

De referentieniveaus getoetst

Inleiding

In het artikel 'Referentieniveaus voor het basisonderwijs' ga ik in op de inhoud van de referentieniveaus. Als vervolg hierop kijk ik nu hoe de referentieniveaus worden getoetst in de verschillende eindtoetsen.

Kun je een goede toets maken bij deze referentieniveaus?

Bij het bestuderen van de referentieniveaus kwam bij mij vaak de vraag op hoe ze nu precies uitgelegd moeten worden, het was niet eenvoudig om deze te vertalen naar leerdoelen voor het onderwijs. Tegelijk vroeg ik me steeds vaker af hoe je deze, niet altijd scherp geformuleerde, referentieniveaus op een goede manier kunt toetsen. Ik heb me dan ook verdiept in de eindtoetsen voor het basisonderwijs, gezien vanuit de referentieniveaus. Dat leverde bijzondere inzichten op. In dit artikel schets ik eerst een beeld van de praktijk en vervolgens verbind ik hier een voorlopige conclusie aan.

Hoe werkt het met de eindtoets?

De Commissie voor Toetsen en Examens (CvTE) is de uitvoeringsorganisatie die in opdracht van het Ministerie van Onderwijs zorgdraagt voor een goed verloop van de eindtoets. Zij hebben aan het bedrijf Cito de opdracht gegeven om een eindtoets te ontwikkelen. Daarnaast mogen ook andere commerciële aanbieders eindtoetsen aanbieden. De verschillende aanbieders zijn de IEP, Route8, AMN en DIA. Zij krijgen per school die hun toetsen afneemt een vergoeding van het Ministerie van Onderwijs.

Het mooie is dat er iets te kiezen is, maar de vraag is dan wel hoe de verschillende toetsaanbieders de referentieniveaus toetsen, de referentieniveaus zijn niet altijd even scherp geformuleerd en kunnen in theorie dus verschillend uitgelegd worden. Geven de verschillende eindtoetsen een gelijkwaardig oordeel en een gelijk advies voor de doorstroming naar het voortgezet onderwijs? In de praktijk bleek dit inderdaad niet het geval. Het CvTE heeft dit opgelost door referentiesets en ankeropgaven te laten ontwikkelen, waarmee de toetsontwikkelaars hun toetsen qua aanbod en cesuur (bij hoeveel opgaven haal je 1F bijvoorbeeld) op elkaar kunnen afstemmen.

Wat zijn referentiesets en ankeropgaven?

Een referentieset bestaat uit een set van opgaven die de referentieniveaus zo goed mogelijk aftoetsen. Het zijn meerkeuzevragen die in alle eindtoetsen opgenomen moeten worden om zo tot een gelijkwaardig oordeel te kunnen komen over het wel of niet beheersen van de referentieniveaus. Daarbij wordt per referentieset afgesproken welke cesuur daarbij hoort. Zo zijn er 80 vragen voor de referentieniveaus voor het 1F niveau voor rekenen. Een leerling haalt 1F als hij 50 of meer vragen goed heeft.

Er zijn twee functies van de referentiesets: het op elkaar afstemmen van de cesuur (bij hoeveel opgaven heb je het referentieniveau behaald) en het is een manier om te laten zien welke type vragen er in de eindtoetsen zullen worden opgenomen. Er zijn steeds twee soorten referentiesets ontwikkeld: een openbare en een niet openbare set. De niet openbare set is naar de toetsontwikkelaars gestuurd, de openbare set is te vinden op een website (zie een link in de bronnen onderaan het artikel) en kan gebruikt worden om te onderzoeken hoe de eindtoetsen zijn opgebouwd. We zoomen nu in op die referentiesets voor de verschillende onderdelen van de eindtoets.

De referentieniveaus voor Nederlandse taal

Ik geef eerst een kort overzicht van de referentieniveaus voor Nederlandse taal, vervolgens kijk ik naar hoe dit wordt getoetst in de eindtoets.

Taal kent in het referentiekader de volgende indeling:

1. Mondelinge Taalvaardigheid

1.1 Gesprekken

1.2 Luisteren

1.3 Spreken

2. Lezen

2.1 Zakelijke teksten

2.2 Fictionele, narratieve en literaire teksten

3. Schrijven

4. Begrippenlijst en taalverzorging (spelling)

Een voorbeeld van een klein onderdeel uit de omschrijving van de referentieniveaus:

2.1 Zakelijke teksten	Niveau 1F	Niveau 2F
Algemene omschrijving Lezen zakelijke teksten	Kan eenvoudige teksten lezen over alledaagse onderwerpen en over onderwerpen die aansluiten bij de leefwereld.	Kan teksten lezen over alledaagse onderwerpen, onderwerpen die aansluiten bij de leefwereld van de leerling en over onderwerpen die verder van de leerling afstaan.
Teksten		
Tekstkenmerken	De teksten zijn eenvoudig van structuur; de informatie is herkenbaar geordend. De teksten hebben een lage informatiedichtheid; belangrijke informatie is gemarkeerd of wordt herhaald. Er wordt niet te veel (nieuwe) informatie gelijktijdig geïntroduceerd. De teksten bestaan voornamelijk uit frequent gebruikte (of voor de leerlingen alledaagse) woorden.	De teksten hebben een heldere structuur. Verbanden in de tekst worden duidelijk aangegeven. De teksten hebben overwegend een lage informatiedichtheid en zijn niet te lang.

