

Ministerie van Onderwijs, Cultuur en
Wetenschap

>Retouradres Postbus 16375 2500 BJ Den Haag

De Voorzitter van de Tweede Kamer der Staten-Generaal
Postbus 20018
2500 EA DEN HAAG

**Wetgeving en Juridische
Zaken**

Rijnstraat 50
Den Haag
Postbus 16375
2500 BJ Den Haag
www.rijksoverheid.nl

Onze referentie
1454840

Datum 13 december 2018
Betreft Beantwoording vragen schriftelijk overleg Toetsbesluit PO (Kamerstuk
31 293, nr. 416)

Hierbij bied ik u aan de beantwoording van de vragen die zijn gesteld in het schriftelijk overleg over het ontwerpbesluit tot wijziging van het Toetsbesluit PO in verband met het realiseren van een goede onderlinge vergelijkbaarheid van de eindtoetsen in het basisonderwijs en speciaal onderwijs.

De Minister voor Basis- en Voorgezet Onderwijs en Media,

Arie Slob

II Reactie minister voor Basis- en Voortgezet Onderwijs en Media

1454840

Graag wil ik de leden van de vaste commissie voor Onderwijs, Cultuur en Wetenschap danken voor de vragen die ze hebben gesteld en de opmerkingen die zijn gemaakt.

Bij deze beantwoording wordt de indeling naar paragrafen aangehouden van de inbreng van het schriftelijk overleg. Ik zal eerst de vragen beantwoorden die gesteld zijn onder het hoofdstuk 'Inleiding'. Hierna worden de vragen onder 'Aanleiding', 'Eindtoetsen onvoldoende vergelijkbaar', 'Wijzigingen Toetsbesluit PO in verband met verbeteren vergelijkbaarheid eindtoetsen' en 'Artikel II. Inwerkingtreding' behandeld. De vragen zijn binnen deze hoofdstukken inhoudelijk gegroepeerd.

Inleiding

De leden van de CDA-fractie vragen de minister nader toe te lichten hoe het doel van de in het ontwerpbesluit genoemde onderlinge vergelijkbaarheid zich verhoudt tot het feit dat de eindtoets niet bedoeld is om scholen onderling te vergelijken, maar een tweede objectief gegeven is voor ouders en leerlingen bij het schooladvies voor het voortgezet onderwijs. De leden van de CDA-fractie vragen of de minister in zijn toelichting aan kan geven of het in dit kader dan ook niet belangrijker is dat de schooladviezen op een vergelijkbare manier tot stand komen en onderling vergelijkbaar zijn dan de eindtoetsen?

Het is belangrijk dat de resultaten van de eindtoetsen onderling vergelijkbaar zijn. Dit draagt namelijk bij aan het primaire doel van de eindtoets: het geven van een toetsadvies dat kan dienen als tweede, objectief gegeven naast het schooladvies. Het is belangrijk dat de verschillende toetsen valide en betrouwbaar zijn en onderling goed vergelijkbare resultaten opleveren, zowel onderling als over de jaren heen. Enkel wanneer de resultaten van de toetsen onderling vergelijkbaar zijn, kan een uitspraak worden gedaan of verschillende leerlingen bij een gelijke prestatie ook een gelijke uitkomst op de eindtoets hebben. Welke toets een leerling maakt zou dus niet uit moeten maken voor de score op de referentieniveaus of de uitkomst in de vorm van het toetsadvies.

De leden van de CDA-fractie vragen ook of het belangrijker is dat schooladviezen op een vergelijkbare manier tot stand komen en onderling vergelijkbaar zijn dan de eindtoetsen.

Dit is niet zo. Scholen zijn vrij om zelf te bepalen hoe het schooladvies tot stand komt. Hierin is het professionele oordeel van de leerkracht leidend, wat ik ook belangrijk vind. Het opstellen van het schooladvies vraagt namelijk om maatwerk van basisscholen, niet om overheidsinterventie op nationaal niveau. De manier waarop de schooladvisering tot stand komt wordt hiernaast tegen het licht gehouden in de eindevaluatie van de Wet Eindtoetsing PO.

De leden van de D66-fractie vinden dat in het genoemde ontwerpbesluit bij de eindtoets teveel de nadruk ligt op meetbaar beleid en te weinig op de ontwikkeling van een kind. Deze leden vragen de minister nader toe te lichten wat hij bedoelt met dat eindtoetsen in de praktijk onvoldoende vergelijkbaar zijn en dit op orde moet worden gebracht. Deze leden menen dat verschillende meetinstrumenten juist bijdragen aan de validiteit en betrouwbaarheid van resultaten. Zij vragen een reactie hierop van de minister.

In de nota van toelichting op het ontwerpbesluit wordt in paragraaf 4 in de onderdelen a tot en met c beschreven op welke manier de vergelijkbaarheid van de resultaten van de eindtoetsen door middel van dit ontwerpbesluit wordt verbeterd. Allereerst moeten de eindtoetsen dezelfde categorieën aanhouden in de toetsadviezen, dus dezelfde lijst van schoolsoorten voor het vervolgonderwijs. Ten tweede dienen de verschillende eindtoetsen een gelijke normering te hebben en dienen, bij gelijke prestaties, tot gelijke scores te leiden ten aanzien van de beheersing van taal en rekenen. Als derde dient ook de vergelijkbaarheid van de resultaten van de eindtoetsen over de jaren heen op orde te zijn. Zonder anker is er geen instrument om een uitspraak te doen over de vergelijkbaarheid van de resultaten van de eindtoetsen. Het invoeren van een anker stelt de Expertgroep Toetsen PO (hierna kortweg 'expertgroep') in staat hier wél een uitspraak over te doen.

In zijn algemeenheid klopt het dat verschillende meetinstrumenten naast elkaar kunnen bijdragen aan de validiteit en betrouwbaarheid van de resultaten. Maar voor de eindtoets geldt dat iedere leerling er daarvan maar één maakt. Er worden wat betreft de eindtoets per leerling dus nooit meer meetinstrumenten naast elkaar afgenomen. Een school kiest vooraf welke eindtoets aan de leerlingen wordt aangeboden, waarna alle leerlingen van die school die ene eindtoets maken. Wanneer de eindtoetsen in normering van elkaar verschillen treedt juist verwarring op ten aanzien van resultaten van leerlingen op andere scholen. Tegelijkertijd blijft alle ruimte bestaan voor iedere school om die toets te kiezen die het beste aansluit bij het onderwijs op de school. Het gaat bij de voorstellen namelijk om vergelijkbaarheid van de *resultaten* van de toets, en niet om vergelijkbaarheid van de inhoud van de toets.

Aanleiding

De leden van de CDA-fractie lezen dat de Inspectie van het Onderwijs in de Staat van het Onderwijs aangeeft dat het voor haar moeilijker wordt een landelijk beeld te schetsen op basis van de resultaten van de eindtoets. Deze leden vragen de minister nader toe te lichten wat het belang hiervan is voor de Inspectie, want de eindtoetsen zijn toch niet ontworpen met als doel om scholen te kunnen beoordelen? Zij vragen of hierbij ook niet geldt dat de eindtoets slechts het tweede objectieve gegeven is en het daarom belangrijker is een goed landelijk beeld van de schooladviezen dan van de eindtoetsen te krijgen?

