variantie is toe te schrijven aan verschillen tussen leerlingen (variantie op leerlingniveau), aan verschillen tussen klassen (variantie op klasniveau) en aan verschillen tussen scholen (variantie op schoolniveau). In Model 0 is te zien dat de variantie op zowel school- als op klas- en leerlingniveau significant is. De intraklasse-coëfficiënten (ICC) geven aan welk deel van de totale residuele variantie op het betreffende niveau ligt. De ICC voor school is berekend door de residuele variantie op schoolniveau te delen door de totale residuele variantie; de ICC voor klas is berekend door de som van de residuele variantie op school- en klasniveau te delen door de totale residuele variantie. De ICC van 0.212 voor school geeft aan dat 21.2% van de variantie op schoolniveau ligt. Zitten leerlingen in dezelfde klas op dezelfde school (ICC klas), dan gaat het om 30.7% van de totale residuele variantie. In Model 1 is aan het random deel te zien dat de variantie op schoolniveau fors is afgenomen na rekening te hebben gehouden met het type onderwijs. Nu ligt nog 3.08% van de totale residuele variantie op schoolniveau en als leerlingen in dezelfde klas zitten dan gaat het om 16.6%. Het overgrote deel van de totale residuele variantie in schrijfvaardigheid is toe te schrijven aan verschillen tussen leerlingen. De onderste regel in Tabel 9.1 geeft aan hoeveel van de totale residuele variantie in schrijfvaardigheid verklaard wordt door rekening te houden met het type onderwijs. De Cohen's f^2 is 0.168. Dit betekent dat 16.8% van de variantie in scores verklaard wordt door de covariaat type onderwijs; een matig effect.

In de volgende modellen is type onderwijs steeds als covariaat opgenomen, maar is het effect van deze variabele niet steeds berekend. Dit omdat in de modellen met leerlinggewicht of schoolcategorie er steeds een categorie 'n.v.t. sbo' is, die samenvalt met het onderscheid tussen regulier basisonderwijs en speciaal basisonderwijs. Weglating van de covariaat geeft dan vertekende resultaten.

Tabel 9.1: Modellen zonder predictoren

	Model 0, leeg model		*	Model 1, met covariaat		*
	<i>b</i>	<i>SE(b)</i>		<i>b</i>	<i>SE(b)</i>	
<i>Fixed part</i>						
Intercept	102.498	1.054	*	107.872	0.829	*
Sbo (ref. bo)				-18.570	1.520	*
<i>Random Part</i>						
Schoolniveau	97.722	18.226	*	14.612	9.043	
Klasniveau	43.652	10.548	*	48.982	9.477	*
Leerlingniveau	319.782	15.035	*	320.061	15.061	*
Totale residuele variantie	461.155			383.655		
ICC school	0.212			0.038		
ICC klas	0.307			0.166		
Aantal scholen	133			133		
Aantal klassen	205			205		
Aantal leerlingen	3358			3358		
Verklaarde variantie t.o.v. Model 0 (Cohen's f^2)				0.168		

* = $p < 0.05$

Model 2, leerlingkenmerken

Tabel 9.2 bevat de gecombineerde effecten van de algemene en domeinspecifieke leerlingkenmerken op schrijfvaardigheid (Model 2). Het type onderwijs is opgenomen als covariaat. Van de algemene leerlingkenmerken hangen geslacht en het voorlopig schooladvies significant samen met de schrijfvaardigheid. Meisjes hebben hogere scores dan jongens (Cohen's $f^2 = 0.056$, een klein effect), en schrijfvaardigheidsscores lijken toe te nemen naarmate leerlingen een hoger voorlopig advies hebben (Cohen's $f^2 = 0.081$, een klein effect). Tabel 9.2 laat zien dat, ten opzichte van de referentiecategorie advies havo/vwo, de schrijfvaardigheid van leerlingen met een lager advies significant lager is. Leeftijd, land van herkomst, thuistaal en leerlinggewicht hangen niet significant samen met schrijfvaardigheid. In het bijzonder voor leerlinggewicht is dit anders dan verwacht. Een aanvullende analyse van Model 2 zonder de predictor voorlopig advies wees uit dat het effect van leerlinggewicht op de schrijfvaardigheid in dat geval wel significant is. Door rekening te houden met het voorlopig advies wordt het effect van leerlinggewicht 'weg'-verklaard.

Van de domeinspecifieke leerlingkenmerken hangen de predictoren schrijven in de vrije tijd, schrijfplezier en self-efficacy significant samen met schrijfvaardigheid. Tabel 9.2 toont dat leerlingen die in hun vrije tijd schrijven significant hogere scores hebben dan leerlingen die nooit in hun vrije tijd schrijven (Cohen's $f^2 = 0.004$, een heel klein effect), behalve de leerlingen die bijna iedere dag in hun vrije tijd schrijven. Dit laatste is een onverwacht effect. De hoogte van de regressiecoëfficiënten doen overigens vermoeden dat het niet zo lijkt te zijn dat hoe meer er geschreven wordt, des te beter de prestaties. Voor de schalen schrijfplezier (Cohen's $f^2 = 0.025$, een klein effect) en self-efficacy (Cohen's $f^2 = 0.006$, een heel klein effect) geldt dat hogere scores samenhangen met hogere prestaties op schrijfvaardigheid. Het domeinspecifieke kenmerk schrijven op de mobiel hangt niet samen met de schrijfvaardigheid.

Het random deel van Model 2, in Tabel 9.2, laat zien dat de leerling kenmerken samen een matig effect hebben op de schrijfvaardigheid. Samen verklaren ze 21.5% van de totale residuele variantie in schrijfvaardigheidsscores ten opzichte van het model met alleen de covariaat type onderwijs (Cohen's $f^2 = 0.215$). Vooral de variantie op leerlingniveau is afgenomen ten opzichte van Model 1, maar ook de varianties op school- en klasniveau zijn een klein beetje afgenomen.

