DEEL C

Achtergrond van de peiling

Doel en werkwijze van de peiling Engels

De peiling Engels 2017/2018

Het is belangrijk om zicht te hebben op de resultaten van leerlingen, de ontwikkeling in die resultaten over de tijd en de relatie van de resultaten met het onderwijsleerproces (bijvoorbeeld aanbod, onderwijstijd en didactisch handelen). Dit geldt niet alleen voor de basisvaardigheden, maar ook voor de vaardigheden die onderdeel zijn van het inhoudsgebied Engels. Sinds 1986 is Engels een verplicht vak in het basisonderwijs in Nederland. Engels speelt bovendien een belangrijke rol in de Nederlandse samenleving en wordt alleen maar belangrijker. Daarom is voldoende aandacht voor de Engelse vaardigheden van leerlingen in het basisonderwijs essentieel, zowel als voorbereiding op het voortgezet onderwijs als voor de verdere toekomst van leerlingen.

Het primaire doel van de peiling Engels einde basisonderwijs was om vast te stellen hoe vaardig leerlingen aan het einde van groep 8 zijn op het gebied van Engelse gespreksvaardigheid, luisteren, lezen en woordenschat. Ook wilden we zicht krijgen op een aantal relevante achtergrondkenmerken van leerlingen en leerkrachten zoals hun houding tegenover Engels en hun buitenschools gebruik van de Engelse taal. Daarnaast wilden we de inhoud en vormgeving van het onderwijs op het gebied van Engels in kaart brengen. Naast een beschrijving hoe basisscholen het onderwijs op het gebied van Engels vormgeven, wilden we nagaan in hoeverre de prestaties van leerlingen aan de ene kant en het onderwijs in Engels aan de andere kant met elkaar samenhangen. Ten slotte wilden we een vergelijking maken met het vorige peilingsonderzoek van Cito dat in 2012 plaatsvond (Geurts & Hemker, 2013). De vaardigheden lezen, luisteren en woordenschat zijn daarom deels op dezelfde manier gemeten als in 2012. De gespreksvaardigheid (interactie) van leerlingen is in 2012 niet gemeten, waardoor we voor deze vaardigheid geen vergelijking konden maken. Wel is in 2012 de spreekvaardigheid (monoloog) van leerlingen getoetst. Om een trend te kunnen rapporteren, is daarom in het huidige peilingsonderzoek een taak voor spreekvaardigheid afgenomen die ook in 2012 is gebruikt.

De peiling Engels einde basisonderwijs vond plaats in het schooljaar 2017/2018. Een consortium bestaande uit GION Onderwijs/Onderzoek, Cito, de Hogeschool van Amsterdam, de afdeling Applied Linguistics en de lerarenopleiding van de Rijksuniversiteit Groningen (RUG) voerde het peilingsonderzoek in opdracht van de Inspectie van het Onderwijs uit.
Zij verzamelden de data in de periode van half januari 2018 tot en met eind maart 2018 op een representatieve steekproef van 95 Nederlandse basisscholen met 2.088 leerlingen. In paragraaf 1.4 van dit hoofdstuk gaan we beknopt in op de steekproef en de representativiteit hiervan.

Van kerndoelen naar toetsbare doelen

De Kerndoelen Primair Onderwijs (Greven & Letschert, 2006) vormen het wettelijk kader voor het domein Engels in het primair onderwijs. Voor Engels zijn er vier kerndoelen, de kerndoelen 13 tot en met 16 (zie onderstaand kader). Ze sluiten aan bij het doel van Engels in het primair onderwijs zoals verwoord in de karakteristiek voorafgaand aan de kerndoelen: het leggen van een eerste basis om te kunnen communiceren met moedertaalsprekers of anderen die buiten de school Engels spreken.

tabel 1.1a kerndoelen engels

- Kerndoel 13: De leerlingen leren informatie te verwerven uit eenvoudige gesproken en geschreven Engelste teksten.
- Kerndoel 14: De leerlingen leren in het Engels informatie te vragen of geven over eenvoudige onderwerpen en zij ontwikkelen een attitude waarbij ze zich durven uit te drukken in die taal.
- Kerndoel 15: De leerlingen leren de schrijfwijze van enkele eenvoudige woorden over alledaagse onderwerpen.
- Kerndoel 16: De leerlingen leren om woordbetekenissen en schrijfwijzen van Engelse woorden op te zoeken met behulp van het woordenboek.

De omschrijvingen van de kerndoelen zijn globaal en geven aan waar het onderwijs in Engels aan kinderen zich op dient te richten. Ze geven geen indicatie van het niveau dat leerlingen aan het einde van het primair onderwijs zouden moeten beheersen. Dat is anders voor het ERK dat in 2001 is gepubliceerd en een raamwerk van niveaubeschrijvingen voor moderne vreemde talen bevat[36]. Het ERK is sectorafhankelijk en kan daarom gebruikt worden voor lesgeven en toetsing van primair onderwijs tot en met hoger onderwijs. Het ERK omschrijft zes niveaus van beheersing waarbij A1 het laagste niveau is en C2 het hoogste.

tabel 1.1b beheersingsniveaus erk

A1/A2	basisgebruiker
B1/B2	onafhankelijke gebruiker
C1/C2	vaardige gebruiker

Het ERK bevat beheersingsniveaus voor de vaardigheden luisteren, lezen, schrijven, spreken (monoloog) en gesprekken voeren (interactie) en is gebaseerd op reallife communicatieve taken. Het beschrijft vaardigheidsniveaus van taalbeheersing in termen van:
- taalomvang: wat moet je kunnen, in welke contexten en voor welke doelen?
- taalcomplexiteit en mate van correctheid: hoe goed moet je het kunnen?

Per vaardigheid en per niveau bevat het ERK globale beschrijvingen en gedetailleerde *can do*-descriptoren. *Can do*-descriptoren zijn korte, kernachtige uitspraken over wat iemand kan in de betreffende taal. Hierbij is het doel duidelijk gericht op taal als communicatiemiddel.

SLO werkte de ERK-descriptoren in Nederland verder uit in de publicatie Taalprofielen (SLO, 2015) op basis van de Nederlandse vertaling van het ERK. Vooralsnog zijn scholen in Nederland niet verplicht om met het ERK te werken. Wel vermelden veel methoden het ERK-niveau waarop ze zich richten en baseren ze hun toetsen op het ERK.