De eindtoets voor Nederlandse taal

In het artikel 'Referentieniveaus voor het basisonderwijs' liet ik al zien dat voor taal alleen de onderdelen *Lezen* en *Taalverzorging* worden afgetoetst in de eindtoets. Alle andere onderdelen, zoals mondelinge taalvaardigheid en het schrijven van bijvoorbeeld een brief worden niet afgetoetst in de eindtoets.

We kijken nu naar deze twee onderdelen, we beginnen met het onderdeel *Taalverzorging*.

De eindtoets taalverzorging

In de referentieniveaus is taalverzorging het onderdeel dat we meestal aanduiden met spelling. De referentieniveaus voor dit onderdeel vind je [hier](#).

Procedure vaststellen referentiecesuren en ankersets voor taalverzorging

De referentiesets zijn opgesteld door Cito, waarbij ze gebruik hebben gemaakt van een expertgroep met experts van onder andere het Kohnstamm instituut, SLO en Hogescholen, waarna er proefafnames zijn geweest die voorgelegd zijn aan een groep experts (oa leerkrachten) die samen de cesuur hebben vastgesteld.

Zij hebben vastgesteld wat de cesuur is bij de referentieopgaven voor Nederlandse taal

- 1F inclusief grammatica: 42 opgaven van de 70 goed (60%)
- 2F inclusief grammatica: 51 opgaven van de 70 goed (72,8%)
- 1F exclusief grammatica: 31 opgaven van de 50 goed (62%)
- 2F exclusief grammatica: 37 opgaven van de 50 goed (74%)

Het onderdeel is grammatica is dus geen verplicht onderdeel op de eindtoets.

Het CvTE geeft aan dat de commerciële toetsaanbieders de ankersets kunnen gebruiken om het niveau van hun eindtoetsen hierop af te stemmen. Dat kan bijvoorbeeld door een ankeropgave mee te nemen in een proeftoets waarin meerdere soortgelijke opgaven worden afgenomen. Door de resultaten op de ankeropgaven te vergelijken met de andere opgaven kan de toetsontwikkelaar aldus het CvTE: *“alle opgaven op één vaardigheidsschaal plaatsen en de referentiecesuur overbrengen.”* Dat wil dus zeggen dat gekeken wordt welke opgaven net zo moeilijk zijn als de referentieopgaven en dus zal de toetsontwikkelaar dit type opgave meenemen in deze en toekomstige toetsen. Het betekent ook dat de verdeling van de soorten opgaven (dus welk domein) gelijk moet zijn aan de referentieopgaven om zo een gelijkwaardig speelveld te creëren. Ook de verhouding tussen het aantal gemaakte opgaven en het aantal opgaven dat een leerling voldoende moet maken zal dus gelijk moeten liggen bij de verschillende toetsontwikkelaars.

Kritiek Expertgroep op de referentieniveaus?

Interessant is om te lezen welke kritiek de expertgroep had op de referentieniveaus. Ze laten optekenen dat ze de referentieniveaus hebben laten prevaleren boven hun professionele inschatting; de ankersets zijn dus niet helemaal in lijn met hun eigen professionele inschatting voor wat betreft de inhoud die afgetoetst zou moeten worden. Ze geven onder andere deze punten aan:

- *De beschrijving van het domein taalverzorging is niet altijd consistent*
- *Een voorbeeld hiervan is de verleden tijd van werkwoorden met een stam op -d of -t: deze categorie wordt zowel bij 1F als 2F genoemd.*
- *Opgaven waarvan de panelleden uit ervaring weten dat deze relatief makkelijk zijn, behoren volgens het referentiekader soms tot een hoog niveau. Dit is bijvoorbeeld aan de orde bij de vervoegingen tegenwoordige tijd (zwakke of klankvaste) werkwoorden enkelvoud zoals bijvoorbeeld ‘hij werkt’. De spelling van deze persoonsvorm levert leerlingen relatief weinig problemen op, maar behoort volgens het referentiekader tot niveau 2F.*
- *Sommige categorieën konden niet in de ankersets worden bevraagd, omdat ze op bepaalde niveaus ontbreken in het referentiekader. Zo is bijvoorbeeld het voltooid deelwoord niet bij spelling werkwoorden in de openbare ankerzet 2F opgenomen, omdat de spelling hiervan volgens het referentiekader óf tot niveau 1F, óf tot niveau 3F behoort.*
- *Ten slotte heeft het expertpanel ervoor gekozen om het referentiekader niet al te strikt te interpreteren als dit ertoe zou leiden dat de ankerzet niet meer naar behoren zou functioneren.*