Tabel 9.2. Model met leerlingkenmerken

	Model 2, leerlingkenmerken			Cohen's f^2
	<i>b</i>	<i>SE(b)</i>		
<i>Fixed Part</i>				
Intercept	97.252	8.580	*	
Sbo (ref. bo)	-5.477	1.832	*	
Leeftijd (in maanden)	-0.031	0.055		
Meisje	9.572	0.818	*	0.056
Niet in NL geboren	-1.334	1.532		
Thuis taal (ref. NL):				
Thuis taal alleen dialect of Fries	-0.252	1.776		
Thuis taal alleen andere taal	-2.490	1.742		
Leerlinggewicht (ref. geen gewicht):				
0.3 leerlinggewicht	-2.994	2.391		
1.2 leerlinggewicht	-1.427	1.958		
N.v.t. sbo	0.000	0.000		
Voorlopig advies (ref. havo/vwo):				0.081
Vmbo g/t	-6.016	0.797	*	
Vmbo b/k	-10.525	1.067	*	
PRO	-19.912	1.944	*	
Schrijven op mobiel (ref. nooit):				
Heb geen telefoon	3.121	2.820		
Paar keer per week	0.873	2.717		
Bijna elke dag	2.047	2.540		
Paar keer per dag	2.458	2.528		
Schrijven in vrije tijd (ref. nooit):				0.004
1 a 2 keer per jaar	3.733	1.216	*	
1 x per maand	3.937	1.244	*	
1x per week	2.804	1.138	*	
Bijna iedere dag	1.931	1.050		
Schrijfplezier (schaal)	0.442	0.067	*	0.025
Self-efficacy (schaal)	0.342	0.103	*	0.006
<i>Random Part</i>				
Schoolniveau	8.625	6.321		
Klasniveau	38.300	8.179	*	
Leerlingniveau	254.306	14.298	*	
Totale residuele variantie	301.231			
ICC school	0.029			
ICC klas	0.156			
Aantal scholen	133			
Aantal klassen	205			
Aantal leerlingen	3358			
Verklaarde variantie t.o.v. Model 1 (Cohen's f^2)	0.215			

* = $p < 0.05$

Model 3, onderwijsleerproces

In Tabel 9.3 staan de gecombineerde effecten op schrijfvaardigheid van de kenmerken van het onderwijsleerproces (Model 3). Het type onderwijs is weer opgenomen als covariaat. Van alle getoetste kenmerken hangen er vijf significant samen met de leerlingprestaties op schrijfvaardigheid. Feedback op inhoud heeft een positieve invloed op de schrijfvaardigheid (Cohen's $f^2 = 0.008$, een heel klein effect). Dit geldt ook voor het frequenter onderwijzen van kennis over genres (Cohen's $f^2 = 0.006$, een heel klein effect) en het leren van anderen (Cohen's $f^2 = 0.005$, een heel klein effect). Kennis over het schrijfproces, daarentegen, heeft een negatief effect op de schrijfvaardigheid van leerlingen (Cohen's $f^2 = 0.005$, een heel klein effect). Tot slot heeft de taalmethode invloed op de schrijfvaardigheid (Cohen's $f^2 = 0.013$, een heel klein effect). De verschillen tussen de onderscheiden taalmethodes lijken, gezien de hoogtes van de regressiecoëfficiënten, nagenoeg verwaarloosbaar. Het meest positieve effect is voor Taalverhaal.nu, maar een aanvullende analyse met Taalverhaal.nu als referentiecategorie wijst uit dat deze methode niet significant verschilt van de andere gespecificeerde taalmethodes. Wel heeft, in vergelijking tot de meest gebruikte taalmethode Taal Actief, het gebruik van meerdere taalmethodes een negatieve invloed en juist het niet gebruiken van een taalmethode een positieve invloed op de schrijfvaardigheid. Mogelijk maken leerkrachten die ervoor kiezen geen taalmethode te gebruiken een weloverwogen(er) selectie van het leerstofaanbod en kiezen zij bewust voor een andere aanpak van het schrijfonderwijs.

Het random deel in Tabel 9.3 laat zien dat alle kenmerken van het onderwijsleerproces samen 4.8% van de totale residuele variantie in schrijfvaardigheid verklaren ten opzichte van het model met alleen de covariaat type onderwijs (Cohen's $f^2 = 0.048$, een klein effect). Ten opzichte van Model 1, met alleen de covariaat type onderwijs, is te zien dat alleen de variantie op klasniveau is afgenomen.