Het een sluit het ander niet uit: het is zowel belangrijk om een landelijk beeld van de schooladviezen als van de taal- en rekenprestaties te krijgen. Zicht op de resultaten aan het einde van het basisonderwijs en eventuele trends daarin zijn van belang voor alle betrokkenen bij het onderwijs. Het landelijk beeld van de taal- en rekenprestaties wordt bepaald aan de hand van de eindtoetsresultaten op de referentieniveaus. Daarnaast verzorgt DUO het landelijk beeld over de schooladviezen, op basis van de registraties die basisscholen doorvoeren in het Basisregister Onderwijs. De schooladviezen laten zien naar welk niveau vervolgonderwijs leerlingen geadviseerd worden om te gaan. Het landelijk beeld over de schooladviezen is dus iets anders dan het landelijk beeld over de taal- en rekenprestaties.

Uit de eindtoets volgt zowel een toetsadvies, bedoeld als tweede, objectief gegeven voor het schooladvies, als een uitspraak over de beheersing van de referentieniveaus. Deze uitspraak over de beheersing van de referentieniveaus gebruikt de inspectie om een landelijk beeld te schetsen over de taal- en

rekenprestaties van leerlingen. Hiernaast gebruikt de inspectie de eindtoetsresultaten ook als één van de onderdelen in haar kwaliteitsbeoordeling van scholen. Op deze manier bewaakt de inspectie de kwaliteit van het onderwijs. Omdat de eindtoets een toets is waarbij leerlingen vaak gemotiveerd zijn om deze goed te maken (*high stakes*) geven de uitkomsten op deze toets een goed beeld van de opbrengsten op het gebied van taal en rekenen van een basisschool.

1454840

De onvergelijkbare eindtoetsen bemoeilijkten het voor de Inspectie om een landelijk beeld van de resultaten van de eindtoets te schetsen. In hoeverre heeft deze onvergelijkbaarheid ook effect gehad op de beoordeling van scholen op basis van de resultaten van de eindtoets, zo vragen de leden van de SP-fractie. Zij vragen zich af of het kan zijn dat scholen benadeeld zijn. Ook vragen zij de minister toe te lichten of de minister daarnaast van mening is dat het resultaat van de eindtoets het juiste middel is om te beoordelen of de resultaten van een school voldoende zijn.

De eindtoetsresultaten worden door de inspectie gebruikt om een beeld te schetsen van de kwaliteit van een school. Daarbij is van belang dat de inspectie voor elk type eindtoets eigen ondergrenzen hanteert. Hierbij kijkt de inspectie naar de gemiddelde score van de leerlingen op de betreffende eindtoets. Door voor elke toets een eigen normtabel te gebruiken, houdt de inspectie zo goed mogelijk rekening met de verschillen tussen de toetsen en zijn scholen niet benadeeld. Omdat de afname van de eindtoets voor leerlingen bovendien *high stakes* is, zorgt dit voor betrouwbare en objectieve gegevens. Deze gegevens zijn echter niet het enige waar de inspectie naar kijkt bij beoordeling van de school. De eindtoetsresultaten van een school wegen mee bij de beoordeling van het onderdeel 'resultaten'. Hiernaast spreekt de inspectie zich uit over het onderwijsproces en het schoolklimaat en de kwaliteitszorg. Hiermee vormen de eindtoetsresultaten één van de indicatoren aan de hand waarvan de inspectie nagaat of de onderwijskwaliteit van voldoende niveau is. Een belangrijke ontwikkeling is het nieuwe onderwijsresultatenmodel van de inspectie, waarmee wordt proefgedraaid in schooljaar 2019/2020 en welke in schooljaar 2020/2021 officieel wordt ingevoerd.¹ Hierbij wordt niet langer naar het gemiddelde op de eindtoets gekeken, maar wordt gekeken naar de scores op de referentieniveaus voor lezen, taalverzorging en lezen. Op deze manier kunnen de resultaten nog beter door de inspectie worden gebruikt in haar kwaliteitsoordeel.

De leden van de D66-fractie lezen dat een gezamenlijk 'anker' in de eindtoetsen de onderlinge vergelijkbaarheid vergroot. Deze leden vrezen dat hierdoor de momentopname van de eindtoetsen een te grote rol gaat spelen in de beoordeling van kwaliteit van het onderwijs en de school. Zij vragen de minister om een reactie.

Het vergelijkbaar maken van de resultaten van de eindtoetsen heeft geen effect op het beoordelingskader van de inspectie. Het 'anker', dat als gevolg van de in paragrafen 'Eindtoetsen onvoldoende vergelijkbaar' en 'Wijzigingen Toetsbesluit PO in verband met verbeteren vergelijkbaarheid eindtoetsen' beschreven methode van 'zaaien' slechts ziet op een beperkt deel van de toetsopgaven, zorgt daarom ook niet voor een verandering in de wijze van beoordeling van kwaliteit van het onderwijs en de school. Het anker zorgt er wel voor dat de uitspraak van de

¹ Zie voor meer informatie:

<https://www.onderwijsinspectie.nl/onderwerpen/onderwijsresultaten-primair-onderwijs/naar-een-nieuw-onderwijsresultatenmodel>

inspectie objectiever wordt. Verder is de inspectie bezig om in de beoordeling van de onderwijskwaliteit van scholen, juist breder te kijken dan enkel alleen naar de eindtoets. Denk hierbij aan het nieuwe onderzoekskader sinds 2017, het nieuwe onderwijsresultatenmodel en de maatregel uit het regeerakkoord: de discretionaire bevoegdheid van de inspectie in de regeling leerresultaten po, waarmee zij scholen die veel leerlingen met een bijzondere onderwijsbehoefte opvangen toch een voldoende beoordeling van de leerresultaten kunnen geven.

1454840

De leden van de SGP-fractie vragen de minister in te gaan op de verhouding tussen twee doelen van de eindtoets, namelijk enerzijds het in kaart brengen van de resultaten van de leerling ten opzichte van de referentieniveaus en anderzijds het geven van een advies ten aanzien van het vervolgonderwijs. Zij constateren dat deze doelen op gespannen voet met elkaar staan en dat het voorgenomen besluit juist de functie van de eindtoets ten behoeve van de individuele leerling onder druk lijkt te zetten. Terwijl bij het eerste doel het belang van vergelijkbaarheid wordt benadrukt, is het tweede doel juist gebaat bij ruimte voor differentiatie, zo geven deze leden aan.

De twee genoemde doelen zijn naar mijn mening goed met elkaar te verenigen. De eindtoets geeft een toetsadvies ten aanzien van het vervolgonderwijs. Hiernaast doet de eindtoets een uitspraak over de beheersing van de referentieniveaus. Voor beide doelen is het belangrijk dat de resultaten van de eindtoetsen onderling goed vergelijkbaar zijn. Het geven van een toetsadvies ten aanzien van het vervolgonderwijs is niet gebaat bij verschillen tussen toetsen qua normering, of differentiatie zoals de leden van de SGP-fractie dit noemen. Indien een toetsadvies hoger is dan het schooladvies dan moet het schooladvies worden heroverwogen. Dit proces dient voor iedere leerling gelijk te zijn, onafhankelijk van de gekozen toets en het jaar van afname, zodat de leerling, ongeacht de eindtoets die hij maakt, een gelijke kans op een heroverweging heeft.