Dit was aan de orde bij het samenstellen van het onderdeel grammatica op niveau 2F. Het referentiekader heeft slechts twee categorieën ingedeeld op dit niveau: de lijdende en de bedrijvende vorm. Het opnemen van opgaven met uitsluitend deze twee categorieën, zou leiden tot te grote eenvormigheid, waardoor de ankerset niet goed meer zou kunnen functioneren als ijkingsinstrument. Daarom heeft het expertpanel besloten opgaven in de ankersets 2F op te nemen die categorieën bevragen op niveau 1F, maar die bijvoorbeeld vanwege woordgebruik of zinslengte als moeilijk worden ervaren.

(bron: Rapportage Ankersets Taalverzorging: verantwoording project)

Een voorbeeld uit de referentieset voor taalverzorging, links 1F vragen, rechts 2F. Wat opvalt is dat het verschil tussen 1F en 2F niet echt groot is in veel gevallen

Spelling, niet werkwoorden

TV1F_01 In welke zin zijn de **dikgedrukte** woorden **allebei** goed gespeld?

- A. De **veldmuisen** renden naar hun **holletjes**.
- B. De **veldmuisen** renden naar hun **holletjes**.

TV1F_42 door de vele regen de laatste tijd **kun** je nauwelijks meer lopen **in** het park **tegenover** ons huis **we** laten de hond daarom maar gewoon op straat uit

Welk onderstreept woord moet met een hoofdletter worden geschreven?

- A. kun
- B. in
- C. tegenover
- D. we

TV1F_43 na een uur **rijden** kwamen de **ruiters** bij een **bos** ze stapten **af** voor een korte pauze

Achter welk onderstreept woord moet een punt (.) worden gezet?

- A. rijden
- B. ruiters
- C. bos
- D. af

TV2F_01 In welke zin zijn de **dikgedrukte** woorden **allebei** goed gespeld?

- A. Isa vulde de **stale** etensbak met het **hondenvoer**.
- B. Isa vulde de **stale** etensbak met het **hondenvoer**.
- C. Isa vulde de **stalen** etensbak met het **hondenvoer**.

TV2F_41 snikkend riep ze tegen haar **vriend** **je** had me nog zo beloofd **dat** **je** **vanavond** met me naar de film zou gaan

Welk onderstreept woord moet met een hoofdletter worden geschreven?

- A. vriend
- B. je
- C. dat
- D. vanavond

TV2F_42 mijn tante zei **lachend** **geef** mij maar een heerlijk stuk **appeltaart** **met** slagroom bij de koffie

Welk onderstreept woord moet met een hoofdletter worden geschreven?

- A. lachend
- B. geef
- C. appeltaart
- D. met

Interpunctie

De eindtoets voor lezen

Het onderdeel Lezen in de eindtoets is Begrijpend lezen. Het gaat dan om verschillende tekstvoorten met daarbij meerkeuzevragen over die tekst.

De referentiesets bestaan uit de volgende tekstvoorten:

- Fictionele teksten (1F en 2F, ten behoeve van het PO) – Boek – Fabel – Gedicht – Historisch – Mop – Parabelachtig – Sprookje – Realistische fictie
- Zakelijke teksten – Betogend – Informatief – Instructief

De referentiesets taal bevatten hoofdzakelijk meerkeuzevragen

Bron: [Rapportage referentiesets Nederlandse taal \(lezen\) en rekenen: verantwoording project](#) 

Om een goed idee te krijgen van de inhoud van de eindtoets op dit onderdeel is het goed om deze referentiesets te bekijken. We laten een aantal voorbeelden zien uit de openbare referentieset, de complete set is hier te vinden: http://www.toetspecials.nl/html/referentiesets_openbaar/taal.shtm

Het eerste voorbeeld is deel van een 1F vraag voor het onderdeel Begrijpend lezen:

De keuzes van Hind, Jasmijn en John

	Hind	Jasmijn	John
(1) Londen of Parijs?	Parijs, voor het stokbrood, de muziek en alle bezienswaardigheden.	Parijs, een gezellige stad met mooie musea.	Parijs. Dan ga ik lekker de hele dag bovenop de Eiffeltoren staan.
(2) Zon of sneeuw?	Zon! Sneeuw is leuk voor een dagje, maar zon is altijd heerlijk.	Sneeuw is wel leuk, maar ik word veel blijer van de zon.	Zon is veiliger. Bij sneeuw krijg ik rare plannen, wil ik met mijn fiets gaan slippen en zo...
(3) Pretpark of museum?	Pretpark. Tegenwoordig durf ik in alle achtbanen.	Ik haat pretparken, durf ook nergens in, dus dan sta ik steeds aan de kant met alle tassen. Doe mij maar een mooi museum.	Pretpark! Vorig jaar heb ik ze met mijn vrienden allemaal in één week gedaan. Moest gewoon even.
(4) Hagelslag of pindakaas?	Vroeger was ik verslaafd aan pindakaas, nu eet ik ook graag hagelslag.	Pindakaas! Wacht nee, toch hagelslag. Of niet...? Mmm... jam is ook lekker.	Pindakaas, daar zit je lekker snel vol van. Af en toe broodje kaas.
(5) Ajax of Feyenoord?	Ajax	Ajax. Ik kom uit Aals-	Hallo zeg, ik ben een Amsterdammer! Ajax