Tabel 9.3. Model met kenmerken van het onderwijsleerproces

	Model 3, onderwijsleerproces			Cohen's f^2
	<i>b</i>	<i>SE(b)</i>		
<i>Fixed Part</i>				
Intercept	97.128	6.318	*	
Sbo (ref. bo)	-18.282	1.893	*	
Schrijftijd per week (in minuten)	0.000	0.011		
Min. 2x p. maand schrijfpodracht	-0.598	1.584		
Min. 2x p. maand schrijfpodracht binnen ander vak	-1.486	1.689		
Wel nadruk doel en publiek	0.339	1.658		
Feedback op inhoud (schaal)	3.732	1.382	*	0.008
Kennis over genre 1x p. maand of meer	3.372	1.544	*	0.006
Feedback op technische aspecten (schaal)	1.042	1.418		
Kennis leestekens hoofdletters 1x p. maand of meer	-0.645	1.910		
Prewriting (schaal)	0.101	1.683		
Revisie (schaal)	1.247	1.230		
Kennis schrijfproces (schaal)	-4.009	1.871	*	0.005
Meestal of vaak feedback op leerdoel	-1.937	1.394		
Samenwerking (schaal)	-0.627	1.192		
Leren van anderen (schaal)	3.055	1.146	*	0.005
Schrijven om te leren (schaal)	1.117	1.308		
Aandacht voor soorten teksten (schaal)	0.303	0.542		
Aandacht voor soorten doelen (schaal)	-0.966	0.944		
Aandacht voor soorten publiek (schaal)	0.006	0.655		
Taalmethode (ref. Taal actief):				0.013
STAAL	-1.703	2.155		
Taal op maat	-2.031	2.419		
Taal in beeld	-3.348	2.906		
Taalverhaal.nu	1.369	2.639		
Andere taalmethode	-2.874	2.247		
Geen taalmethode	7.712	3.708	*	
Meer dan 1 taalmethode	-6.078	2.501	*	
Taalmethode + extra leermiddelen	-0.448	1.379		
Wel zelf ontwikkeld materiaal schrijfonderwijs	-1.321	1.267		
Beoordeling o.b.v. aspecten tekst meestal of vaak	-0.262	1.424		
Beoordeling o.b.v. rubrics meestal of vaak	3.296	1.984		
Wel differentiatie	1.072	1.815		

* = $p < 0.05$

Tabel 9.3. Model met kenmerken van het onderwijsleerproces (vervolg)

	Model 3, onderwijsleerproces		
	<i>b</i>	<i>SE(b)</i>	Cohen's f^2
<i>Random Part</i>			
Schoolniveau	17.908	7.388	*
Klasniveau	27.668	6.032	*
Leerlingniveau	319.765	15.055	*
Totale residuele variantie	365.341		
ICC school	0.049		
ICC klas	0.125		
Aantal scholen	133		
Aantal klassen	205		
Aantal leerlingen	3358		
Verklaarde variantie t.o.v. Model 1 (Cohen's f^2)	0.048		

* = $p < 0.05$

Model 4, leerkrachtkenmerken

Tabel 9.4 toont de effecten op schrijfvaardigheid van de algemene en domeinspecifieke leerkrachtkenmerken (Model 4). Type onderwijs is weer als covariaat opgenomen. De algemene kenmerken geslacht en vooropleiding van de leerkracht hangen niet significant samen met de schrijfvaardigheid van leerlingen, maar de werkervaring van de leerkracht wel (Cohen's $f^2 = 0.006$, een heel klein effect). Leerlingen die les krijgen van leerkrachten die nul tot vijf jaar lesgeven hebben iets hogere scores op schrijfvaardigheid dan leerlingen die les krijgen van leerkrachten met meer dan 15 jaar werkervaring. De schrijfvaardigheid van leerlingen die les krijgen van leerkrachten met zes tot tien jaar of 11 tot 15 jaar werkervaring verschilt niet significant van de referentiecategorie nul tot vijf jaar werkervaring. De twee domeinspecifieke leerkrachtkenmerken, professionalisering en voldoende kennis en vaardigheid op het gebied van schrijfonderwijs/-les, hangen niet significant samen met de schrijfvaardigheid van leerlingen. Alle leerkrachtkenmerken samen verklaren slechts 0.9% extra van de totale residuele variantie in schrijfvaardigheid ten opzichte van het model met alleen de covariaat type onderwijs (Cohen's $f^2 = 0.009$, een heel klein effect). Vergeleken met Model 1, met alleen de covariaat type onderwijs, is de variantie op klasniveau een fractie gedaald en de variantie op schoolniveau een fractie gestegen.

Tabel 9.4. Model met leerkrachtkenmerken

	Model 4, leerkrachtkenmerken		
	<i>b</i>	<i>SE(b)</i>	Cohen's f^2
<i>Fixed Part</i>			
Intercept	109.121	1.995	*
Sbo (ref. bo)	-17.947	1.681	*
Man	0.908	1.386	
Vooropleiding (ref. PABO):			
Vervolgopleiding SEN of OSO	-0.338	3.084	
Verkorte pabo	-3.274	2.452	
Werkervaring als leerkracht (ref. 0-5 jaar):			0.006
6-10 jaar	-2.178	2.110	
11-15 jaar	-2.716	1.906	
16-20 jaar	-5.075	2.383	*
Meer dan 21 jaar	-4.923	2.061	*
Wel professionalisering schrijfonderwijs	-0.357	2.208	
Wel kennis over schrijfles	2.231	1.673	
<i>Random Part</i>			
Schoolniveau	18.183	8.939	*
Klasniveau	42.184	9.059	*
Leerlingniveau	320.013	15.063	*
Totale residuele variantie	380.381		
ICC school	0.048		
ICC klas	0.159		
Aantal scholen	133		
Aantal klassen	205		
Aantal leerlingen	3358		
Verklaarde variantie t.o.v. Model 1 (Cohen's f^2)	0.009		

* = $p < 0.05$

Model 5, schoolkenmerken

In Tabel 9.5 staan de gecombineerde effecten op schrijfvaardigheid van de algemene en domeinspecifieke schoolkenmerken (Model 5). Ook nu is het type onderwijs weer opgenomen als covariaat. Van de algemene kenmerken vinden we een significante samenhang met schrijfvaardigheid voor de schoolcategorie en voor denominatie. Op scholen zonder gewichtenleerlingen (categorie S1) hebben leerlingen een hogere schrijfvaardigheid dan op scholen met 10% of meer gewichtenleerlingen (categorieën S3 en S4). Het gaat om een heel klein effect (Cohen's $f^2 = 0.013$). De schrijfvaardigheid van leerlingen op scholen met maximaal 10% gewichtenleerlingen (categorie S2) verschilt niet significant van de referentiecategorie S1. Voor denominatie is het effect eveneens heel klein (Cohen's $f^2 = 0.004$). Leerlingen op scholen met als denominatie 'overig levensbeschouwelijk' hebben iets lagere prestaties voor schrijfvaardigheid dan leerlingen op openbare scholen (de referentiecategorie). Daarnaast is te zien dat geen van de domeinspecifieke schoolkenmerken een significante samenhang heeft met schrijfvaardigheid.