De leden van de SGP-fractie merken allereerst op dat differentiatie al onder druk staat door het beperkte aantal vaardigheden uit het referentiekader dat getoetst wordt. Door de dominante rol van de ankeropgaven wordt bovendien de differentiatie in de wijze van toetsen verder ingesnoerd door een eenzijdige psychometrische benadering. Waarom wordt het belang van een breed schooladvies zo sterk onder druk gezet door een beperkte definitie van vergelijkbaarheid, zo vragen zij.

De wijze van totstandkoming van de schooladvisering blijft ongewijzigd. Hierin is het oordeel van de basisschool nog steeds doorslaggevend. Voor het brede schooladvies, waaraan de SGP-fractie waarde hecht, blijft daarom alle ruimte. Vanaf 2019 zullen de eindtoetsen alleen nog maar resulteren in brede toetsadviezen of in het enkelvoudige advies vwo.² Daarmee stimuleren de eindtoetsen juist tot het geven van een breed schooladvies. De SGP-fractie benoemt hierbij dat de rol van de ankeropgaven in de eindtoetsen een dominante is. Ook dit beeld deel ik niet. De meeste leerlingen krijgen in een eindtoets slechts twee of drie opgaven uit het anker voorgelegd. Hoe groter de groep leerlingen die een eindtoets afneemt, hoe kleiner het aantal ankeropgaven is dat zij krijgen voorgelegd. De ankeropgaven hebben daarom slechts een beperkte invloed op de inhoud van de eindtoetsen.

² Zie ook hierna het antwoord op de vraag van leden van de CDA-fractie over gelijkklidende categorieën van schoolsoorten door toetsaanbieders in paragraaf 'Wijziging Toetsbesluit PO'.

De leden van de SGP-fractie vragen ook waarom de afweging van deze belangen niet bij uitstek een thema is dat thuis hoort in de beleidsreactie op de evaluatie van de wet, mede gelet op de aangenomen motie van de leden Bisschop en Rog.³ Uit de scenario's voor ijking aan de referentieniveaus blijkt dat ook kwalitatieve winst te behalen valt zonder vooralsnog tot verplichte ankeropgaven over te gaan.

Betrouwbare en valide uitkomsten van de eindtoetsen vind ik van groot belang voor leerlingen en scholen. Ongelijkheid op dit punt mag niet langer voortbestaan dan nodig. Daarom zijn hierop naar aanleiding van de tussenevaluatie Wet Eindtoetsing PO,⁴ de evaluatie die twee jaar na de aanpassing van de Wet Eindtoetsing PO plaatsvond, direct maatregelen genomen. Over deze maatregelen heb ik u in januari 2017 geïnformeerd.⁵ Bovendien geeft de Expertgroep in zijn rapportage 'Scenario's voor ijking aan de referentieniveaus' alsmede in het rapport 'Vergelijkbaarheid van de eindtoetsen' juist aan wat de noodzaak is voor de invoering van het anker. Er is niet genoeg kwalitatieve winst te halen als het anker niet wordt ingevoerd. Het is niet mogelijk om op een andere manier dan via een anker te zorgen voor onderling vergelijkbare eindtoetsen.

In voornoemde motie van de leden Bisschop en Rog is de regering verzocht om, in de beleidsreactie op de evaluatie van de eindtoets, in te gaan op de wijze waarop de eindtoets recht kan gaan doen aan alle vaardigheden van de Wet referentieniveaus, op een veelzijdiger wijze dan enkel met meerkeuzevragen. Deze motie kijkt inhoudelijk naar de referentieniveaus en naar de manier van vraagstellen. Beide punten hebben echter geen directe relatie met de wijzigingen in dit ontwerpbesluit. Deze overwegingen zal ik daarom meenemen in de eindevaluatie van de Wet Eindtoetsing PO. Deze zal ik u voor het zomerreces doen toekomen.

Eindtoetsen onvoldoende vergelijkbaar

De leden van de SP-fractie vragen de minister of het daadwerkelijk zo is dat de eindtoets als tweede, objectief gegeven dient bij het schooladvies voor het voortgezet onderwijs, terwijl uit cijfers van het Centraal Bureau voor Statistiek (CBS) blijkt dat in schooljaar 2016-2017 bijna 58.000 leerlingen na de uitslag van de eindtoets in het reguliere basisonderwijs in aanmerking kwamen voor een bijstelling naar boven van het oorspronkelijke leerkrachtadvies, maar dit advies slechts bij twee op de tien daadwerkelijk naar boven is bijgesteld. Deze leden willen graag weten wat de minister van deze cijfers vindt en welke oorzaken hij ziet.

De eindtoets dient inderdaad als tweede, objectief gegeven bij het schooladvies. Zoals ik ook heb toegelicht in eerdere antwoorden op uw Kamervragen laat het feit dat bijna een kwart van de adviezen wordt bijgesteld, zien dat scholen het resultaat van de toets als onafhankelijk tweede gegeven serieus nemen en serieus

³ Kamerstukken 34 775-VIII, nr. 101.

⁴ Wet van 11 december 2013 tot wijziging van onder meer de Wet op het primair onderwijs, de Wet op de expertisecentra en de Wet College voor examens in verband met de invoering van een centrale eindtoets, de invoering van een leerling- en onderwijsvolgsysteem en invoering van bekostigingsvoorschriften voor minimumleerresultaten voor speciale scholen voor basisonderwijs en scholen voor speciaal onderwijs en speciaal en voortgezet speciaal onderwijs (centrale eindtoets en leerling- en onderwijsvolgsysteem primair onderwijs) (Stb. 2014, 13).

⁵ Kamerstukken II 2016/17, 31289, nr. 343.

de mogelijkheid van bijstelling overwegen.⁶ De trends rondom het percentage bijstellingen worden verder toegelicht in de Kamerbrief eindtoetsresultaten 2018 die ik recent naar uw Kamer heb gestuurd.⁷

1454840

De leden van de CDA-fractie vragen de minister nader toe te lichten waarom het nodig is dat eindtoetsen onderling vergelijkbaar zijn en of dit niet op gespannen voet staat met de wettelijke vrijheid van scholen om te kiezen voor een eindtoets die het beste bij hen past: hetzij de Cito-eindtoets, hetzij een alternatieve eindtoets? Deze leden vragen de minister of een betere onderlinge vergelijkbaarheid niet het gevaar in zich heeft van dusdanig vergaande standaardisering dat de toegevoegde waarde van de alternatieve toetsen, namelijk een gelijkwaardig alternatief bieden voor scholen dat beter bij hun profiel of pedagogisch-didactische concept past dan de Cito-eindtoets, hiermee steeds kleiner wordt?

Het is belangrijk dat de resultaten van de eindtoetsen onderling vergelijkbaar zijn omdat het toetsadvies voor leerlingen dient als het objectief, tweede gegeven bij het schooladvies. Indien het toetsadvies hoger is dan het schooladvies, moet de school het schooladvies heroverwegen. De kans op heroverweging moet voor elke leerling gelijk zijn, onafhankelijk van de eindtoets die de school heeft gekozen en het jaar van afname. De basis hiervoor is artikel 4, onderdeel b, van het Toetsbesluit PO waarin staat dat 'de toets inhoudelijke valide (en), betrouwbaar (is) en een deugdelijke normering' heeft.