[36] in 2018 verscheen een geactualiseerde versie van het erk, de zogenaamde companion volume. aangezien de constructie van meetinstrumenten en de dataverzameling voor dit peilingsonderzoek voor die tijd plaatsvonden, laten we deze geactualiseerde versie van het erk hier buiten beschouwing.

tabel 1.1c voorbeeld uitwerking globale beschrijving en *can do*-descriptor gesprekken

VOORBEELDEN (SLO, 2015)

Globale beschrijving gesprekken A1

Ik kan deelnemen aan een eenvoudig gesprek, wanneer de gesprekspartner bereid is om zaken in een langzamer spreektempo te herhalen of opnieuw te formuleren en mij helpt bij het formuleren van wat ik probeer te zeggen. Ik kan eenvoudige vragen stellen en beantwoorden die een directe behoefte of zeer vertrouwde onderwerpen betreffen.

***Can do*-descriptor gesprekken A1**

Kan zichzelf en anderen voorstellen en reageren als iemand voorgesteld wordt. Kan op een eenvoudige manier groeten en afscheid nemen.

Voor het peilingsonderzoek Engels einde basisonderwijs maakte SLO een domeinbeschrijving (Rose, 2016). In deze domeinbeschrijving staan de wettelijke eisen voor Engels, hoe deze terugkomen in het onderwijs en welke resultaten eerdere peilingen op het gebied van Engels opleverden. Ook specificeert SLO in de domeinbeschrijving, op basis van een veldraadpleging met domeinexperts, schoolleiders en leerkrachten, welke vaardigheden minimaal onderdeel moeten zijn van het peilingsonderzoek. De deelnemers aan de veldraadpleging waren het er unaniem over eens dat gespreksvaardigheid, woordenschat en luistervaardigheid de hoogste prioriteit hebben. Met name gespreksvaardigheid achtten zij belangrijk, omdat deze vaardigheid een "realistisch beeld geeft van de taalvaardigheid van een leerling". Ook veronderstelden de deelnemers aan de veldraadpleging dat hier door de communicatieve aanpak van methoden veel aandacht naar uit gaat in het onderwijs. Zij achtten leesvaardigheid met name relevant in relatie tot de overgang naar het voortgezet onderwijs. Daarnaast vonden de deelnemers aan de veldraadpleging het wenselijk de peilingsresultaten te koppelen aan de ERK-niveaus.

De instrumenten

De prioritering van vaardigheden door de experts en de wens om de resultaten te koppelen aan ERK-niveaus en –beschrijvingen waren leidend voor de constructie van het instrumentarium voor het peilingsonderzoek. De instrumenten zijn gericht op de communicatieve en sociale functie van taal, met andere woorden op het taalgebruik van leerlingen in zo authentiek mogelijke situaties. De toetsen bevatten vragen (of in het geval van gespreksvaardigheid taken) op de ERK-niveaus A1, A2 en B1 om zo aan te sluiten op de verschillen tussen leerlingen in Engelse taalvaardigheid.

Vanwege de tijd die het afnemen van de instrumenten van de leerlingen vroeg, maakten niet alle leerlingen alle onderdelen van het instrumentarium. De afname startte voor alle leerlingen met het invullen van de leerlingvragenlijst, zodat de leerlingen ongeacht hun waargenomen prestatie op de toetsen hun mening gaven over hun attitude tegenover Engels en hun ervaring met de Engelse taal. Na de vragenlijst maakten de leerlingen de toetsen voor lezen, luisteren en woordenschat. De totale hoeveelheid opgaven voor deze vaardigheden was dusdanig groot dat leerlingen een toetsboekje met een selectie van de opgaven kregen. Om de resultaten over de toetsboekjes heen samen te kunnen nemen, overlapten de verschillende toetsboekjes elkaar inhoudelijk. Aan de klassikale afname van de toetsen voor lezen, luisteren en woordenschat namen telkens twee leerlingen geen deel. Zij voerden op dat moment buiten de klas in een separate ruimte de taken voor spreek- en gespreksvaardigheid uit. Zo namen in totaal maximaal zes leerlingen per klas deel aan de taken voor spreek- en gespreksvaardigheid. Deze leerlingen zijn random geselecteerd, mits er sprake was van actieve toestemming van hun ouders in verband met de video-opnamen. Die waren noodzakelijk voor de beoordeling van de spreek- en gesprekprestaties. De afname van alle instrumenten duurde voor leerlingen in totaal twee uur.

Gespreksvaardigheid

Aangezien het belangrijk is voor leerlingen om mondeling te kunnen communiceren in het Engels, is gespreksvaardigheid gemeten met behulp van drie praktische taken. Deze taken sloten zoveel mogelijk aan bij de belevingswereld van leerlingen en bootsten gesprekssituaties na waarin zij mogelijk terechtkomen. Tijdens de uitvoering van de taken voerden de leerlingen in tweetallen een (gestructureerd) gesprek. Om drie redenen is ervoor gekozen leerlingen met elkaar in gesprek te laten gaan in plaats van met de testleider, zoals meestal toegepast bij het meten van gesprekvaardigheid:

- Een gesprek voeren met een medeleerling sluit beter aan bij authentieke situaties waarin leerlingen mogelijk terechtkomen. In authentieke situaties is interactie met een leeftijdsgenoot immers realistischer dan met een (onbekende) volwassene. Bovendien is een gesprek met een medeleerling naar verwachting meer ontspannen en spontaner dan een gesprek met de testleider als onbekende volwassene.
- Een gesprek voeren met een medeleerling sluit beter aan bij een klassensituatie waarin alle leerlingen gestimuleerd zouden moeten worden om zoveel mogelijk te praten.
- Door leerlingen met elkaar gesprekken te laten voeren, konden zoveel mogelijk leerlingen in de beschikbare tijd deelnemen aan het onderdeel gespreksvaardigheid.

Aangezien de taken opliepen in moeilijkheid, besloot de testleider na elke taak of een of beide leerling(en) door mocht(en) naar de volgende taak. Uitgangspunt daarbij was om leerlingen door te laten gaan, tenzij er sterke twijfels waren. Als duidelijk was dat een van beide leerlingen duidelijk onder niveau presteerde en de andere leerling niet, werd het gesprek gevoerd tussen de leerling die het volgende niveau mogelijk wel aankon en de testleider. Die was inmiddels al enigszins vertrouwd voor de leerling. Niet alle leerlingen deden dus alle drie gesprekstaken. De leerlingen die het niveau niet aankonden, werkten ondertussen in een speciaal samengesteld activiteitenboekje met Engelse oefeningen om hen niet het gevoel te geven te hebben gefaald.