Het random deel van Model 5 laat zien dat de schoolkenmerken 2.9% extra van de totale residuele variantie in schrijfvaardigheid verklaren ten opzichte van het model met alleen de covariaat type onderwijs (Cohen's $f^2 = 0.029$, een klein effect). De iets lagere variantie ten opzichte van Model 1, met alleen het type onderwijs als covariaat, komt voornamelijk voor rekening van het schoolniveau.

Tabel 9.5. Model met schoolkenmerken

	Model 5, schoolkenmerken		
	<i>b</i>	<i>SE(b)</i>	Cohen's f^2
<i>Fixed Part</i>			
Intercept	113.600	5.569	*
Sbo (ref. bo)	-22.818	3.565	*
Schoolcategorie (ref. S1, 0% gewichtenleerlingen):			0.013
S2, max 10% gewichtenleerlingen	-1.481	3.156	
S3, 10-25% gewichtenleerlingen	-8.567	3.498	*
S4, >25% gewichtenleerlingen	-8.851	4.137	*
N.v.t. sbo	0.000	0.000	
Aantal leerlingen op school	-0.005	0.005	
Stedelijkheid (ref. niet stedelijk):			
Weinig stedelijk	0.244	2.529	
Matig stedelijk	-0.197	2.434	
Sterk stedelijk	1.125	2.422	
Zeer stedelijk	2.580	2.684	
Denominatie (ref. openbaar):			0.004
Protestants-christelijk	-1.317	2.187	
Rooms-katholiek	-2.770	2.165	
Algemeen bijzonder	0.805	2.433	
Overig levensbeschouwelijk	-6.113	3.038	*
Wel doorgaande leerlijn vanaf groep 3	1.514	1.867	
Wel systematisch monitoren schrijfprestaties	0.567	2.412	
Schoolbrede focus schrijfonderwijs (schaal)	-0.055	0.137	
Wel expliciet taalbeleidsplan	0.106	1.987	
Wel taalcoördinator	-1.561	1.645	
<i>Random Part</i>			
Schoolniveau	6.048	7.760	
Klasniveau	46.415	9.033	*
Leerlingniveau	320.058	15.061	*
Totale residuele variantie	372.521		
ICC school	0.016		
ICC klas	0.141		
Aantal scholen	133		
Aantal klassen	205		
Aantal leerlingen	3358		
Verklaarde variantie t.o.v. Model 1 (Cohen's f^2)	0.029		

* = $p < 0.05$

Model 6, compact totaalmodel

Voorafgaand aan de opbouw van het totaalmodel zijn een aantal interactie-effecten tussen type onderwijs en kenmerken van het onderwijsleerproces onderzocht. De selectie van variabelen is gebaseerd op modeluitkomsten voor regulier en speciaal basisonderwijs apart en op basis van een inhoudelijke expert. Het gaat om interacties met de volgende kenmerken van het onderwijsleerproces: samenwerking (schaal), wel/niet zelf ontwikkeld materiaal schrijfonderwijs, frequentie schrijfopdrachten binnen andere vakken, leren van anderen (schaal), wel/niet nadruk doel en publiek, schrijven om te leren (schaal), aandacht voor soorten teksten (schaal) en feedback op technische aspecten (schaal). Elke interactie is apart getoetst in het model met de kenmerken van het onderwijsleerproces (Model 3). Van alle getoetste interacties hebben er twee een significant effect. Het gaat om type onderwijs x leren van anderen (schaal) en type onderwijs x aandacht voor soorten teksten (schaal).

De twee significante interacties zijn toegevoegd aan het totaalmodel, dat als basis heeft: alle kenmerken van het onderwijsleerproces, alle algemene leerling-, leerkracht- en schoolkenmerken en de significante domeinspecifieke leerling-, leerkracht- en schoolkenmerken. Middels backwards regressie is het model gereduceerd tot een compact totaalmodel dat met zo weinig mogelijk predictoren maximale variantie verklaart in schrijfvaardigheid. Tabel 9.6 bevat de gecombineerde effecten op schrijfvaardigheid van de vijf datasets voor het compacte totaalmodel (Model 6). Vergeleken met Model 1, met alleen de covariaat type onderwijs, verklaart het compacte totaalmodel 25.5% extra variantie in schrijfvaardigheid (Cohen's $f^2 = 0.255$, een matig effect). De grootste daling in variantie komt voor rekening van het leerlingniveau, maar ook op klasniveau is de variantie duidelijk afgenomen. De variantie op schoolniveau is een beetje afgenomen. De intraklasse coëfficiënten (ICC) laten zien dat ten opzichte van Model 1, met alleen type onderwijs als covariaat, de variantie op klasniveau is afgenomen van 16.6% naar 10.6% en de variantie op schoolniveau is afgenomen van 3.8% naar 3.6%.