Dit laat voor scholen voldoende vrijheid om te kiezen voor een eindtoets die het beste bij hen past. Ook brengt een betere onderlinge vergelijkbaarheid van de resultaten van de toetsen de toegevoegde waarde van de alternatieve eindtoetsen niet in gevaar. Hier zijn twee redenen voor. Allereerst heeft het 'anker' maar een beperkt effect op de eindtoetsen. Als gevolg van de te gebruiken methode van 'zaaien' maken de meeste leerlingen slechts twee of drie opgaven uit het anker. Hoe meer leerlingen een eindtoets maken, hoe minder opgaven van het anker zij hoeven te maken. Hiernaast zorgt het anker enkel voor vergelijkbaarheid van de *resultaten* van de eindtoets. Meer specifiek zorgt het anker er voor dat de uitkomsten op de referentieniveaus en van de toetsadviezen vergelijkbaar zijn. Toetsen kunnen op allerlei andere manieren, bijvoorbeeld in vraagstelling en eventuele extra modules van elkaar verschillen, om daarmee goed aan te sluiten op de pedagogisch-didactische visie van een school.

De toegevoegde waarde van de alternatieve toetsen blijft hiermee dus onverminderd groot. Op dit moment kunnen scholen kiezen uit vijf verschillende eindtoetsen, waaronder de Centrale Eindtoets die onder verantwoordelijkheid van het CvTE door Stichting Cito wordt ontwikkeld. Daarnaast hecht ik eraan dat alle eindtoetsen, zowel die van de overheid als die van private partijen van goede kwaliteit zijn. Ik vind het vanuit het oogpunt van keuzevrijheid ongewenst om de Centrale Eindtoets als maatstaf ten opzichte van de andere eindtoetsaanbieders te hanteren.

De leden van de CDA-fractie vragen de minister ook toe te lichten wat en op welke onderdelen de eigen invulling voor de alternatieve toetsaanbieders zal zijn na aanneming van dit Toetsbesluit. Waarin zullen deze alternatieve eindtoetsen nog kunnen afwijken van de Cito-eindtoets, zo vragen deze leden.

⁶ Aangangsels Handelingen II 2018/19, 685.

⁷ Kamerstukken II 2018/19, 31293, nr. 422.

Zoals hiervoor uiteengezet hebben de maatregelen in het ontwerpbesluit slechts beperkte invloed op de inhoud van de eindtoetsen. De toetsaanbieders houden, behoudens het beperkte 'anker', op verschillende manieren ruimte voor eigen invulling van de eindtoetsen. Denk hierbij aan de volgende onderdelen:

- o Afnamevorm: papier of digitaal;
- o Inhoudelijk: welke onderdelen naast taal en rekenen worden aangeboden (bijvoorbeeld andere referentieniveaus of facultatieve modules als werkhouding en motivatie);
- o Vraagvorm: open, gesloten, *multiple choice*;
- o Vraagregistratie: in het boekje, op een los antwoordvel, digitaal;
- o Niveau: adaptief of op één niveau;
- o Uiterlijk: de look-en-feel van de toets.

De leden van de D66-fractie menen dat het meten van referentieniveaus voldoende ijkpunt vormt om de onderlinge vergelijkbaarheid te vergroten zonder de diversiteit uit de meetinstrumenten te halen. De minister geeft aan dat uit praktijkervaring blijkt dat verschillende eindtoetsen er niet in zijn geslaagd op gelijke wijze referentieniveaus te meten. Deze leden vragen de minister om eerst de producenten van de eindtoetsen aan te spreken op dit gebrek, alvorens een wijziging toe te passen in de wet.

Er zijn hierover inderdaad vele gesprekken gevoerd, ook tussen de Expertgroep en de eindtoetsaanbieders. Deze gesprekken heb ik inderdaad – en veelvuldig – met de eindtoetsaanbieders gevoerd. Het klopt dus dat de eindtoetsaanbieders er niet in zijn geslaagd om de referentieniveaus zonder anker op gelijke wijze te meten. Het is inderdaad het doel de referentieniveaus als ijkpunt binnen de eindtoetsen te gebruiken. Door het anker in te voeren zorgen we er voor dat de normeringen op de referentieniveaus door alle toetsaanbieders alsnog op dezelfde wijze worden gedaan, zodat de uitspraken over de referentieniveaus ook gelijk zijn. Om die reden is in samenspraak met de eindtoetsaanbieders afgesproken om een anker met nieuwe vragen in de eindtoetsen op te nemen. De eindtoetsaanbieders dragen allen opgaven aan voor het nieuwe anker.

De leden van de SGP-fractie lezen dat de veronderstelling van de wetgever was dat ijkking van de eindtoetsen aan de referentieniveaus voldoende waarborg zou vormen voor betrouwbaarheid en vergelijkbaarheid. Zij constateren dat de minister met dit conceptbesluit meteen ook grijpt naar het instrument van ankeropgaven ondanks het feit dat alternatieven bestaan die beter aansluiten bij de bedoeling en verwachting van de wetgever. Hoe is dit vanuit het oogpunt van proportionaliteit te rechtvaardigen en waarom is het niet geboden om op basis van de evaluatie, waarin ook de door de Kamer gevraagde alternatieven aan bod komen, eerst ten principale de bedoeling van de wet te bespreken?

Ik heb verschillende alternatieven overwogen om de eindtoetsen onderling vergelijkbaar te maken. Dit heb ik gedaan naar aanleiding van de tussenevaluatie van de Wet Eindtoetsing PO, waarover ik u in januari 2017 heb geïnformeerd.⁸ In de bijlage bij deze Kamerbrief is opgenomen het advies van de onafhankelijke Expertgroep Toetsen PO. In dit advies worden verschillende opties genoemd en afgewogen. De Expertgroep adviseert hierin het invoeren van een nieuw anker in de eindtoetsen, omdat dit de meest effectieve oplossing is. Naar aanleiding van de Tussenevaluatie worden nu maatregelen genomen om de vergelijkbaarheid van de resultaten van de eindtoetsen te verbeteren. Vanwege het grote belang

⁸ Kamerstukken II 2016/17, 31289, nr. 343.

van betrouwbare en onderling vergelijkbare eindtoetsen, is invoering van het anker in het Toetsbesluit PO noodzakelijk en proportioneel.

1454840

Wijzigingen Toetsbesluit PO in verband met verbeteren vergelijkbaarheid eindtoetsen

De leden van de CDA-fractie lezen in het ontwerpbesluit dat er de afgelopen jaren zorgen zijn gerezen over hoe bij de verschillende eindtoetsen de referentiekaders als ijkpunt worden genomen; deze bleken in de praktijk van elkaar te verschillen. Deze leden vragen de minister dan ook toe te lichten waarom niet samen met de toetsaanbieders en toetsafnemers en andere betrokken deskundigen is gekeken hoe kan worden geborgd dat dit wel op een goede wijze gaat gebeuren? Want als deze ijking goed gebeurt, dan is dat toch voldoende waarborg voor betrouwbaarheid en goede vergelijkbaarheid door de jaren heen? Waarom heeft de minister dit niet gedaan, maar gelijk gekozen voor het via 'ankers' beter onderling vergelijkbaar maken, zo vragen deze leden.