T1F_11

Wie houdt niet van pretparken?

- A. Hind
- B. Jasmijn
- C. John

T1F_12

Waarom wonen Hind, Jasmijn en John liever in de stad dan op het platteland?

- A. Ze genieten van het lawaai in de stad.
- B. Ze houden meer van mensen dan van dieren.
- C. Ze houden van de drukte in de stad.
- D. Ze vinden het maar stinken op het platteland.

T1F_13

Vul onderstaande tabel in: zet per onderwerp een kruisje onder Alle drie eens, Twee van de drie eens of Niemand met elkaar eens.

Onderwerp	Alle drie eens	Twee van de drie eens	Niemand met elkaar eens
a. Londen of Parijs?			
b. Zon of sneeuw?			
c. Pretpark of museum?			
d. Hagelslag of pindakaas?			
e. Ajax of Feyenoord?			
f. Stad of platteland?			
g. Fruit of snoep?			
h. Vlees of vis?			
i. Hond of kat?			

T1F_14

Hoeveel vragen worden er gesteld over voeding?

- A. 1
- B. 2
- C. 3
- D. 4

De antwoorden voor deze vraag zijn te vinden op de laatste bladzijde van deze referentieset. Wat mij opvalt is dat sommige vragen vooral een soort strikvragen zijn, bijvoorbeeld vraag T1F_13, onderdeel **d** over hagelslag of pindakaas. Je kunt als argument aanvoeren dat ze hier vooral goed moeten lezen, maar bij mij blijft wel steeds het gevoel hangen dat sommige vragen op meerdere manier te interpreteren zijn. De onderliggende vraag is wat je nu precies aftoetst met dit soort vragen, kun je hiermee bepalen of een leerling bijvoorbeeld een TL of een Havo leerling is?

Nog een voorbeeld, nu uit 2F:

- 1 **Afzien voor het goede doel**
- 2 Zweeten is het nieuwe collecteren
- 3 Geld ophalen voor het goede doel. Met een
- 4 collectebus haal je hooguit een paar honderd
- 5 euro op. Wie het serieus wil aanpakken gaat
- 6 beulen en sporten. Elk jaar reizen bijvoorbeeld
- 7 duizenden mensen af naar Frankrijk om te
- 8 fietsen op de bergen Alpe d'Huez en Mont
- 9 Ventoux.
- 10 De collectebus verliest het van de massale
- 11 sportactiviteiten en sponsorlopen: rennen of
- 12 fietsen voor het goede doel. Met het geld van
- 13 de sponsors worden onderzoeken betaald naar
- 14 kanker, taaislijmziekte of andere ziekten. De
- 15 actie Alpe d'HuZes leverde in 2012 maar liefst
- 16 28,5 miljoen euro op voor kankeronderzoek.
- 17 Als je met de deelnemers praat, kom je
- 18 erachter dat ze allemaal wel mensen kennen
- 19 die aan zo'n ziekte lijden. Het kan een vader of
- 20 moeder zijn die voor een ziek kind loopt, of een
- 21 vriendengroep die op deze manier steun wil
- 22 betuigen aan een zieke vriend. De motivatie die
- 23 de deelnemers geven is vaak: "De zieken
- 24 moeten elke dag topsport leveren om hun dag
- 25 door te komen. Laten wij ons dan een keer in
- 26 het zweet werken om ze te helpen."

T2F_01

Waar gaat het in deze tekst **vooral** om?

- A. Gesponsorde sportactiviteiten zijn populair omdat het lekker is om jezelf in het zweet te werken, zeker als het voor het goede doel is.
- B. Het heeft niet zoveel zin om met een collectebus geld op te halen voor onderzoek naar ziekten.
- C. Omdat steeds meer mensen lijden aan ziekten als kanker, moet er steeds meer geld ingezameld worden voor onderzoek.
- D. Tegenwoordig wordt veel geld voor onderzoek naar ziekten opgehaald door middel van gesponsorde sportactiviteiten.

T2F_02

Wat is het belangrijkste doel van deze tekst?

- A. Aanwijzingen geven over het organiseren van sportieve acties voor het goede doel.
- B. Informeren over sportieve acties voor het goede doel.
- C. Kritisch beoordelen van sportieve acties voor het goede doel.
- D. Overhalen van de lezer mee te doen aan sportieve acties voor het goede doel.