In Tabel 9.6 is te zien dat niet elke variabele een significant effect heeft. Toch zijn deze niet-significante variabelen in het model gebleven, omdat weglating tot een slechter model fit leidde. Van alle variabelen in het model hangt het advies het sterkst samen met de schrijfvaardigheid, gevolgd door geslacht. Beide hebben met een Cohen's f^2 van rond de 0.07 een klein effect. Daarna volgt schrijfplezier met een Cohen's f^2 van 0.02; ook een klein effect. De taalmethode en de schaal leren van anderen hebben met een Cohen's f^2 van 0.01 een heel klein effect, waarbij het effect van de laatste niet significant is. Wel is het heel kleine effect voor de interactie tussen type onderwijs en de schaal leren van anderen significant. Het interactie-effect geeft aan dat in het speciaal basisonderwijs een hogere score op de schaal leren van anderen een grotere positieve invloed heeft op de schrijfvaardigheid dan in het regulier basisonderwijs. De effecten van de overige variabelen zijn met een Cohen's f^2 van onder de 0.01 heel klein en niet altijd meer significant. De richting van de effecten is gelijk aan die in de Modellen 2 t/m 5, waarbij drie dingen opvallen. Als eerste: voor schrijven in de vrije tijd is het verschil tussen nooit schrijven en bijna iedere dag schrijven in de vrije tijd nu wel significant. Als tweede: het verschil tussen de taalmethode Taal actief en Taal in beeld is nu wel significant. Tot slot, voor denominatie is het verschil tussen openbaar en Rooms-katholiek nu wel significant.

Tabel 9.6. Compact totaalmodel

	<i>Model 6, totaal model</i>		
	<i>b</i>	<i>SE(b)</i>	<i>Cohen's f²</i>
<i>Fixed Part</i>			
Intercept	97.955	4.341	*
Sbo (ref. bo)	-16.355	4.533	*
Feedback op inhoud (schaal)	2.203	1.204	0.004
Kennis over genre 1x p. maand of meer	3.312	1.176	* 0.006
Kennis schrijfproces (schaal)	-2.801	1.256	* 0.006
Leren van anderen (schaal)	1.472	0.917	0.010
Interactie Sbo x Leren van anderen (schaal)	3.559	1.472	* 0.003
Taalmethode (ref. Taal actief):			0.013
STAAL	-1.066	1.558	
Taal op maat	-3.004	1.881	
Taal in beeld	-4.974	2.319	*
Taalverhaal.nu	0.962	2.559	
Andere taalmethode	-1.863	1.955	
Geen taalmethode	6.103	2.767	*
Meer dan 1 taalmethode	-5.177	2.316	*
Schrijven in vrije tijd (ref. nooit):			0.005
1 a 2 keer per jaar	2.966	1.154	*
1x per maand	2.246	1.032	*
1x per week	4.219	1.247	*
Bijna iedere dag	3.962	1.207	*
Schrijfplezier (schaal)	0.412	0.068	* 0.019
Self-efficacy (schaal)	0.375	0.101	* 0.006

* = $p < 0.05$

Tabel 9.6. Compact totaalmodel (vervolg)

	Model 6, totaal model			
	b	SE(b)		Cohen's f^2
Meisje	9.781	0.814	*	0.065
Voorlopig advies (ref. havo/vwo):				0.073
Vmbo g/t	-6.111	0.787	*	
Vmbo b/k	-10.753	1.034	*	
PRO	-19.739	1.874	*	
Vooropleiding (ref. PABO):				0.006
Vervolgopleiding SEN of OSO	1.547	2.850		
Verkorte pabo	-3.404	2.118		
Werkervaring als leerkracht (ref. 0-5 jaar):				0.005
6-10 jaar	-1.525	1.793		
11-15 jaar	-3.436	1.572	*	
16-20 jaar	-3.547	2.019		
Meer dan 21 jaar ¹⁵¹	-3.515	1.746	*	
Schoolcategorie (ref. S1, 0% gewichtenleerlingen):				0.008
S2, max 10% gewichtenleerlingen	-0.835	2.162		
S3, 10-25% gewichtenleerlingen	-4.215	2.595		
S4, >25% gewichtenleerlingen	-6.819	3.381	*	
N.v.t. sbo	0.000	0.000		
Denominatie (ref. openbaar):				0.007
Protestants-christelijk	-1.550	1.690		
Rooms-katholiek	-3.157	1.561	*	
Algemeen bijzonder	-0.102	2.402		
Overig levensbeschouwelijk	-6.139	3.382		
<i>Random Part</i>				
Schoolniveau	10.406	4.973	*	
Klasniveau	19.768	5.385	*	
Leerlingniveau	255.472	14.381	*	
Totale residuele variantie	285.646			
ICC school	0.036			
ICC klas	0.106			
Aantal scholen	133			
Aantal klassen	205			
Aantal leerlingen	3358			
Verklaarde variantie t.o.v. Model 1 (Cohen's f^2)	0.255			

* = $p < 0.05$

¹⁵¹ Omdat uit de schattingen blijkt dat de werkervaring vanaf meer dan 10 jaar geen verschil meer uitmaakt, zou een spaarzamer model kunnen worden verkregen door deze categorieën samen te nemen. Om twee redenen is hier niet voor gekozen. Ten eerste leek het ons inhoudelijk relevant om de categorieën te blijven onderscheiden. Daarnaast liet het model met alleen de leerkrachtkenmerken (Tabel 9.4) nog wel meer verschillen tussen de groepen zien. Die verschillen zijn in dit model weliswaar verdwenen, maar dat had net zo goed niet gekund. In het gepresenteerde model is dat zichtbaar.