Voor het antwoord op deze vraag wil ik graag verwijzen naar het antwoord hierover op de vraag van de leden van de D66-fractie in de paragraaf 'Eindtoetsen onvoldoende vergelijkbaar'. De eindtoetsaanbieders hebben de normering van de referentieniveaus niet allemaal op dezelfde manier kunnen overbrengen op hun toetsen. Daarmee is – zoals hierboven bij de vraag van de leden van de D66-fractie over het meten van referentieniveaus ook uiteengezet – de gebruikte methode van normeringen op de referentieniveaus niet meer bruikbaar om de resultaten op de toetsen onderling vergelijkbaar te maken. Zie ook het hiervoor genoemde advies van de Expertgroep Toetsen PO dat ik u in januari 2017 aan uw Kamer toezond.⁹

De leden van de CDA-fractie lezen dat de Expertgroep heeft aangegeven dat er drie opties zijn voor het bovengenoemde 'anker', te weten een gezamenlijk 'anker' (gezamenlijke set aan opgaven) invoeren in de pretests; een gezamenlijk 'anker' invoeren in zowel pretests als in operationele toetsen; of een gezamenlijk 'anker' invoeren in de operationele toetsen. Volgens de Expertgroep heeft invoering van de derde optie het grootste effect op de vergelijkbaarheid van de toetsen. Deze leden vragen de minister een nadere onderbouwing te geven voor de keuze voor de derde optie alsmede een opsomming van de voor- en nadelen van de eerste en tweede optie zodat duidelijk wordt op basis van welke argumenten de derde optie als meest geschikt wordt beschouwd door de Expertgroep en de minister.

In de voornoemde 'rapportage vergelijkbaarheid eindtoetsen' beveelt de Expertgroep Toetsen PO aan om in alle operationele eindtoetsen een gezamenlijk anker op te nemen als deel van iedere eindtoets.¹⁰ In het rapport 'Scenario's van ijking van de referentieniveaus' zet de Expertgroep uiteen wat de voor- en nadelen zijn van het zaaien van het anker in de pretest, in de operationele toets, of in beiden.¹¹ Deze laten zich als volgt samenvatten:

⁹ Glas, C.A.W., Emons, W.H.M., & Berding-Oldersma, P.K. (2016) *Scenario's voor ijking van de eindtoetsen op de referentieniveaus*. Utrecht: Expertgroep Toetsen PO, p. 9.

¹⁰ Emons, W.H.M., Glas, C.A.W., & Berding-Oldersma, P.K. (2016) *Rapportage vergelijkbaarheid eindtoetsen*. Utrecht: Expertgroep Toetsen PO, p.20.

¹¹ Glas, C.A.W., Emons, W.H.M., & Berding-Oldersma, P.K. (2016) *Scenario's voor ijking van de eindtoetsen op de referentieniveaus*. Utrecht: Expertgroep Toetsen PO, p. 13.

Het voordeel van het opnemen in het gezamenlijk anker in de operationele toetsen is dat de ankeropgaven in een *high stakes* situatie worden afgenomen. Dit zorgt er voor dat de ankeritems veel beter functioneren dan wanneer zij enkel in de pretest worden gebruikt omdat de leerlingen de items in een operationele toets veel gemotiveerder beantwoorden dan in een relatief vrijblijvende pretest. De nadelen van ankeren via de pretesten waarbij het anker vooraf is samengesteld zijn het genoemde mogelijk gebrek aan motivatie van de leerlingen en het eventueel uitlekken van items. Het voordeel is dat de eigenschappen van de items voor de eigenlijke afname van de eindtoets bekend zijn zodat uit de pretest de best functionerende items voor het gezamenlijke anker gekozen kunnen worden. De voor- en nadelen afwegend is voor het schatten van de eigenschappen van het anker gekozen voor een oplossing die een combinatie is van pretesten en testen in de operationele fase. Zo werden de ankeritems van 2019 gepretest in de operationele afname van 2018 en worden de ankeritems voor 2020 gepretest in de operationele afname van 2019 door deze items in de operationele eindtoets te zaaien. Dit houdt in dat iedere leerling twee á drie items van het anker krijgt voorgelegd die niet meetellen voor de eindtoetsscore. Voor de leerling zijn deze items gewone opgaven, net als de overige opgaven. Hierdoor worden er voldoende gegevens verzameld om een kwalitatief goed anker samen te stellen waarmee alle eindtoetsen wederom op een gezamenlijke schaal kunnen worden geplaatst.

1454840

De leden van de CDA-fractie zien het belang van een gelijke moeilijkheidsgraad door de jaren heen van een eindtoets. Deze leden vragen de minister nader toe te lichten waarom hij van mening is dat deze gelijke moeilijkheidsgraad alleen te bereiken is met het door hem gekozen 'anker' en niet door het verbeteren van de borging van de referentieniveaus als ijkpunten. Deze leden vragen de minister onderbouwd aan te geven of de toegevoegde waarde van de alternatieve eindtoetsen, te weten beter aansluiten bij het profiel of de pedagogisch-didactische visie van de school, beter gewaarborgd wordt door het verbeteren van de borging van de referentieniveaus als ijkpunt dan door het voorgestelde anker. In mijn vorige antwoord ben ik ingegaan op de noodzaak van ankeren in de operationele toets. Graag verwijs ik naar dat antwoord. Hiernaast geeft het rapport 'Scenario's voor ijking aan de referentieniveaus' weer waarom de borging van de referentieniveaus als ijkpunten niet verbeterd kan worden: "Voor het daadwerkelijke ankeronderzoek hebben enkele toetsaanbieders meegedaan aan het Head Start project, waarbij zij de op dat moment niet openbare referentiesets konden gebruiken voor de ankering. Andere toetsaanbieders hebben niet meegedaan met het Head Start project, waardoor zij de reeds openbare referentiesets moesten gebruiken voor de ankering."¹²

In de beantwoording van vragen van de leden van de CDA-fractie in de paragraaf 'Eindtoetsen onvoldoende vergelijkbaar' heb ik uiteengezet dat het anker ziet op de vergelijkbaarheid van de *resultaten* van de eindtoetsen. Hierbij behouden eindtoetsaanbieders veel ruimte voor een eigen invulling. Noch het hypothetisch verbeteren van de borging van de referentieniveaus als ijkpunt noch het voorgestelde anker hebben daarom een effect op het aansluiten van de eindtoetsen op het profiel of de pedagogisch didactische visie van een school.

¹² Glas, C.A.W., Emons, W.H.M., & Berding-Oldersma, P.K. (2016) *Scenario's voor ijking van de eindtoetsen op de referentieniveaus*. Utrecht: Expertgroep Toetsen PO, p. 9.