De gesprekken zijn op video opgenomen en achteraf met behulp van gestandaardiseerde beoordelingsformulieren beoordeeld op drie aspecten: inhoud, kwaliteit en vloeiendheid. De beoordeling op inhoud betrof de mate waarin de leerling de taak inhoudelijk uitvoerde. De kwaliteit werd vastgesteld op basis van woordenschat, grammatica en samenhang. Bij vloeiendheid was het tempo een belangrijke indicator. Per aspect waren er meerdere beoordelingscriteria. Elk criterium werd gescoord op een driepuntsschaal (0-onvoldoende, 1-voldoende en 2-goed). Daarnaast zijn ook criteria gedefinieerd om de gesprekken op interactie te beoordelen.

Om de uniformiteit van de afname te bevorderen, startte de testleider altijd met een gestandaardiseerde instructie op film waarin het rollenspel werd uitgelegd. Na deze korte instructiefilm kwam een kort filmpje waarin twee studenten een voorbeeldtaak uitvoerden, inclusief de geïnstrueerde gedragingen als een leerling het niet weet (elkaar om hulp vragen). In de instructiefilm werd leerlingen gevraagd alleen Engels te spreken en zo nodig gebaren of andere omschrijvingen te gebruiken. Alleen als de twee leerlingen er samen echt niet uit kwamen, schoot de testleider te hulp door bijvoorbeeld te vragen of de leerling het wellicht in andere woorden kon omschrijven. In het uiterste geval kon de testleider aangeven dat een leerling het in het Nederlands mocht zeggen, zodat de andere leerling kon reageren. Testleiders beschikten hiertoe over een protocol met daarin gestandaardiseerde instructies, richtlijnen en praktische handreikingen. Ook zijn zij voorafgaand aan de afname getraind. Bovendien zijn alle testleiders geselecteerd op basis van hun kennis van en ervaring met het ERK.

De leerlingen voerden voor gespreksvaardigheid de volgende drie taken uit:

- Taak 1 (Wie ben ik?) is ontworpen als taak op niveau A1, met elementen op niveau A2. De taak is een rollenspel, waarbij leerlingen ieder een praatplaat over een fictieve persoon kregen. Aan de hand van vraagkaartjes probeerden zij informatie van de ander te krijgen en informatie over hun fictieve persoon te geven. Binnen de taak spreken de leerlingen met elkaar over eenvoudige, alledaagse onderwerpen.
- Taak 2 (Wat spreken we af?) is ontworpen als taak op niveau A2, met elementen die een gesprek op een iets hoger niveau dan A2 (A2+) uitlokten. De leerlingen blijven dezelfde fictieve persoon als bij de eerste taak en proberen een afspraak te maken (een alledaagse bezigheid). De leerlingen krijgen ieder een eigen agenda en ontvangen beurtelings kaartjes met daarop vragen/antwoorden. Gezien hun volle agenda moeten ze onderhandelen om tot de afspraak te komen. Hierbij is het aanbrengen van samenhang tussen zinnen noodzakelijk.

- Taak 3 (Wat vind je ervan?) is ontworpen als taak op niveau A2+/B1. De leerlingen gaan in debat over de vraag of het goed of slecht is om veel te gamen. De ene leerling is voor, de ander tegen. Deze posities moeten zij onderbouwen aan de hand van informatie op kaartjes, die door de testleider worden verdeeld.

Elke taak bestond uit verschillende deeltaken. De deeltaken zorgden voor structurering van het gesprek en waren zodanig samengesteld dat beide leerlingen minimaal een beurt hadden en zowel initiator als reactor waren. Het gesprek bij taak 1 bestond uit achttien deeltaken, zoals vragen naar naam, vragen naar leeftijd en vertellen over familie. Taak 2 bestond uit vijf deeltaken, zoals voorstellen om af te spreken. Taak 3 bestond uit zes deeltaken, zoals het geven van argumenten voor de stelling. De *can do*-descriptoren binnen het ERK vormden de basis voor de drie taken. In tabel 1.3a zijn de *can do*-descriptoren van elke taak opgenomen.

tabel 1.3a beschrijving van de *can do*-descriptoren per gesprekstaak

Taak	ERK-niveau	*Can do*-descriptor
1. Wie ben ik?	a1	• kan zichzelf en anderen voorstellen en reageren als iemand voorgesteld wordt. kan op een eenvoudige manier groeten en afscheid nemen. • kan eenvoudige vragen stellen en beantwoorden en eenvoudige uitspraken doen en beantwoorden op het gebied van primaire behoeften of over zeer vertrouwde, concrete onderwerpen. • kan vragen beantwoorden en stellen over zichzelf en over anderen, waar zij wonen, wie zij kennen, wat zij bezitten. • kan zeggen wat hij of zij leuk of lekker vindt en wat niet.
	a2	• kan beperkte informatie uitwisselen over vertrouwde en alledaagse zaken van praktische aard. • kan beperkte informatie uitwisselen over vertrouwde en alledaagse zaken van praktische aard. • kan informatie van persoonlijke aard vragen en geven.
2. Wat spreken we af?	a2	• kan eenvoudige alledaagse beleefdheidsvormen gebruiken om anderen te begroeten en aan te spreken. • kan uitnodigingen doen, op uitnodigingen ingaan of deze afslaan, suggesties opperen. • kan op eenvoudige wijze voorkeur en mening uitdrukken over vertrouwde alledaagse onderwerpen. • kan communiceren over eenvoudige en alledaagse taken in eenvoudige bewoordingen om dingen te vragen en te verschaffen, eenvoudige informatie te verkrijgen en te bespreken wat er vervolgens moet gebeuren. • kan afspraken maken. • kan vragen stellen en beantwoorden over wat men op het werk en in de vrije tijd doet.
	a2+	• kan te kennen geven het (on)eens te zijn met anderen. • kan deelnemen aan korte gesprekken over belangwekkende onderwerpen in een alledaagse context. • kan relevante informatie uitwisselen en zijn of haar mening geven over praktische problemen wanneer dat rechtstreeks gevraagd wordt, mits hij of zij enige hulp krijgt bij het formuleren en indien nodig om herhaling van belangrijke punten kan vragen. • kan bespreken wat er vervolgens moet gebeuren, daarbij suggesties doen en beantwoorden, en aanwijzingen vragen en geven. • kan vragen stellen en beantwoorden over gewoonten en routines. • kan vragen stellen en beantwoorden over tijdverdrijf en vroegere of toekomstige activiteiten.
3. Wat vind jij ervan?	a2+	• kan te kennen geven het (on)eens te zijn met anderen. • kan deelnemen aan korte gesprekken over belangwekkende onderwerpen in een alledaagse context.
	b1	• kan gevoelens uiten en op gevoelens van anderen reageren. • kan persoonlijke standpunten, commentaar en meningen geven of er om vragen in discussies over belangwekkende onderwerpen. • kan zijn of haar meningen en reacties overbrengen met betrekking tot oplossingen voor problemen of praktische vragen. • kan op beleefde wijze een overtuiging, mening, instemming en afkeuring uitdrukken. • kan een standpunt helder overbrengen, maar heeft er moeite mee om in debat te gaan. • kan deelnemen aan een alledaagse formele discussie over vertrouwde onderwerpen die gaan om de uitwisseling van feitelijke informatie, het ontvangen van en reageren op aanwijzingen of het bespreken van oplossingen voor praktische problemen. • kan zijn of haar mening geven en reacties overbrengen met betrekking tot mogelijke oplossingen van problemen en het nemen van praktische beslissingen, en daarbij beknopte redenen en verklaringen geven.