Hoofdstuk 10. Overzicht van databestanden en variabelen

In dit hoofdstuk geven we een beschrijving van de bestanden van de peiling schrijfvaardigheid en de variabelen die in de verschillende bestanden zijn opgenomen. In de eerste paragraaf van dit hoofdstuk geven we een toelichting van de opbouw van de verschillende bestanden. Op basis daarvan kan een potentiële gebruiker bepalen welke bestanden voor hem of haar relevant zijn. In de tweede paragraaf geven we voor het hoofdbestand een overzicht van alle variabelen met een beschrijving van elke variabele. Dit overzicht kan een gebruiker helpen bij de selectie en interpretatie van relevante variabelen. Een overzicht van alle variabelen in de bestanden van de vragenlijstbestanden zijn gegeven in Bijlage 14, 15 en 16. Een overzicht van de variabelen in de bestanden van de interviews met leerkrachten en leerlingen zijn gegeven in Bijlage 17 en 18. Tenslotte worden in Bijlage 19 een overzicht gegeven van de geaggregeerde observatiedata en in Bijlage 20 een overzicht van de leerlingsscores. Bijlagen zijn weergegeven in een apart supplement.

10.1 Opbouw van databestanden

De peiling schrijfvaardigheid heeft een set aan bestanden opgeleverd. Er is gekozen voor een hoofdbestand met een aantal deelbestanden vanwege het zeer grote aantal variabelen dat in de peiling is verzameld. Het samenvoegen van alle variabelen tot één bestand had geleid tot een niet werkbaar breed bestand.

Het belangrijkste bestand is het hoofdbestand. In dit bestand zijn de belangrijkste gegevens samengebracht die in de peiling schrijfvaardigheid zijn verzameld. In het overzicht dat is opgenomen in paragraaf 10.2 is een extra kolom opgenomen die verwijst naar de bron waaruit de variabele is voortgekomen (bijvoorbeeld administratie van aangemelde leerlingen, toetsen, leerlingvragenlijst).

Een belangrijk onderdeel van de verschillende bestanden betreft de identificatie variabelen. Deze identificatie variabelen zijn beschikbaar op verschillende niveaus, namelijk op leerlingniveau in de vorm van 'Combi_LI_Id'. Op schoolniveau is Combi_Sch_ID beschikbaar. Door middel van dit School ID-nummer kunnen gegevens van de schoolvragenlijst gekoppeld worden aan de gegevens van de leerlingen. Ook kan deze schoolidentificatie variabele gebruikt worden voor zogenaamde meerniveau analyse, aangezien deze weergeeft welke leerlingen samen op een school leskrijgen. Het derde niveau van identificatievariabelen betreft de klas (Combi_Klas_id). In het hoofdbestand is gebruik gemaakt van een geanonimiseerde klasnaam. Deze variabele 'klas' kan gekoppeld worden aan de gegevens van de leerlingen en is tevens gebruikt in de meerniveau analyses ter identificatie van niveau 2. Aangezien er slechts één leerkracht per klas de vragenlijst heeft ingevuld valt de leerkrachtidentificatie samen met de klasidentificatie.

10.2 Overzicht van variabelen in het hoofdbestand

Variabele	Label	Bron
Combi_LI_ID	leerling id	Admin.
Combi_Klas_ID	klas id	Admin.
Combi_Sch_ID	school id	Admin.
ngd	Geboortedag	School
ngm	Geboortemaand	School
ngj	Geboortjaar	School
Geslacht	Geslacht	School
dyslexie	Dyslexie	School
lIn_gw_num	Leerlinggewicht	School
advies	advies vo	School
TypeBO	type school	Admin.
OFI_4_WML	vaardigheidsscore	toets
ref_cat	Referentieniveau	toets
refcat2	Referentieniveau in categorieën	toets
V01_MC	Geboren in NL?	leerlingvragenlijst
V02_MC	Taal leren praten	leerlingvragenlijst
V03_MC	Taal ouders onderling	leerlingvragenlijst
V04_MC	Taal met ouders	leerlingvragenlijst
V05_MC	Schrijven op mobiele telefoon per week	leerlingvragenlijst
V07_MC	Schrijven vrije tijd	leerlingvragenlijst
SP7	Schrijfplezier 7 (vraag 19, items.01-.07)	leerlingvragenlijst
SE5	Self Efficacy 5 (vraag 19, items.08-.12)	leerlingvragenlijst
SErvaring	Schrijfervaring buiten de klas om (som	leerlingvragenlijst
Groep	Jaargroep	School
taaltijd_pw	Tijd besteed aan taal, per week	leerkrachtvragenlijst
schrijftijd_pw	Tijd besteed aan schrijven, per week	leerkrachtvragenlijst
Taalcoördinator	Aanwezigheid taalcoördinator op school	Schoolvragenlijst+interne
Sc_invuller	Functie invuller	Schoolvragenlijst
Sc_Q1	In hoeverre is er op uw school sprake van taalbeleid	Schoolvragenlijst
Sc_Q3_1_dich	doorgaande leerlijn va groep 3, dichotoom	Schoolvragenlijst
Sc_Q3_10_dich	systematisch volgen schrijfprestaties, dichotoom	Schoolvragenlijst
Focusschool_brede_facilitering	Focus schrijfonderwijs: faciliteren schijfonderwijs in brede zin	Schoolvragenlijst
gewicht_cat	Percentage gewichtenleerlingen op	DUO
regio	regio	DUO
grootte	schoolgrootte	DUO
den_def	denominatie	DUO
stedelijkheid	stedelijkheid	DUO
formdate	referentiedatum leeftijd	Admin.
age	leeftijd op 1-10-2018	School
dum_sco	heeft leerling wml score en refcat?	Admin.
dum_vrl	bruikbare gegevens uit vragenlijst	Admin.