De leden van de SGP-fractie vragen een toelichting op de legitimatie van het voorstel in het licht van de tekst en de bedoeling van de wet. Deze leden merken dat voortdurend wordt gewezen op het belang van het ijken van toetsen aan de referentieniveaus. Zonder af te doen aan het belang van het relateren aan de referentieniveaus constateren zij dat deze verplichting tot op heden nog steeds niet in werking is getreden. Zij vragen hoe het te rechtvaardigen is dat via de route van het Toetsbesluit aanpassingen worden verricht die uitgaan van een verplichting die niet bestaat. Uit het feit dat de wetgever kennelijk heeft gemeend dat de beoogde relatie tussen de referentieniveaus en de toetsresultaten nog niet in werking dient te treden maken de leden op dat een grondslag voor het stellen van eisen in het Toetsbesluit dienaangaande ontbreekt. Zij zijn ook van mening dat de thematiek zozeer raakt aan de essentie van het systeem van de eindtoets dat het inmiddels voor de hand ligt deze eerst integraal te bespreken in het kader van de evaluatie.

De betreffende bepalingen omtrent referentieniveaus uit de Wet Eindtoetsing PO zijn nog niet direct per 1 augustus 2014 in werking getreden om scholen voldoende tijd te geven om praktijkervaring op te doen met het werken met de sinds 1 augustus 2010 verplichte referentieniveaus. Om een soepele overgang te kunnen realiseren naar een eindtoets die verplicht rapporteert over de vorderingen van leerlingen en groepen leerlingen ten opzichte van de referentieniveaus, hebben de eindtoetsaanbieders de periode vanaf 1 augustus 2014 benut om ervaring op te doen met het meten van de referentieniveaus. In praktijk heeft daarom sinds de invoering van de Wet Eindtoetsing PO wel steeds ijking van de eindtoetsen plaatsgevonden aan de referentieniveaus. Het beoordelingskader van de Expertgroep bepaalt zijn advies ook op het feit of de eindtoetsen rapporteren over de referentieniveaus. Inmiddels hebben scholen ruim de tijd gekregen om praktijkervaring op te doen met het werken met de verplichte referentieniveaus. Daarom zal ik de reeds tot stand gekomen wettelijke bepaling per 1 augustus 2019 in werking laten treden door een te treffen inwerkingtredingsbesluit. Op deze manier komt de wet in overeenstemming met de staande praktijk.

Betrouwbare en qua moeilijkheid vergelijkbare eindtoetsen zijn bovendien van groot belang voor zowel leerlingen als scholen. Ongelijkheid op dit punt wil ik niet langer laten voortbestaan dan nodig. Daarom zijn hierop direct naar aanleiding van de Tussenevaluatie van de Wet Eindtoetsing PO maatregelen genomen, die met dit ontwerpbesluit worden verankerd. Ik heb de Kamer geïnformeerd over deze maatregelen en daarbij ook aangegeven dat wijziging van het Toetsbesluit PO zou volgen.¹³

De leden van de CDA-fractie lezen in het ontwerpbesluit dat voortaan de toetsaanbieders gelijklopende categorieën van schoolsoorten gaan hanteren. Deze leden vragen een toelichting van de minister wat deze verschillen precies waren. Hoe konden dezelfde schoolsoorten onder verschillende categorieën worden gehanteerd: de namen van de categorieën staan toch vast? Tevens vragen de leden van de CDA-fractie een bevestiging van de minister dat ook in het nieuwe systeem een gecombineerd advies kan worden gegeven zoals vmbo-tl/havo of havo-vmbo.

De leden van de CDA-fractie vragen of dezelfde schoolsoorten onder verschillende namen werden gehanteerd. Dit is niet het geval. De namen van de categorieën staan inderdaad vast, juist om verwarring hierover bij scholen en ouders te

¹³ Kamerstukken II 2016/17, 31289, nr. 343.

voorkomen. De verschillen in 2018 en hiervoor richtten zich voornamelijk op de (combinatie van) de vmbo-adviescategorieën. Zo bevatten de Centrale Eindtoets en de IEP Eindtoets in 2018 niet het enkelvoudige advies 'vmbo bb'. Zij bevatten wel de meervoudige adviezen 'pro/vmbo bb' en 'vmbo bb/ vmbo kb'. Ook bevatte de Centrale Eindtoets in 2018 niet het meervoudige advies 'vmbo kb/ vmbo gl-tl'. De Centrale Eindtoets bevatte wel de enkelvoudige adviezen 'vmbo kb' en 'vmbo gl-tl'. Dit kon leiden tot onduidelijkheid voor scholen, leerlingen en ouders. Ik bevestig dat ook in het nieuwe systeem een gecombineerd toetsadvies kan worden gegeven. Vanaf het schooljaar 2018/2019 zijn de adviescategorieën van alle eindtoetsen gelijk, namelijk:

- pro/vmbo bb
- vmbo bb/vmbo kb
- vmbo kb/vmbo gl-tl
- vmbo gl-tl/havo
- havo/vwo
- vwo¹⁴

Het definitieve schooladvies van de basisschool kan zowel een enkelvoudig als een meervoudig schooladvies inhouden.

De verschillende eindtoetsen dienen in hun toetsadviezen dezelfde categorieën te hanteren van schoolsoorten voor het vervolgonderwijs. De leden van de SP-fractie vragen de minister waarom het praktijkonderwijs niet als advies uit de eindtoets kan komen.

Om toe gelaten te worden tot het praktijkonderwijs is een toelaatbaarheidsverklaring van het samenwerkingsverband nodig. Het toetsadvies 'praktijkonderwijs' kan niet uit de eindtoets komen omdat de eindtoets geen voldoende geschikt instrument is om te bepalen of een leerling toegelaten kan worden tot het praktijkonderwijs.

De leden van de CDA-fractie vragen de minister waarom het nodig is om niet alleen dezelfde eindtoets van jaar tot jaar met elkaar te vergelijken, maar ook verschillende eindtoetsen.

Het is belangrijk om de verschillende eindtoetsen met elkaar te vergelijken omdat bij een gelijk niveau van een leerling dit op de eindtoets moet leiden tot zowel een gelijke uitspraak op de beheersing van de referentieniveaus als tot een gelijk toetsadvies. Indien een toetsadvies hoger is dan het schooladvies dan moet het schooladvies worden heroverwogen. Dit proces dient voor iedere leerling gelijk te zijn, onafhankelijk van de gekozen toets en het jaar van afname, zodat in gelijke gevallen op gelijke wijze heroverweging plaatsvindt van het schooladvies. Welke toets een leerling maakt zou niet uit moeten maken voor de score op de referentieniveaus of de uitkomst in de vorm van het toetsadvies. Om dit te verzekeren is niet alleen een vergelijking over de jaren, maar ook tussen de verschillende eindtoetsen nodig. Op de doelen van de eindtoets ben ik ook ingegaan bij mijn antwoorden in paragrafen 'Aanleiding' en 'Eindtoetsen onderling onvoldoende vergelijkbaar' naar aanleiding van vragen van respectievelijk de leden van de SGP-fractie en leden van de CDA-fractie.

¹⁴ De afkortingen horen respectievelijk bij de schoolsoorten praktijkonderwijs; vmbo basisberoepsgerichte leerweg; vmbo kaderberoepsgerichte leerweg, vmbo gemengde leerweg /vmbo theoretische leerweg; havo en vwo.