Er volgen dan nog een aantal vragen over deze tekst, maar dit even als een voorbeeld van een 2F vraag.

Hoe goed doen we het bij taal?

In 2019 beheerst een groot deel van de leerlingen niveau 1F voor lezen (98 procent) en taalverzorging (97 procent). Een percentage ruim boven de ambitie van 85 procent. Voor lezen beheerst 78 procent van de leerlingen het streefniveau 1S/2F. Dat is ook boven de ambitie van minimaal 65 procent. Voor taalverzorging haalt 60 procent van de leerlingen het 2F niveau. Dit zijn behoorlijk hoge percentages, in de meeste gevallen ruim boven de verwachting van de toetsontwikkelaars en ook wel gezien de zorgen die er zijn rond het niveau van taal en lezen.

Voor de verdeling van de leerlingen over het vervolgonderwijs maakt een hoge score geen verschil, de verdeling is al jaren: 20% VWO, 30% HAVO, en 50% voor de verschillende onderdelen van het VMBO. Het is dus niet zo dat als alle leerlingen hoger scoren dan verwacht dat er dan meer leerlingen naar een bepaalde vorm van voortgezet onderwijs gaan, het gaat bij deze verdeling om de score ten opzichte van andere leerlingen.

De eindtoets voor rekenen

In het artikel 'Referentieniveaus voor het basisonderwijs' beschreef ik hoe de referentieniveaus voor rekenen zijn opgebouwd. Een kleine samenvatting hiervan, de hoofdindeling is:

Getallen

Verhoudingen

Meten en meetkunde

Verbanden

Daarna komt steeds de indeling:

A. Notatie, taal en betekenis

B. Met elkaar in verband brengen

C. Gebruiken

Deze drie onderdelen hebben dan steeds weer de verdeling in:

Paraat hebben,

Functioneel gebruiken

Weten waarom

en dan volgt de doelbeschrijving.

Zoals ik al aangaf is het in de praktijk lastig om deze kopjes goed te interpreteren, wat is *functioneel gebruiken* bij B. *Met elkaar in verband brengen* en wat is *functioneel gebruiken* bij C. *Gebruiken*? Daarbij is de verwoording van de doelen ook niet altijd meteen helder: soms wordt er een voorbeeldsom gegeven, soms staat er alleen een woord, bijvoorbeeld: - schaal. Dan is er nog best wat interpretatie nodig om hier een concreet doel van te maken.

Het onderdeel *Getallen*, B. *Met elkaar in verband brengen* uit de referentieniveaus heb ik ter illustratie bijgevoegd. Deze referentieniveaus vormen de basis voor de eindtoets.

Getallen	Niveau 1F	Niveau 1S
B. Met elkaar in verband brengen	Paraat hebben	
Getallen en getalrelaties	-tienstructuur -getallenrij	-getallenlijn, ook met decimale getallen en breuken
Structuur en samenhang	-getallenlijn met gehele getallen en eenvoudige decimale getallen	
	Functioneel gebruiken	

	-vertalen van eenvoudige situatie naar berekening -afroonden van gehele getallen op ronde getallen -globaal beredeneren van uitkomsten -splitsen en samenstellen van getallen op basis van het tientallig stelsel	-vertalen van complexe situaties naar berekening - decimaal getal afronden op geheel getal - afronden binnen gegeven situatie: 77,4 dozen berekend dus 78 dozen kopen
	Weten waarom	
	-structuur van het tientallig stelsel	-opbouw decimale positiestelsel -redeneren over breuken, bijvoorbeeld: is er een kleinste breuk?

De openbare referentiesets voor rekenen

In dit artikel proberen we op basis van de openbare referentieset een beeld te krijgen van hoe de referentieniveaus worden afgetoetst en wat in essentie het verschil is tussen 1F en 1S. Het aantal opgaven voor 1F is in totaal 80, voor 1S zijn er 50 opgaven, ik heb geteld hoeveel dat er zijn per domein. Ik gebruik daarvoor de indeling van het boek 'Leerlijnen voor het basisonderwijs'. Zelf vind ik de indeling in Getallen, Verhoudingen etc niet altijd duidelijk. Zo is het onderdeel breuken in twee verschillende domeinen terug te vinden. In de tabel zie je het aantal opgaven per domein, het wil niet zeggen dat dit exact gelijk is aan de uiteindelijke eindtoetsen, maar het geeft wel een goede indicatie. Het uiteindelijke aantal opgaven in de eindtoets kan hiervan afwijken, maar de verhouding zal gelijk blijven. Dat geldt ook voor de cesuur die hieraan gekoppeld is: een leerling haalt 1F als hij 50 van de 80 opgaven goed heeft gemaakt (62,5%). Voor 1S moet een leerling 30 van de 40 opgaven goed maken om 1S te halen. (75%).