Vr1_a_mk	geslacht leerkracht	leerkrachtvragenlijst
Vr1_b_mk	Leeftijd leerkracht	leerkrachtvragenlijst
v1c1	Opleiding: reguliere pabo	leerkrachtvragenlijst
v1c2	Opleiding: verkorte pabo	leerkrachtvragenlijst
v1c3	Opleiding: academische pabo	leerkrachtvragenlijst
v1c4	Opleiding: pabo nog niet gehaald	leerkrachtvragenlijst
v1c5	Opleiding: anders	leerkrachtvragenlijst
V1c5_cat	Vr 1C - anders	leerkrachtvragenlijst
Vr1_d_mk	hoe lang als leerkracht in (S)BO	leerkrachtvragenlijst
Vr1_e_mk	hoe lang les aan gp 6/7/8 van (S)BO	leerkrachtvragenlijst
Vr1_f_mk	deelgenomen aan prof.act. mbt schrijfonderwijs	leerkrachtvragenlijst
v2a1	Taal Actief	leerkrachtvragenlijst
v2a2	Taalverhaal.nu	leerkrachtvragenlijst
v2a3	STAAL	leerkrachtvragenlijst
v2a4	Taal op maat	leerkrachtvragenlijst
v2a5	Taal in beeld	leerkrachtvragenlijst
v2a6	Anders	leerkrachtvragenlijst
v2a7	Geen methode	leerkrachtvragenlijst
Vr2_b_mk	naast taalmethode nog andere	leerkrachtvragenlijst
Vr2_c_mk	zelf ontwikkeld materiaal voor Schrijven	leerkrachtvragenlijst
Sc_1_f_rec	deelname aan professionaliseringsactiviteit, REC	leerkrachtvragenlijst
Sc_1g_dich	voldoende kennis en vaardigheid, dich	leerkrachtvragenlijst
Sc_4_1_dich	Integratie taalonderwijs_dich	leerkrachtvragenlijst
Sc_4_2_dich	integratie_andere_vakken_dich	leerkrachtvragenlijst
Sc_9_1_dich	nadruk op doel en publiek	leerkrachtvragenlijst
Sc_9_2_dich	Doel schrijfaanpak, dichotoom	leerkrachtvragenlijst
Sc_14_b_dich	Feedback enkel op leerdoelen, dichotoom	leerkrachtvragenlijst
soorten_teksten_NEW	index soorten teksten, obv gedichotomiseerde items	leerkrachtvragenlijst
soorten_doelen_NEW	index soorten doelen, obv gedichotomiseerde items	leerkrachtvragenlijst
soorten_publiek_NEW	index soorten publiek, obv gedichotomiseerde items	leerkrachtvragenlijst
Sc_15_2_dich	analytisch beoordelen	leerkrachtvragenlijst
Sc_15_3_dich	beoordelen met rubrics	leerkrachtvragenlijst
differentiatie_dich	differentiatie dichotoom, ja/nee (vraag 18)	leerkrachtvragenlijst
Sc_16_1_dich	kennis verschillende tekstsoorten, dichotoom	leerkrachtvragenlijst
Sc_16_4_REC_dich	kennis leestekens en hoofdletters, dichotoom	leerkrachtvragenlijst
schaal_fb_inhoud_mean	schaal feedback op inhoud, gemiddelde	leerkrachtvragenlijst

schaal_fb_correctheid_mean	schaal feedback op correctheid, gemiddelde	leerkrachtvragenlijst
Schaal_aanpak_mean	schaal schrijfaanpak obv gemiddelde	leerkrachtvragenlijst
Schaal_revisie_mean	schaal revisie obv gemiddelde	leerkrachtvragenlijst
Samenwerking_mean	samenwerking tussen leerlingen obv	leerkrachtvragenlijst
Schaal_leren_van_anderen_mean	schaal leren van anderen obv	leerkrachtvragenlijst
Schaal_proceskennis_mean	schaal kennis over schrijfprocessen obv gemiddelde	leerkrachtvragenlijst
schrijven_om_te_leren_mean	schrijven om te leren obv gemiddelde	leerkrachtvragenlijst

Geraadpleegde literatuur

Besluit referentieniveaus Nederlandse taal en rekenen. Staatsblad 2010 265.

Bonset, H. & Hoogeveen, M. (2007). *Schrijven in het basisonderwijs: Een inventarisatie van empirisch onderzoek in het perspectief van leerplanontwikkeling*. Enschede: SLO.

Bonset, H., & Hoogeveen, M. (2015). *Schrijven in het basisonderwijs opnieuw onderzocht: Een inventarisatie van empirisch onderzoek van 2004 tot 2014*, Enschede: SLO.

Cizek G.J. and Bunch. M.B. (2007). *Standard Setting*. Thousand Oaks: Sage.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

De Glopper, K., Minnaert, A.E.M.G., Van Gelderen A.J.S., Pander Maat, H. & Kuijpers, J.M., 2017. *Voorstel peilingsonderzoek Schrijfvaardigheid*. Groningen: Rijksuniversiteit Groningen.

De Miliano, I. (2013). *Lees- en schrijfontwikkeling van vmbo-leerlingen: De rol van betrokkenheid bij lezen en schrijven op school*. Proefschrift. Amsterdam: UVA.

Dekker, B., Krooneman, P.J., Brekelmans, J. & Groenwoud, M. (2012). *Onderwijsinspanningen taal en rekenen in PO, VO en MBO*. Amsterdam: Regioplan.

Espin, C. A. (1999). Criterion-related validity of curriculum-based measures in writing for secondary school students, *Reading & Writing Quarterly*, 15:1, 5-27, <https://doi.org/10.1080/105735699278279>

Evers, A., Sijtsma, K., Lucassen, W. I., & Meijer, R. R. (2010). Het COTAN-beoordelingssysteem voor de kwaliteit van tests herzien. *De Psycholoog*, 45, 48 - 55.