De leden van de CDA-fractie lezen dat door de introductie van een set van opgaven die voor alle toetsaanbieders hetzelfde moet zijn – een zogenoemd 'anker' – de onderlinge vergelijkbaarheid wordt geborgd. Deze leden vragen de minister om een nadere toelichting, op wat dit 'anker' precies betekent en in hoeverre verschilt dit van de huidige praktijk. Ook vragen deze leden hoe deze gelijke set van opgaven zich verhoudt tot de vrijheid van scholen om te kiezen voor een toets die het beste bij hun profiel of pedagogisch-didactische concept past en op welke wijze dit dan nog mogelijk is na inwerkingtreding van dit ontwerpbesluit. Ook wensen deze leden een uitgebreide toelichting op de vraag of de invoering van dit 'anker' niet leidt tot dusdanig vergaande standaardisering dat er de facto sprake zal zijn van 'Citoïsering' van de schooltoetsen. Dit anker betekent dat in alle eindtoetsen een set van dezelfde vragen wordt opgenomen. De invoering van dit anker leidt niet tot vergaande standaardisering; de ankeropgave hebben slechts een geringe invloed op de inhoud van de toets. In de beantwoording van vragen van de leden van de CDA-fractie in de paragraaf 'Eindtoetsen onvoldoende vergelijkbaar' is op dit onderwerp reeds een uitgebreide toelichting gegeven. Ook wordt in deze antwoorden beschreven hoe deze ankerset zich verhoudt tot de vrijheid van scholen om te kiezen voor een toets die het beste bij hun profiel of pedagogisch-didactisch concept past.

De leden van de SGP-fractie hebben de indruk dat de keuzes ten aanzien van de ankeropgaven leiden tot een rekolonisatie van de eindtoets door Cito. Deze leden hebben namelijk begrepen dat de inhoud van de ankeropgaven in ieder geval het eerste jaar afkomstig is geweest van Cito. Zij vragen waarom niet in een zoveel mogelijk gelijkwaardige positie van toetsaanbieders wordt voorzien als het gaat om ankeropgaven, zoals voorgesteld in de scenario's voor ijking van eindtoetsen. En indien dat voornemen voor de toekomst wel bestaat, waarom wordt niet in het besluit gewaarborgd dat alle aanbieders, mogelijk naar rato van hun aandeel, kunnen voorzien in een deel van de ankeropgaven, zo vragen deze leden de minister.

Voor het anker van 2018 en 2019 is gebruik gemaakt van opgaven die Stichting Cito reeds had ontwikkeld voor de Centrale Eindtoets, omdat de andere eindtoetsaanbieders op dat moment nog niet over geproefde opgaven beschikten. Voordat een opgave in een eindtoets kan worden gebruikt moet deze niet alleen ontwikkeld worden, maar ook worden geproefde. Bij voorkeur gebeurt dit onder vergelijkbare condities als bij de wettelijke afname van de eindtoets (*high stakes*). Alle eindtoetsaanbieders hebben gezamenlijk besloten om de eerste twee jaar de opgaven die reeds door stichting Cito waren ontwikkeld te gebruiken.

Het proces van het ontwikkelen van de ankeropgaven tot opname in de eindtoets duurt ongeveer twee jaar. Om die reden zijn de eindtoetsaanbieders dit jaar al gestart met het ontwikkelen van een gezamenlijk anker voor de eindtoets van 2020. Iedere toetsaanbieder draagt daarvoor opgaven aan. Deze opgaven worden dit schooljaar geproefde. Dit gezamenlijk anker zal worden opgenomen in de eindtoets van 2020.

Er is hierbij niet gekozen voor het aanbieden van ankeropgaven naar rato van het aandeel van de verschillende toetsaanbieders. Iedere aanbieder heeft opgaven aangeleverd, hoe de verdeling precies wordt gemaakt, laat ik aan de toetsaanbieders, en dat gebeurt in goed onderling overleg. Het aandeel van de eindtoetsaanbieders varieert over de jaren en is daarom, los van een ontbrekende noodzaak, lastig in regelgeving te vervatten.

De leden van de SGP-fractie vragen of de minister onderkent dat ook de wijze van toetsen en presenteren van opgaven een belangrijk argument was voor het toelaten van alternatieve eindtoetsen. Deze leden vragen de minister, zeer concreet te beschrijven welke ruimte de toetsaanbieders hebben om de ankeropgaven zodanig vorm te geven dat zij passen bij het onderscheidende karakter van hun toets. Eveneens vragen zij waarom het besluit hieromtrent geen bepalingen bevat.

1454840

De wijze van toetsen en presenteren van opgaven maakt geen deel uit van het beoordelingskader van de Expertgroep. Daarmee is dit ook geen argument in het toelaten van alternatieve eindtoetsen. Wel wordt in het beoordelingskader gekeken naar de validiteit van de opgaven en naar helderheid in taalgebruik. Het is van belang dat elke eindtoetsaanbieder de lay-out van de ankeropgaven zodanig aanpast aan de lay-out van de eigen toets, dat een opgave uit de ankeropgaven niet te onderscheiden is van een opgave die een eindtoetsaanbieder zelf heeft ontwikkeld. Aanpassingen dienen echter zo minimaal mogelijk plaats te vinden, om uit te sluiten dat verschillen in scores worden veroorzaakt door verschillen in opmaak of formulering. Om deze reden hebben de eindtoetsaanbieders afgesproken dat zij deze aanpassingen zo beperkt mogelijk houden. De voorgestelde bepaling laat deze ruimte, De onderlinge afspraken binnen deze ruimte zijn van een dermate gedetailleerde en specifieke aard dat dit slecht leent voor vastlegging in een algemeen verbindend voorschrift. De noodzaak daartoe is ook niet aanwezig.

De leden van de CDA-fractie lezen dat in het huidige Toetsbesluit nog staat dat volledige verversing van alle toetsopgaven nodig is voor de validiteit van de toets, maar dat bij gedeeltelijke herhaling van het 'anker' deze validiteit ook voldoende gewaarborgd is door onder meer de omvangrijke digitale itembanken waaruit de digitale eindtoetsen putten, waardoor het niet van te voren bekend is welke opgaven worden gebruikt. Deze leden vragen de minister hoe voorkomen wordt dat dezelfde opgaven in de daaropvolgende jaren worden gebruikt. Zij vragen of het niet zo is dat, omdat je niet weet van te voren welke opgaven worden gebruikt, het in theorie ook zo kan zijn dat dezelfde opgaven meerdere keren achtereenvolgend worden gebruikt. Ook vragen zij of het mogelijk is om ervoor te zorgen dat dit niet kan voorkomen gedurende een bepaalde periode. Voorts willen deze leden weten of het niet nodig is om ook te voorkomen dat leerlingen die met de oude opgaven hebben geoefend, hetzij op school hetzij via een vorm van huiswerkbegeleiding en deze oude opgaven weer tegenkomen in hun eigen eindtoets, hier geen voordeel van hebben ten opzichte van de leerlingen bij wie de eindtoets eerder is afgenomen omdat dat toch ook een effect zou kunnen hebben op de moeilijkheidsgraad. Deze leden vragen de minister tevens hoe dit geborgd wordt in het geval de eindtoets (nog) niet digitaal wordt afgenomen. Tot slot vragen deze leden de minister nader toe te lichten of de ervaringen met de examens Nederlands als tweede taal op dit punt nog aanleiding geven tot een nadere afweging.