Domein	Aantal opgaven 1F	Aantal opgaven 1S
Getallen en getalrelaties	5	2
Optellen en aftrekken	5	3
Vermenigvuldigen en delen	12	4
Breuken en procenten	13	10
Meetkunde (figuren)	1	2
Lengte en omtrek	6	2
Oppervlakte	7	5
Inhoud	4	1
Gewicht	6	2
Tijd en snelheid	5	1
Geld		3
Tabellen en grafieken	7	4
Niet in referentieniveaus: gemiddelde		1
Totaal aantal opgaven	80	40

Voorbeeld referentieopgaven bij rekenen

Dan naar de inhoud van de referentieopgaven, ik laat een aantal voorbeelden zien. Het eerste voorbeeld gaat over Getallen en getalrelaties, onderdeel **afronden**.

Omschrijving referentieniveaus

Bij het onderdeel 1 Getallen, bij B: Met elkaar in verband brengen, bij functioneel gebruiken staat als omschrijving:

1F: - afronden van gehele getallen op ronde getallen

1S: - decimaal getal afronden op geheel getal

- afronden binnen gegeven situatie: 77,6 dozen berekend dus 78 dozen kopen

In de referentiesets zijn daarbij deze opgaven te vinden voor 1F en 1S:

R1F_49

Rond af op honderdtallen.

343 748

R1S_09



Het muziekfestival werd bezocht door 117 812 mensen. De krant rondt dit aantal af op tienduizend mensen nauwkeurig.

Welk aantal komt in de krant te staan?

- A. 110 000
- B. 117 800
- C. 118 000
- D. 120 000

R1F_54

Rond af op een heel getal.

1987 gram is afgerond:

_____ kg

R1F_64



In dit krantenartikel is 1,2 miljoen een afronding.

Van welk werkelijk aantal kan dit de afronding zijn?

- A. 19 642
- B. 1 189 764
- C. 1 290 361
- D. 12 091 012

Bij deze 1F opgaven zijn er nog twee andere vaardigheden van belang: bij de eerste het DHTE schema, wat is hier het honderdtal, bij de tweede het omzetten van grammen naar kilogrammen en vervolgens het afronden. De derde opgave bij 1F gaat ook al om een groter getal. Bij de 1S opgave gaat het ook om de DHTE indeling, wat is hier 10.000, de getallen zijn groter en er staat, (zoals vaker) een overbodig plaatje bij om de leerling te laten zoeken naar de juiste informatie in de opgave.

Het verschil tussen 1F en 1S in de eindtoets

Bij al deze opgaven kun je jezelf steeds de vraag stellen: als ik niet weet of dit een 1F of een 1S opgave is, zou ik dan kunnen aangeven of het een 1F of 1S vraag is? Wat is in essentie de bedoeling van de 1S toets? In de toetswijzer, te vinden op de website van de centrale eindtoets po staat deze tekst:

3.1.9 Nog niet te toetsen vaardigheden en perspectief op termijn

In de huidige vorm van de Centrale Eindtoets rekenen kunnen nog niet alle doelen die in het referentiekader zijn geformuleerd, getoetst worden. Dit heeft onder meer te maken met het feit dat de eindtoets op dit moment alleen meerkeuzevragen bevat. **Hiermee zijn doelen als 'kunnen uitleggen waarom...' en 'inzien dat...' moeilijk te toetsen.** Ook vaardigheden als 'hanteren van meetinstrumenten' en 'tekenen van diagrammen en grafieken' zijn met de huidige afnamevorm nog niet goed te meten. Vanaf 2018 worden andersoortige items opgenomen in de Centrale Eindtoets en kan deze situatie wijzigen.

Bron <https://www.centraleeindtoetspo.nl/publicaties/publicaties/2015/05/01/toetswijzer-bij-de-centrale-eindtoets-po-taal-en-rekenen>

Deze toetswijzer is niet aangepast sinds 2015, van andere afnamevormen is dan ook nog niets te vinden. Het streefniveau is bedoeld om leerlingen meer inzicht te laten ontwikkelen in het rekenen, maar juist het aftoetsen hiervan is, met alleen meerkeuzevragen, niet altijd goed te doen. De moeilijkheid van de 1S opgaven in de eindtoets zit hem over het algemeen in: meer tekst, grotere getallen, meer inzicht nodig (want er zijn meer denkstappen nodig om de opgave te kunnen oplossen), meer overbodige informatie zodat de leerling de juiste informatie moet kiezen, en zoals al genoemd is de cesuur zwaarder: je moet 30 van de 40 opgaven goed hebben (75%) om 1S te halen, voor 1F is dat 50 van de 80 opgaven (62,5%).