Expertgroep doorlopende leerlijnen taal en rekenen. (2008). *Over de drempels met taal en rekenen: hoofdrapport van de Expertgroep doorlopende leerlijnen taal en rekenen*. Enschede: SLO.

Franssen, H. M. B., & Aarnoutse, A. (2003). Schrijfonderwijs in de praktijk. *Pedagogiek*, 23(3), 185-198.

Hacquebord, H. I., & Lenting-Haan, K. (2012). *Kunnen we de moeilijkheid van teksten meten? Naar concrete maten voor de referentieniveaus*. *Levende Talen Tijdschrift*, 13(2), 14 - 23.

Hoogeveen, M. (2017). *Schrijfvaardigheid in het basisonderwijs. Domeinbeschrijving ten behoeve van peilingsonderzoek*. Enschede: SLO.

Inspectie van het Onderwijs (2012). *Focus op schrijven. Het onderwijs in het schrijven van teksten (stellen)*. Utrecht: Inspectie van het Onderwijs.

Inspectie van het Onderwijs (2014). *De kwaliteit van schoolleiders in het basisonderwijs, speciaal onderwijs en voortgezet onderwijs*. Utrecht: Inspectie van het Onderwijs.

Janssen, T. & D. van Weijen (2017). *Effectief schrijfonderwijs op de basisschool: een didactisch kader ten behoeve van landelijk peilingsonderzoek*. Amsterdam: Universiteit van Amsterdam.

- Keuning, J., Straat, H., and R.C.W. Feskens (2017). The Data-Driven Direct Consensus (3DC) Procedure: A New Approach to Standard Setting. In *Standard Setting – International State of Research and Practices in the Nordic Countries* (Eds: Sigrid Blömeke and Jan-Eric Gustafsson). Springer series Methodology of Educational Measurement and Assessment. pp 263-278).
- Kingston, N. M., Kahl, S. R., Sweeney, K., & Bay, L. (2001). Setting performance standards using the Body of Work method. In G. J. Cizek (Ed.). *Standard setting: Concepts, methods, and perspectives* (pp. 219-248). Mahwah, NJ: Erlbaum.
- Koster, M., Bouwer, R., & Van den Bergh, H. (2017). Professional development of teachers in the implementation of a writing intervention program for elementary students. *Contemporary Educational Psychology*, 49, 1-20. <https://doi.org/10.1016/j.cedpsych.2016.10.002>
- Koster, M., Tribushinina, E., De Jong, P.F., & Van den Bergh, B. (2015). Teaching children to write: A meta-analysis of writing intervention research. *Journal of Writing Research*, 7(2), 249-274.
- Krom, R., Gein, Van der, J., Van der Hoeven, J., Van der Schoot, F., Verhelst, N., Veldhuijzen, N., Hemker, B. (2004). *Balans van het schrijfonderwijs op de basisschool: Uitkomsten van de peiling in 1999: halverwege en einde basisonderwijs en speciaal basisonderwijs*. Arnhem: Citogroep.
- Kuhlemeijer, H., Til, A. van, Hemer, B., Klijn, W. de, Feenstra, H. (2013). *Balans van de schrijfvaardigheid in het basis- en speciaal basisonderwijs in groep 5, groep 8 en de eindgroep SBO*. Arnhem: Cito.
- Lampe, T. & Hemker, B. (2020). *Standaardbepaling referentieniveaus 1F en 1S/2F bij Peil.Schrijfvaardigheid 2019*. Arnhem: Cito.
- Ministerie van Onderwijs, Cultuur en Wetenschap (OCW) (2006). *Kerndoelen primair onderwijs*. Den Haag: Ministerie van Onderwijs, Cultuur en Wetenschap.
- Pajares, F., & Valiante, G. (1997). Influence of self-efficacy on elementary students' writing. *The Journal of Educational Research*, 90(6), 353–360. <https://doi.org/10.1080/00220671.1997.10544593>
- Pander Maat, H. (2002). *Tekstanalyse. Wat teksten tot teksten maakt*. Bussum: Coutinho.
- Rasbash, J., Browne, W. J., Healy, M., Cameron, B. & Charlton, C. (2013). *MLwiN Version 2.27*. Centre for Multilevel Modelling, University of Bristol.
- Reezigt, G.J., Reker, L.T.M., & Spithoff, M. (2019). *Kwaliteit van sturing in het primair onderwijs: bestuurlijk vermogen van schoolbesturen die met krimp te maken hebben*. Groningen: GION onderwijs/onderzoek
- Rietdijk, S., van Weijen, D., Janssen, T., van den Bergh, H., & Rijlaarsdam, G. (2018). Teaching writing in primary education: Classroom practice, time, teachers' beliefs and skills. *Journal of Educational Psychology*, 110(5), 640–663. <https://doi.org/10.1037/edu0000237>
- Rubin, D. B. and Schenker, N. (1986). Multiple imputation for interval estimation from simple random samples with ignorable nonresponse. *Journal of the American Statistical Association*, 81, 366-374.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). Mice: multivariate imputation by chained equations in R. *Journal of Statistical Software*, 45(3), 1-67.

Van Gelderen, A. & Blok, H. (1991). De praktijk van het stelonderwijs in de groepen 7 en 8 van de basisschool; observaties en interviews. *Pedagogische Studiën*, 68(4), 159–175.

Van Gelderen, A., Oosterloo, A., & Paus, H. (2010). Waarheen, waarlangs, tot waar? Schrijven in de basisschool. *Tijdschrift Taal voor opleiders en onderwijsadviseurs*, 1(2), 6-11.

Van der Leeuw, B., Pauw, I., Smits, M., & Van de Ven, P-H. (2012). Schrijven op de pabo; wat weten we uit onderzoek? *Levende Talen Tijdschrift*, 11(2), 30-39.