Naast de individuele prestatie is ook de interactie tussen de leerlingen beoordeeld. Deze beoordeling betrof zowel de mate waarin de leerlingen als paar een gesprek konden voeren als de mate waarin dat gesprek in het Engels was. Ook de interactie werd op basis van de videobeelden beoordeeld, maar door een andere beoordelaar dan diegene die de individuele gespreksvaardigheid van de betreffende leerlingen beoordeelde. Bij het beoordelen van de interactie gaven beoordelaars een score voor het geheel aan verbaal en non-verbaal gedrag dat de gesprekspartners tijdens de verschillende deeltaken lieten zien. Zij scoorden de aanwezigheid van interactie door een keuze te maken voor een van de volgende categorieën:

1. volledig Engelstalig;
2. deels Engelstalig, deels non-verbaal;
3. deels Engelstalig, deels Nederlandstalig;
4. deels Engelstalig, deels Nederlandstalig, deels non-verbaal;
5. deels Nederlandstalig, deels non-verbaal;
6. volledig non-verbaal;
7. volledig Nederlandstalig;
8. geheel afwezig (niet communicatief; stil; amper beweging).

Op basis van de beoordeling zijn verschillende typen interactiescores berekend. Ten eerste een 'totale interactiescore' waarbij de paren punten kregen voor elk type interactie, behalve wanneer de interactie volledig afwezig was (categorie 1 tot en met 7). Daarnaast een 'milde interactiescore' waarbij de paren punten kregen voor alle interactie die zij volledig of deels in het Engels uitvoerden (categorie 1 tot en met 4). Tot slot een 'strenge interactiescore' waarbij de paren alleen een score kregen voor volledig Engelstalige interactie (categorie 1).

In totaal zijn 577 leerlingen van 94 scholen geselecteerd om de gesprekstaken uit te voeren. Daarvan voerde 98 procent taak 1 uit, 79 procent ook taak 2 en 26 procent voerde vervolgens ook taak 3 uit. Uiteindelijk zijn de scores voor de aspecten inhoud, kwaliteit en vloeiendheid op de 3 taken omgezet naar 1 totaalscore voor gespreksvaardigheid. Daarmee is voor 561 leerlingen van 94 scholen een score voor gespreksvaardigheid beschikbaar. Van 283 leerlingparen van 94 scholen is ook de interactie beoordeeld.

Spreekvaardigheid

Spreekvaardigheid is uitsluitend getoetst met het oog op de trendvergelijking met de vorige peiling. Om de ontwikkeling van de spreekvaardigheid van leerlingen in groep 8 vast te stellen, is net als in de peilingsonderzoeken in 2006 en 2012 de beeldverhalentaak De brug gebruikt. In deze taak moest de leerling in het Engels een verhaal vertellen op basis van een serie afbeeldingen die hij te zien kreeg. De spreekprestatie van leerlingen werd beoordeeld aan de hand van het beoordelingsmodel uit 2012.

De beeldverhalentaak is in het huidige peilingsonderzoek afgenomen bij 297 leerlingen van 94 scholen. Van deze 297 leerlingen waren de prestaties van 11 leerlingen niet te beoordelen. Bij 1 leerling had dit te maken met de slechte opnamekwaliteit en bij 10 andere leerlingen omdat zij bij geen van de plaatjes een te beoordelen antwoord gaven.

Luistervaardigheid

De Engelse luistervaardigheid is in de peiling gemeten met een toets waarin leerlingen een of meerdere vragen bij audio(visuele) fragmenten beantwoordden. Deze fragmenten sloten wat betreft medium en onderwerp zoveel mogelijk aan bij de belevingswereld van leerlingen en hadden betrekking op het dagelijks leven, de school of opleiding of het publieke domein. Gezien het streven naar zo authentiek mogelijke taken en de ontwikkeling in de maatschappij op audiovisueel gebied, bestond ongeveer driekwart van de opgaven uit opgaven bij audiovisuele fragmenten en ruim een kwart van de opgaven uit opgaven bij zuivere audiofragmenten.

In totaal waren er 48 opgaven, waarvan 20 opgaven ook tijdens het peilingsonderzoek in 2012 zijn voorgelegd (zogenaamde ankeropgaven). Alle opgaven waren meerkeuzeopgaven. Elke leerling maakte in totaal 16 opgaven, waarvan 9 audio-opgaven en 7 audiovisuele opgaven. Alle opgaven waren in het Nederlands geformuleerd om te voorkomen dat de vraagstelling de moeilijkheid van de opgaven zou kunnen beïnvloeden.

Binnen het ERK worden de niveaus van luisteren gedefinieerd aan de hand van luisterstrategieën en de tekstkenmerken woordgebruik en zinsbouw, tempo en articulatie en tekstlengte (SLO, 2015).
De niveaubeschrijving richt zich op het begrip van gesproken teksten, onderverdeeld naar:

1. een moedertaalspreker als gesprekspartner begrijpen;
2. gesprekken tussen moedertaalsprekers begrijpen;
3. luisteren als lid van het aanwezige publiek[37];
4. luisteren naar mededelingen en instructies;
5. kijken en luisteren naar audiovisuele media.