Een snelle optelling (ongeveer, om een indicatie te geven) van het verschil tussen het aantal 'kale' opgaven of opgaven zonder veel tekst

	1F	1S
Eenduidige sommen, zonder veel context of zonder tekst:	60 van de 80.	10 van de 30
Sommen met tekst of veel aflezen uit grafiek of een foto:	20 van de 80	20 van de 30

Hoe goed doen we het bij rekenen?

In 2019 beheerst een groot deel van de leerlingen niveau 1F voor rekenen (94 procent). Een percentage ruim boven de ambitie van 85 procent. Voor 1S is de ambitie minimaal 65%, maar dat streefpercentage wordt niet gehaald: 47% van de leerlingen haalt 1S bij rekenen.

Eindconclusie

Ik begon mijn artikel met de stelling dat de referentieniveaus niet altijd scherp zijn, wat het lastig maakt om dit goed af te toetsen. Onbedoeld lijkt het ook alsof de referentieniveaus de status van een eindexamen hebben gekregen, dit wordt versterkt door de exacte manier waarop de uitslag wordt gepresenteerd, terwijl die exactheid in de referentieniveaus ver te zoeken is. Mijn eerste gevoel, dat het vaak wel diffuus is, blijft zeker staan. Ja, het is moeilijk om precies te omschrijven wat nu bijvoorbeeld precies 2F is, een taalniveau is niet makkelijk vast te pakken, maar we suggereren in de uitslagen van de eindtoetsen wel een hoge graad van nauwkeurigheid. Ik krijg het idee dat de uitwerking en normering heel nauwkeurig bedoeld is, maar dat het onderliggende bouwwerk (de referentieniveaus zelf) op zijn minst erg onoverzichtelijk is beschreven. In de praktijk is al gebleken dat wanneer je twee groepen taal- of rekenexperts onafhankelijk van elkaar een eindtoets laten maken gebaseerd op de referentieniveaus, dat je dan twee behoorlijk verschillende eindtoetsen zult

krijgen. Als dit niet zo zou zijn waren de referentieopgaven niet nodig geweest. Dat zegt dus iets over de referentieniveaus, maar dat zegt ook iets over de huidige eindtoetsen. Zijn deze referentieniveaus en deze eindtoetsen nu wel zo'n goed instrument voor de onderwijsinspectie om het oordeel over een school op te baseren en geeft dit een eerlijk advies voor het vervolgonderwijs voor alle leerlingen? Laten we dit per doel bekijken. De eindtoetsen hebben drie verschillende doelen:

1. Het aftoetsen van de referentieniveaus, nagaan op welk niveau een leerling deze referentieniveaus beheerst.

Op papier halen we dit doel, we beoordelen immers elke leerling. Als je echter inzoomt op de inhoud van de referentieniveaus, welk beperkt deel van de referentieniveaus we aftoetsen en de vaak onduidelijke omschrijving van de referentieniveaus zelf, dan valt hier in ieder geval nog veel te verbeteren.

2. De verdeling van de leerlingen over de diverse vormen van voortgezet onderwijs

De verdeling van de percentages staat van tevoren al redelijk vast: 20% VWO, 30% HAVO, 50% VMBO. Hoe goed of slecht een eindtoets is, hoe goed of slecht deze gemaakt is, voor de onderlinge verdeling maakt dit dus niet uit. Alleen voor de verdeling van de leerlingen onderling maakt het uit, je bent altijd een concurrent van je medeleerlingen. Krijgt iedereen hier dan een eerlijke kans?

Als je ziet welk beperkt deel van wat een leerling moet kennen en kunnen hier afgetoetst wordt is mijn mening dat dit niet altijd eerlijk is, er zit een zekere mate van willekeur in. Hoe kun je op basis van deze toetsen zeggen of iemand een TL of een HAVO leerling is? Of een basis-of kaderleerling? Ja, ook het oordeel van de leerkracht telt, maar die is ook gebaseerd op (vaak) het CITO leerling volgsysteem, met hetzelfde type vraagstelling en onderwerpen, dus hoe eerlijk en nauwkeurig is dit?

3. De beoordeling van de onderwijsinspectie van het presteren van een school

Bijna alle leerlingen halen voor taal het verwachte niveau. Dat is natuurlijk mooi, maar hoe realistisch is dit? We horen regelmatig dat we te laag scoren op taal en dat zelfs een groot deel van de leerlingen niet goed genoeg kan lezen om goed te kunnen functioneren in de praktijk. Wat zegt dan een beoordeling van de school op basis van deze eindtoetsen? Als bijna iedereen de doelen voor taal haalt is het oordeel over een school dus afhankelijk van de score op rekenen en dan vooral op de 1S onderdelen, want die zijn gemiddeld gezien onder de verwachte scores. Dus van alle inspanningen van een school is (een deel van het oordeel) dus vooral gebaseerd op de 40 1S vragen bij rekenen. Een erg magere start om naar te kijken lijkt me. Gelukkig kijkt de inspectie breder dan dit, maar deze scores worden wel gebruikt voor verschillende ranglijsten.