In de luistertoets zijn alleen 'luisteren naar mededelingen en instructies' en 'kijken en luisteren naar audiovisuele media' gepeild. De *can do*-descriptoren waarop de toets betrekking had, staan in tabel 1.3b.

tabel 1.3b overzicht van de *can do*-descriptoren in de luistertoets

ERK-niveau	Deelvaardigheid	*Can do*-descriptoren
A1	luisteren naar mededelingen in instructies	• kan tot hem of haar gerichte vragen en opdrachten verstaan en korte, eenvoudige aanwijzingen volgen. • kan in korte, duidelijk gesproken teksten getallen en bekende woorden verstaan die gericht zijn aan de luisteraar. • kan korte, eenvoudige waarschuwingen begrijpen die gericht zijn aan de luisteraar.
	kijken en luisteren naar audiovisuele media	• kan het onderwerp bepalen van korte kijk-/luisterteksten.
A2	luisteren naar mededelingen in instructies	• kan in vertrouwde situaties eenvoudige feitelijke informatie begrijpen. • kan een korte uitleg begrijpen. • kan aanwijzingen begrijpen over de werking van vertrouwde apparaten, mits het apparaat voorhanden is.
	kijken en luisteren naar audiovisuele media	• kan de wezenlijke informatie verstaan en onderscheiden in korte opgenomen passages over voorspelbare alledaagse zaken. • kan zich een indruk vormen van de belangrijkste inhoud in een feitelijk nieuwsbericht op televisie. • *Kan het belangrijkste punt herkennen van nieuwsberichten op televisie waarin verslag wordt gedaan van gebeurtenissen, ongelukken en dergelijke en waarin het beeld het commentaar ondersteunt (A2+).*
B1	luisteren naar mededelingen en instructies	• kan eenvoudige technische informatie begrijpen. • kan gedetailleerde aanwijzingen volgen.
	kijken en luisteren naar audiovisuele media	• kan de hoofdpunten verstaan van nieuwsberichten op de radio en van eenvoudiger opgenomen materiaal over vertrouwde onderwerpen. • *Kan inhoudelijke informatie van het meeste opgenomen of uitgezonden geluidsmateriaal over onderwerpen van persoonlijk belang verstaan (B1+).* • kan de hoofdpunten vatten van televisieprogramma's over vertrouwde onderwerpen wanneer deze betrekkelijk langzaam en helder worden gepresenteerd. • *Kan een groot deel begrijpen van veel televisieprogramma's over onderwerpen van persoonlijk belang, zoals vraaggesprekken, korte voordrachten en nieuwsverslagen, wanneer deze betrekkelijk langzaam en helder worden gepresenteerd (B1+).*

In totaal maakten 1.819 leerlingen van 95 scholen de toets voor luistervaardigheid.

[37] voor deze taak zijn er nog geen niveaubeschrijvingen.

Leesvaardigheid

De leesvaardigheid in het Engels is gemeten met een toets waarin leerlingen korte en wat langere Engelstalige teksten en berichten moesten lezen. Daarna beantwoordden zij hierover een of meerdere vragen. De leesteksten sloten, net zoals bij luistervaardigheid, zoveel mogelijk aan bij de belevingswereld van leerlingen uit groep 8, zowel wat betreft onderwerp als medium (bijvoorbeeld sociale mediaberichten). De teksten hadden betrekking op het dagelijks leven, de school of opleiding of het publieke domein.

In totaal waren er 51 opgaven, waarvan 17 opgaven die ook tijdens het peilingsonderzoek in 2012 zijn voorgelegd (zogenaamde ankeropgaven). Alle opgaven waren meerkeuzeopgaven. Elke leerling maakte in totaal 17 opgaven. Net als bij luistervaardigheid waren alle opgaven in het Nederlands geformuleerd om te voorkomen dat de moeilijkheid van de opgaven veroorzaakt zou kunnen worden door de vraagstelling in plaats van de tekst.

Binnen het ERK worden de niveaus van lezen gedefinieerd aan de hand van leesstrategieën en de tekstkenmerken woordgebruik en zinsbouw, tekstindeling en tekstlengte (SLO, 2015). De niveaubeschrijving richt zich op het begrip van geschreven teksten, onderverdeeld naar het doel van de leestaak:

1. correspondentie lezen;
2. oriënterend lezen;
3. lezen om informatie op te doen;
4. lezen van instructies.

Tabel 1.3c geeft per niveau en leesdoel weer op welke *can do*-descriptoren de opgaven in de leestoets een beroep deden.

tabel 1.3c overzicht van de *can do*-descriptoren in de leestoets

ERK-niveau	Deelvaardigheid	*Can do*-descriptoren
A1	correspondentie lezen	• kan korte, eenvoudige mededelingen begrijpen, bijvoorbeeld via sociale media of op brief- of ansichtkaarten. • kan voorgedrukte kaarten begrijpen met standaardboodschappen.
	oriënterend lezen	• kan een korte standaardmededeling lezen. • kan dingen opzoeken in of kiezen uit een lijst. • kan eenvoudige informatie op een poster, mededelingenbord of in een brochure lezen.
	lezen om informatie op te doen	• kan zich een idee vormen van de inhoud van een korte tekst die waar mogelijk visueel ondersteund wordt. • kan in korte informatieve teksten informatie over personen en plaatsen begrijpen.
	instructies lezen	• kan zeer eenvoudige, korte en goed gestructureerde instructies begrijpen.

ERK-niveau	Deelvaardigheid	*Can do*-descriptoren
A2	correspondentie lezen	• kan een korte, eenvoudige (standaard)brief of e-mail begrijpen. • *Kan elementaire soorten standaardbrieven, faxen (inlichtingen, bestellingen, bevestigingsbrieven, enzovoort) of een algemene kennisgeving over vertrouwde onderwerpen begrijpen (A2+).*
	oriënterend lezen	• kan specifieke informatie vinden en begrijpen in eenvoudig, alledaags materiaal. • kan eenvoudige advertenties met weinig afkortingen begrijpen. • kan in lijsten, overzichten en formulieren specifieke informatie vinden en begrijpen. • kan veelvoorkomende borden en mededelingen begrijpen.
	lezen om informatie op te doen	• kan specifieke informatie begrijpen in eenvoudige teksten. • kan de hoofdlijn begrijpen van eenvoudige teksten in een tijdschrift, krant of op een website. • kan korte, beschrijvende teksten over vertrouwde onderwerpen begrijpen. • kan door meelezen eenvoudig audiovisueel materiaal begrijpen.
	instructies lezen	• kan eenvoudige, korte en goed gestructureerde instructies begrijpen. • *Kan regels en bepalingen begrijpen, wanneer deze in eenvoudige taal zijn gesteld (A2+).*
B1	correspondentie lezen	• kan persoonlijke brieven, e-mails en vormen van sociale media voldoende begrijpen om met iemand te kunnen corresponderen. • kan een eenvoudige formele brief of e-mail voldoende begrijpen om adequaat te kunnen reageren.
	oriënterend lezen	• kan relevante informatie vinden en begrijpen in brochures en korte officiële documenten op internet of in andere media. • *Kan in langere teksten op internet of in andere media informatie zoeken over thema's binnen het eigen interessegebied (B1+).*
	lezen om informatie op te doen	• kan belangrijke feitelijke informatie begrijpen in korte verslagen en artikelen. • kan eenvoudige jeugdliteratuur lezen. • kan door meelezen alledaags audiovisueel materiaal begrijpen. • *Kan hoofdthema en belangrijkste argumenten begrijpen in eenvoudige teksten in tijdschriften, kranten of op internet (B1+).*
	instructies lezen	• kan helder geschreven, ondubbelzinnige instructies begrijpen.