Wat dan wel?

Mijn pleidooi is dan ook om de referentieniveaus in ieder geval aan te scherpen en te integreren met de kerndoelen. Twee soorten doelen (kerndoelen en referentieniveaus) heeft geen zin, dat leidt alleen maar tot verwarring. De volgende stap zou kunnen het onderzoeken van een betere vorm van toetsing, in samenwerking met het voortgezet onderwijs. Het zou dan niet meer moeten gaan om de onderlinge verdeling tussen de leerlingen, maar om het objectief halen van gestelde doelen. Erg belangrijk is dan om dit meteen zo concreet mogelijk te maken, zodat veel ruis wegvalt, nu is er vaak veel interpretatie nodig om een onderdeel van de referentieniveaus concreet te maken. De inhoud van de referentieniveaus hoeft misschien niet eens echt aangepast te worden, de beschrijving ervan wel. Tegelijk zou de afstemming met de eerste twee leerjaren van het voortgezet onderwijs nu eindelijk eens goed geregeld moeten worden, op een concreet niveau. Nu zijn de verschillen tussen het beheersingsniveau van de leerlingen van verschillende basisscholen vaak groot, zodat alles nog maar een keer opnieuw aangeboden wordt. Een beter begrip van waar een leerling staat (dus betere toetsen), gecombineerd met heldere doelen voor het leerstofaanbod en een veel flexibeler systeem in de onderbouw van het voortgezet onderwijs zou al helpen. Op deze manier zouden alle leerlingen

een eerlijke kans moeten krijgen om op de juiste plek in het voortgezet onderwijs te starten en zich te ontwikkelen.

Kansenongelijkheid vergroot door referentieniveaus en eindtoets?

We staan in een lange traditie van de zoektocht naar een eerlijke doorverwijzing naar het voortgezet onderwijs. Er zijn geen makkelijke oplossingen, veel is al geprobeerd (bijvoorbeeld alleen het oordeel van de schooldirecteur, toelatingsexamens etc), maar is de eindconclusie niet dat we gewoon te vroeg, te definitief, leerlingen willen indelen? Blokkeren we geen mogelijkheden en kansen van leerlingen door ze te snel in een bepaalde richting te plaatsen? Deze discussie zullen we denk ik veel intensiever moeten voeren de komende tijd. Graag verwijs ik hierbij ook naar het indrukwekkende proefschrift van Karen Heij over dit onderwerp: <https://www.parrhesiaonderwijsadvies.nl/mijn-publicaties/van-de-kat-en-de-bel>. Zij schetst een verhelderend beeld over het ontstaan en de praktijk van de eindtoets en welke gevolgen dit heeft voor kansengelijkheid.

Ik ga graag in gesprek met wie mee wil denken over dit boeiende onderwerp.

Maarten van der Steeg

April 2022

info@nieuwleren.nl

Bronnen:

De referentiesets: http://www.toetsspecials.nl/html/referentiesets_openbaar/default.shtm

https://www.expertgroepoetsenpo.nl/uimg/ept/b65275_att-rapport_vergelijkbaarheid-eindtoetsen_161216.pdf

<https://www.rijksoverheid.nl/documenten/rapporten/2018/12/01/vergelijkbaarheid-eindtoetsen-en-invoering-van-gezamenlijk-anker>

<https://didactiefonline.nl/artikel/eindtoets-meten-met-twee-maten>

<https://didactiefonline.nl/artikel/wie-toetst-de-toetsmaker>

<https://didactiefonline.nl/artikel/eindtoets-po-de-belofte-van-een-gelijk-speelveld-voor-toetsaanbieders>

<https://didactiefonline.nl/artikel/exit-de-centrale-eindtoets>

<https://www.trouw.nl/nieuws/vanaf-morgen-is-het-weer-tijd-voor-de-citotoets-al-gebruikt-de-helft-van-de-scholen-een-andere-eindtoets~bba9b073/>

<https://www.trouw.nl/nieuws/en-nu-moet-de-cito-toets-juist-weer-belangrijker-worden-vindt-het-cpb~b65a48d4/>

[Rapportage referentiesets Nederlandse taal \(lezen\) en rekenen: verantwoording project](#)

<https://www.onderwijsinspectie.nl/binaries/onderwijsinspectie/documenten/rapporten/2016/04/13/taal-en-rekenen-aan-het-einde-van-de-basisschool/taal-en-rekenen-aan-het-einde-van-de-basisschool.pdf>

<https://www.rijksoverheid.nl/binaries/rijksoverheid/documenten/rapporten/2020/03/31/peilingsonderzoek-taal-en-rekenen-in-het-basisonderwijs-2018-2019/peilingsonderzoek-taal-en-rekenen-in-het-basisonderwijs-2018-2019.pdf>

<https://www.parrhesiaonderwijsadvies.nl/mijn-publicaties/van-de-kat-en-de-bel>.