In totaal maakten 1.860 leerlingen van 95 scholen de toets voor leesvaardigheid.

Woordenschat

Voor de ontwikkeling van de woordenschattoets was het CATTS-model (Laufer & Goldstein, 2004) het uitgangspunt. Dit model onderscheidt actieve (productieve) en passieve (receptieve) kennis. Actief gaat daarbij om het produceren van de juiste woordvorm vanuit de gegeven betekenis en passief om het produceren van de juiste betekenis vanuit de gegeven vorm. Aan dit onderscheid tussen actief en passief voegt het CATTS-model ook onderscheid tussen het oproepen en het herkennen van woorden toe. Dit leidt tot de vier niveaus van woordenschatverwerving zoals opgenomen in tabel 1.3d. In de peiling is het onderscheid in de vier niveaus verwerkt in de opgaven voor woordenschat[38].

[38] in de praktische uitvoering van de woordenschattaken is op twee belangrijke punten afgeweken van de catts-methode. ten eerste is een woord slechts in een van de vier niveaus uit tabel 1.3d getoetst en niet in alle vier niveaus. daarnaast is er geen sprake van opgaven die volledig in de tweede taal (in dit geval engels) gesteld zijn. in de peiling is hiervan afgeweken omdat de moeilijkheid niet in de bevraging (de stam van de vraag of de alternatieven) diende te zitten, maar in het bevraagde woord.

tabel 1.3d niveaus van woordenschatverwerving

	Oproepen	Herkennen van woorden
Actief (productief)	productief oproepen	productief herkennen
Passief (receptief)	passief oproepen	passief herkennen

Daarnaast was het uitgangspunt dat de toets een bij de doelgroep aansluitende range in moeilijkheid van de te selecteren woorden had. Hoewel formeel gezien losse woorden niet op de taalvaardigheidsniveaus van het ERK-niveau kunnen worden ingeschaald, is het wel mogelijk om met behulp van frequentielijsten en databases een inschatting te maken van de woordkennis van leerlingen. Om binnen de vier niveaus uit tabel 1.3d rekening te kunnen houden met de verschillende taalvaardigheidsniveaus van leerlingen, is voor productieve woordenschat gebruik gemaakt van het English Vocabulary Profile (EVP)[39]. In het EVP is op grond van een zeer grote database van tweede-taalgebruikers van het Engels een betrouwbaar woordenschatprofiel afgeleid voor elk van de ERK-niveaus. In dit peilingsonderzoek is de beheersing van de productieve woordenschat gepeild voor items die zijn geconstrueerd op de ERK-niveaus A1, A2 en B1, op basis van het EVP. Voor receptieve woordenschat is gebruik gemaakt van een aanpassing van een bestaande toets, de New Vocabulary Levels Test-level 1 (McLean & Kramer, 2015). Deze test is gekoppeld aan een corpus van zowel Brits- als Amerikaans-Engelse woorden. Level 1, het uitgangspunt voor de receptieve woordenschat in dit peilingsonderzoek, bevat 1.000 hoogfrequente woorden waaruit geput is voor de constructie van de opgaven voor receptieve woordenschat.

In totaal waren er 110 opgaven, waarvan 33 opgaven die ook tijdens het peilingsonderzoek in 2012 zijn voorgelegd (zogenaamde ankeropgaven). 70 opgaven waren productieve opgaven (waarvan 21 oproepen van woorden en 49 herkennen van woorden) en 40 receptieve opgaven (waarvan 6 oproepen van woorden en 34 herkennen van woorden). Elke leerling maakte in totaal 36 of 37 opgaven, afhankelijk van de toetsversie. De opgaven over de receptieve woordkennis waren meerkeuzevragen. De productieve opgaven waren open opgaven. Bij de beoordeling daarvan is op twee aspecten gelet. Ten eerste of leerlingen het woord actief konden benoemen en daarnaast of zij het woord goed hadden gespeld. Antwoorden werden dus alleen goed gerekend als deze ook correct waren gespeld. Hiervoor is gekozen omdat onderzoek laat zien dat de betrouwbaarheid bij soepele beoordelingswijzen lager is (Westhoff, 1981).

In totaal is voor 1.834 leerlingen van 95 scholen een score voor woordenschat beschikbaar.

Vragenlijst voor leerlingen

De leerlingen die deelnamen aan dit peilingsonderzoek, vulden een vragenlijst in. Deze vragenlijst is grotendeels gebaseerd op de vragenlijst die in het peilingsonderzoek van 2012 is gebruikt. Van alle 2.088 leerlingen vulden 2.075 leerlingen van 95 scholen de leerlingvragenlijst in. Dat is 99,4 procent van de leerlingen die aan het peilingsonderzoek deelnamen. De leerlingvragenlijst bevatte vragen over:

- houding tegenover Engels (7 items attitude en 5 items *linguistic self-confidence*);
- buitenschools gebruik van het Engels (15 items);
- ingeschatte eigen vaardigheid voor lezen, luisteren, spreken en schrijven (19 items);
- thuistaal: de taal die leerlingen thuis spreken, de taal die zij met hun vader en moeder spreken en de herkomst van hun vader en moeder (5 items).

We gebruikten de gegevens uit de vragenlijst om een beschrijving te geven van kenmerken van de leerlingen die in het licht van hun Engelse taalvaardigheid relevant zijn. Waar mogelijk vergeleken we de resultaten in dit peilingsonderzoek met de resultaten van het peilingsonderzoek dat in 2012 plaatsvond. Dat was alleen mogelijk voor de afzonderlijke vragen, omdat er destijds geen schalen zijn geconstrueerd voor attitude, *linguistic self-confidence*, buitenschoolse gebruik en de inschatting van de eigen vaardigheden. Tot slot keken we naar de samenhang tussen onderwijsleerproces en Engelse taalvaardigheid, als we er rekening mee houden dat leerlingen op de kenmerken van de leerlingvragenlijst van elkaar verschillen.

[39] het evp is samengesteld door cambridge university press op basis van het cambridge learner corpus.

Vragenlijst voor leerkrachten

De leerkrachten die het onderwijs in Engels aan groep 8 verzorgen, vulden een vragenlijst in. In deze vragenlijst stonden vragen over de volgende aspecten van het onderwijsleerproces op het gebied van Engels:

- de doelen die het meest bepalend zijn voor de inrichting van hun onderwijs;
- de methoden, lesactiviteiten en werkvormen die zij gebruiken;
- het geven van huiswerk;
- toetsing en ondersteuning van leerlingen.

Naast vragen over het onderwijsleerproces bevatte de vragenlijst vragen over een aantal relevante achtergrondkenmerken van de leerkracht:

- ervaring en opleiding;
- buitenschools gebruik van het Engels;
- zelfinschatting van Engelse vaardigheden;
- inschatting van de leerlingvaardigheden op het gebied van Engels;
- waardering voor Engels in het primair onderwijs;
- algemene achtergrondkenmerken, bijvoorbeeld moedertaal en leeftijd.

In totaal vulden 109 leerkrachten van 90 deelnemende scholen (94,7 procent) de leerkrachtvragenlijst in. Bij sommige scholen vulden meer dan 1 leerkracht de leerkrachtvragenlijst in. Dit gebeurde op scholen waarvan meerdere groepen 8 deelnamen en bij scholen met duo-leerkrachten in groep 8. Op 15 scholen vulden 2 leerkrachten de vragenlijst in, op 2 scholen 3 leerkrachten. Op 5 scholen is de leerkrachtvragenlijst niet ingevuld.

We gebruikten de gegevens uit de leerkrachtvragenlijst om het onderwijsleerproces in kaart te brengen en een beschrijving te geven van relevante achtergrondkenmerken van de leerkrachten die het onderwijs in Engels aan groep 8 verzorgen. Ook gebruikten we de gegevens om de samenhang tussen kenmerken van het onderwijsleerproces en prestaties van de leerlingen op het gebied van de Engelse taalvaardigheid te onderzoeken als we rekening houden met de verschillen tussen leerkrachten in hun achtergrond.

Gegevens uit de intakevragenlijst

Voor de afname zijn met alle deelnemende scholen intakegesprekken gehouden. Deze gesprekken vonden plaats met de schoolleider en, indien mogelijk, de groepsleerkracht(en) van de deelnemende groep(en) 8. Tijdens deze intakegesprekken is een aantal vragen voorgelegd om de organisatie en inrichting van het onderwijs in Engels op de scholen in kaart te brengen. Het betrof vragen over:

- het startmoment van het onderwijs in Engels;
- de aanwezigheid van een coördinator of werkgroep voor Engels;
- de faciliteiten die de school voor Engels biedt aan leerkrachten;
- de overdracht van leerlinggegevens aan het voortgezet onderwijs.

Op 16 van de 95 deelnemende scholen vond de intake op verzoek van de school telefonisch of via Skype plaats. Deze scholen vulden de vragen in het intakeformulier na afloop in en stuurden ze digitaal op.

In totaal ontvingen we van 89 van de 95 scholen (93,7 procent) gegevens. Deze gegevens gebruikten we om de betreffende kenmerken van het onderwijsleerproces te beschrijven en om de samenhang tussen deze kenmerken van het onderwijsleerproces en de prestaties van de leerlingen op het gebied van de Engelse taalvaardigheid te onderzoeken.

Steekproef van deelnemende scholen en leerlingen

Voor de peiling is een steekproef getrokken van leerlingen die in het schooljaar 2017/2018 in groep 8 van hun basisonderwijs zaten. De steekproef van leerlingen is indirect getrokken door een steekproef van scholen te trekken en in principe alle leerlingen in groep 8 op een geselecteerde school op te nemen in de steekproef. Voor de steekpoeftrekking zijn de scholen in twee groepen verdeeld op basis van het startmoment met het onderwijs in Engels (eibo en vvto). Vervolgens zijn de scholen per startmoment verdeeld in vier groepen op basis van het percentage gewichtenleerlingen[40] op een school[41]. Binnen deze acht groepen zijn de scholen geordend naar mate van stedelijkheid en schoolgrootte. Hieruit is een steekproef van honderd scholen getrokken (met twee reservesteekproeven). De tweehonderd scholen uit de reservesteekproeven dienden als vervangers voor scholen uit de eerste steekproef.

Uiteindelijk deden 95 scholen mee aan het peilingsonderzoek[42]. Deze scholen zijn representatief voor de populatie Nederlandse basisscholen als het gaat om stedelijkheid, schoolgrootte, denominatie en toezicht-arrangement. Wat betreft het percentage gewichtenleerlingen en regio verschilde de steekproef van de totale populatie basisscholen. Bij het rapporteren van de data op schoolniveau hielden we hier rekening mee door terug te wegen naar het percentage gewichtenleerlingen. Ook is er teruggewogen naar start-moment van het onderwijs in Engels (eibo en vvto), omdat er verhoudingsgewijs meer vvto-scholen aan dit onderzoek deelnamen dan de totale verhouding vvto- tegenover eibo-scholen in Nederland.

In de opzet van de peiling was het de bedoeling dat van een deelnemende school alle leerlingen in leerjaar 8 aan het onderzoek zouden deelnemen. Dit had tot gevolg dat van 8 van de 95 scholen 2 of meer groepen 8 aan het peilingsonderzoek deelnamen. Ter voorbereiding op de afname leverden deze scholen gegevens van 2.343 groep 8-leerlingen aan. Ongeveer 95 procent van de aangemelde leerlingen (2.219) had van hun ouders toestemming om aan het onderzoek mee te doen. Van deze leerlingen was op de afnamedag in totaal 5 procent afwezig (110 leerlingen) vanwege ziekte of andere redenen. Daarnaast viel een hele klas met 21 leerlingen uit wegens ziekte van de leerkracht. Dit betekent dat er uiteindelijk 2.088 leerlingen aan het peilingsonderzoek deelnamen. De deelnemende leerlingen wijken niet significant af van de landelijke populatie groep 8-leerlingen wat betreft leeftijd en de verhouding jongens/meisjes. Wel verschilt de steekproef significant van de populatie door een groter aandeel leerlingen met leerlinggewicht 1,2. Deze oververtegenwoordiging kan (mede) verklaard worden doordat de steekproef bewust meer scholen bevatte met een hoog percentage gewichtenleerlingen. Reden hiervoor was met name om voor het verdiepend onderzoek scholen te kunnen selecteren binnen 4 groepen die varieerden wat betreft startmoment Engels (vvto- of eibo-scholen) en percentage gewichtenleerlingen (omwille van de herkenbaarheid van de resultaten voor verschillende typen scholen). Bij de analyses hielden we hier rekening mee door terug te wegen voor het percentage gewichtenleerlingen. Ook wogen we, zoals eerder gemeld, terug naar start-moment met het onderwijs in Engels (eibo en vvto) omdat er verhoudingsgewijs meer vvto-scholen aan dit onderzoek deelnamen dan de totale verhouding vvto- tegenover eibo-scholen in Nederland (omwille van de vergelijking tussen vvto- en eibo-scholen).

Naast de representativiteit van de steekproef op leerlingniveau is het ook relevant na te gaan in hoeverre de groep leerlingen, die van hun ouders toestemming had voor deelname aan de gesprekstaken, representatief was. Van de 2.088 groep 8-leerlingen is voor 229 van de leerlingen (15,8 procent) geen actieve toestemming van de ouders ontvangen voor de afname van de gesprekstaken[43]. Om informatie te verkrijgen over de

[40] het betreft hier het 'gewicht' dat een leerling bij instroom in het basisonderwijs krijgt toegekend en dat de eventuele extra bijdrage bepaalt die een school voor de leerling krijgt. dit als onderdeel van de gewichtenregeling basisonderwijs die zich richt op het verminderen van onderwijsachterstanden van risicoleerlingen in het basis-onderwijs. het 'gewicht' van een leerling wordt momenteel berekend op basis van het opleidingsniveau van de ouders/verzorgers: hoe hoger de opleiding van ouders, hoe lager het gewicht.

[41] Zowel het aantal vvto-scholen als het aantal scholen met veel gewichtenleerlingen was in de steekproef bewust oververtegenwoordigd.

[42] in totaal benaderden we 230 scholen voor deelname en hadden aanvankelijk 101 scholen deelname toegezegd (44 procent). om diverse redenen zijn later nog 6 scholen uitgevallen.

[43] dit is inclusief een klas van 20 leerlingen waarbij de schoolleider weigerde actieve toestemming bij de ouders op te vragen.

representativiteit van de groep leerlingen die aan de gesprekstaken deelnam, zijn de resultaten op de toetsen voor lezen, luisteren en woordenschat van de groep leerlingen met actieve toestemming vergeleken met de resultaten van de groep leerlingen zonder actieve toestemming. Toetsing van de verschillen in de gemiddelde scores op de toetsen voor lezen en luisteren liet zien dat de prestaties van beide groepen niet significant van elkaar verschilden. Voor woordenschat scoorde de groep leerlingen met actieve toestemming significant hoger dan de groep leerlingen zonder actieve toestemming. Het absolute verschil in gemiddelden is echter zeer klein en op de totale range aan scores niet relevant.

In het technisch rapport van het consortium dat de instrumentontwikkeling en dataverzameling voor deze peiling uitvoerde, staat een uitgebreidere verantwoording van de steekproeftrekking, werving, dataverzameling en het instrumentarium (Ritzema et al., 2019).

Verdiepend onderzoek

Na de afname van het peilingsonderzoek in schooljaar 2017/2018, is in september en oktober 2018 een verdiepend onderzoek uitgevoerd op een selectie van bovengemiddeld presterende scholen. Doel van dit verdiepend onderzoek was om inzicht te krijgen in de (vak)didactische invulling van de lessen Engels in groep 8 en, waar mogelijk, inspirerende lespraktijken te verzamelen.

Bij de selectie van scholen voor dit verdiepend onderzoek is rekening gehouden met de steekproefopzet voor het peilingsonderzoek door binnen vier groepen scholen die varieerden qua startmoment Engels (vvto- of eibo-scholen) en percentage gewichtenleerlingen de bovengemiddeld presterende scholen te selecteren. Om deze scholen te identificeren, is gebruik gemaakt van de schoolprestaties op enerzijds de gesprekstaak en anderzijds de luister-, lees- en woordenschattoets. Hierbij is rekening gehouden met een aantal leerlingkenmerken waarop de scholen verschillen. Voor de luister-, lees- en woordenschattoets ging dit om leerlinggewicht, geslacht, leeftijd, thuistaal en frequentie van het gebruik van Engels (spreken, luisteren, lezen en schrijven) door de leerlingen. Voor de gesprekstaak ging dit om leerlinggewicht, geslacht en leeftijd en de score op de luister-, lees- en woordenschattoets. In totaal zijn zo 49 scholen geselecteerd, waarvan 21 scholen (met 22 klassen) uiteindelijk deelnamen.

Het verdiepend onderzoek bestond uit lesobservaties in en interviews met leerkrachten van groep 8. Zeven, specifiek voor dit onderzoek getrainde, inspecteurs in de rol van onderzoeker voerden het onderzoek uit. Voor de observaties van de Engelse lessen is een instrument ontwikkeld dat bestond uit een generiek deel en een vakdidactisch inhoudelijk deel. Voor het generieke deel is gebruik gemaakt van een bestaand observatie-instrument (ICALT-PO van Van de Grift, Van der Wal & Torenbeek, 2011). Dit instrument brengt pedagogisch-didactische vaardigheden van leerkrachten in kaart die bewezen effectief zijn. Het betreft 24 observatiepunten over 6 dimensies geordend van meer basale naar complexe vaardigheden:

- een veilig en stimulerend klasklimaat;
- efficiënte lesorganisatie;
- duidelijke en gestructureerde instructie;
- intensieve en activerende les;
- afstemmen op verschillen;
- leerstrategieën aanleren.

De observatoren scoorden ieder observatiepunt op een vierpuntsschaal van ‘overwegend zwak’ tot ‘overwegend sterk’. Daarnaast is met drie observatiepunten op dezelfde vierpuntsschaal ook de betrokkenheid van de leerlingen bij de les in kaart gebracht.

Het vakdidactische deel van het observatie-instrument en de vragen van het interview zijn gebaseerd op theoretische uitgangspunten over succesvol onderwijs in de moderne vreemde talen en praktische bevindingen over de vakdidactiek Engels in het primair onderwijs (Tuin & Thijs, 2015). Het vakdidactische gedeelte van het observatie-instrument bevatte 22 observatiepunten over de volgende 6 dimensies van de les:

- gebruik van taaldoelen;
- uitleg en instructie;
- interactie;
- (vak)didactisch repertoire;
- evaluatie van de les;
- algemene indruk van de les.

Bij elk observatiepunt, dat soms uit meerdere subitems bestond, konden de observatoren aangeven of zij het wel of niet hadden gezien.

Het interview is na afloop van de observatie van de les Engels gehouden aan de hand van een semi-gestructureerde interviewleidraad. Het ging in op drie centrale thema's:

- visie op Engels in het primair onderwijs;
- informatie over een standaardles Engels, in relatie tot de geobserveerde les;
- opleiding en professionele ontwikkeling van de leerkracht.

De belangrijkste bevindingen van het verdiepend onderzoek staan in hoofdstuk 5 van deel B van deze rapportage. Een uitgebreide beschrijving van de selectie van scholen, de samenstelling van het observatie-instrument en de interviewleidraad, de uitvoering van het verdiepend onderzoek en de resultaten staat in het technisch rapport van het consortium dat de instrumentontwikkeling en de dataverzameling voor deze peiling uitvoerde (Ritzema et al., 2019).