Naar aanleiding van deze vragen wil ik graag uitleggen welke procedure voor de papieren en digitale eindtoetsen wordt gebruikt om geheimhouding van de opgaven te verzekeren. Allereerst zijn er de papieren toetsen. De ankeropgaven worden 'gezaaid' in alle papieren toetsen. Dit houdt in dat de ankeropgaven in deelsetjes worden verdeeld over een groot aantal opgavenboekjes. Het is in de toetsboekjes niet zichtbaar welke opgaven ankeropgaven zijn en welke niet. Daarom is voor derden niet inzichtelijk welke opgaven eventueel in een volgende toets kunnen terugkeren. Leerlingen die een boekje met hierin ankeropgaven

maken zullen in de meeste gevallen twee of drie vragen uit het anker tegenkomen. Kijkende naar het totaal aantal opgaven van een eindtoets is het aantal vragen uit het anker dus erg beperkt.

1454840

Hoewel er opgaven zullen worden herhaald, is de kans uiterst klein dat scholen meerdere jaren achter elkaar dezelfde opgaven ontvangen. Mocht een school toch dezelfde opgaven ontvangen zal het om een klein deel van de opgaven gaan. Ten tweede zijn er digitale eindtoetsen. Hierbij is in de nota van toelichting op het Toetsbesluit PO (Stb. 2014, 209) destijds aangegeven dat geheimhouding van de opgaven hier niet in het geding komt wanneer opgaven worden herhaald. De digitale toetsen bevatten omvangrijke, adaptieve itembanken. Door deze procedure is het van te voren onduidelijk welke toetsopgaven een leerling zal ontvangen.

De maatregelen voorkomen dus niet dat dezelfde ankeropgaven in de daaropvolgende jaren worden gebruikt. Hierboven heb ik uiteen gezet waarom dit geen probleem is. Ik heb ook uitgelegd hoe door de procedure van zaaien in papieren toetsen en itembanken van digitale toetsen het praktisch onmogelijk is dat één school meerdere jaren achtereen dezelfde ankeropgaven maakt.

Verder is het niet nodig om te voorkomen dat leerlingen met de oude opgaven oefenen, het is namelijk uitgesloten dat een leerling oefent met opgaven uit het anker. Afnames waarbij ankeropgaven aan leerlingen worden voorgelegd, ofwel in een proeftoets, ofwel tijdens de daadwerkelijke afname, worden nauwgezet bijgehouden. In het geval van een papieren toets worden alle toetsboekjes na afloop weer ingenomen. Bij een digitale afname krijgt de leerling geen uitdraai van de gemaakte opgaven. Verspreiding van de toetsopgaven is daarmee niet mogelijk.

Bij de examens Nederlands als tweede taal was iets anders aan de hand. Van dit examen circuleerden beschrijvingen van taken van de schrijfexamens op sociale media. Het is echter niet waarschijnlijk dat een leerling de opgaven van de eindtoets, waarvan een voor de leerling onbekend en zeer klein deel bestaat uit ankeropgaven, zal kunnen onthouden en vervolgens verspreiden.

De leden van de SGP-fractie vragen een reactie op het risico dat de meerjarige aanwezigheid van ankeropgaven de betrouwbaarheid van toetsen aantast. Weliswaar zal de bekendheid met de inhoud van de ankeropgaven voor leerlingen en buitenstaanders beperkt zijn, maar dat geldt naar de mening van deze leden niet voor de leraren. Leraren zullen na verloop van jaren in beeld hebben wat de inhoud van ankeropgaven is en aan de hand van hun kennis en ervaring hun leerlingen, al dan niet bewust, voorbereiden op de toets. In hoeverre kan de status van het anker op termijn worden ondergraven, zo vragen zij.

De status van het anker zal op termijn niet worden ondergraven. De redenen hiervoor zijn tweeledig. De meeste leerlingen (en daarmee ook leraren) zullen slechts twee of drie opgaven van de ankeropgaven in hun eindtoets maken. Wanneer leerlingen een digitale adaptieve eindtoets maken zal het ook tussen de leerlingen in één klas verschillen welke ankeropgaven zij te zien krijgen. Deze vragen zijn niet te onderscheiden van de andere vragen in de eindtoets. Ook wanneer een leerkracht een eindtoets volledig doorloopt zal deze in de meeste gevallen dus maar twee of drie opgaven uit het anker bevatten, zonder dat de leerkracht weet welke dit zijn. Als laatste geldt ook dat het anker jaarlijks wordt verversd. Na een aantal jaar is het gehele anker vervangen. Zelfs indien een leerkracht er in slaagt enkele vragen uit het anker te achterhalen zal deze informatie snel achterhaald zijn. Zie verder ook het antwoord op de vorige vraag.

De leden van de SGP-fractie constateren dat het voornemen is de ankeropgaven per 1 augustus te verplichten, terwijl deze opgaven al met het oog op het schooljaar 2018-2019 als verplicht zijn gepresenteerd.¹⁵ Deze leden vragen waarom zonder rechtsgrond wezenlijke veranderingen in de toets zijn aangebracht die de vrijheid van de aanbieders van alternatieve toetsen beperken. De toelichting vermeldt weliswaar dat de huidige aanbieders de maatregel steunen, maar dat doet niets af aan het feit dat deze maatregel door de toetsaanbieders publiekelijk als een beperking van de ruimte voor eigen keuzes is benoemd. Zij constateren dat de gang van zaken veel trekken vertoont van het verschijnsel fantoomwetgeving dat door de commissie Dijsselbloem gewraakt is en zij vragen de minister kritisch te reflecteren op dit proces en de lessen die eruit getrokken dienen te worden.

Betrouwbare en valide uitkomsten van de eindtoetsen vind ik van groot belang voor leerlingen en scholen. Ongelijkheid op dit punt mag niet langer voortbestaan dan nodig. Niet alleen ikzelf, maar ook de inspectie van het onderwijs, de PO-raad en de toetsaanbieders hechten veel belang aan deze thema's. Daarom is in goed onderling overleg besloten om, vooruitlopend op de in het onderhavige ontwerpbesluit opgenomen aanpassingen die voortkomen uit de tussenevaluatie Wet Eindtoetsing PO, reeds afspraken met de eindtoetsaanbieders te maken om al in 2018 de set van ankeropgaven op te nemen in de eindtoets. Over deze maatregelen heb ik u in januari 2017 geïnformeerd.¹⁶

Ik deel de analyse niet dat deze afspraken overeenkomsten hebben met het verschijnsel van fantoomwetgeving dat door de commissie-Dijsselbloem is genoemd. Fantoomwetgeving ging over veranderingen die niet wettelijk waren voorgeschreven, maar door scholen wel zo werden ervaren. In de situatie rondom eindtoetsing is de afspraak in overleg met toetsaanbieders gemaakt om een set van ankeropgaven in te voeren. Deze zijn zich bewust van het feit dat de wettelijke verplichting nog niet geldt en benadrukken zelf ook het belang van de vergelijkbaarheid van de resultaten van eindtoetsen. Betrouwbare en valide uitkomsten van de eindtoetsen vinden zij net als ik van groot belang voor leerlingen en scholen. Ongelijkheid op dit punt mag niet langer voortbestaan dan nodig. Zij zijn dan ook op de hoogte van deze afspraak en van de op dit moment voorliggende wijzigingen van het Toetsbesluit PO.

¹⁵ Zie bijvoorbeeld Kamerstuk 34 775 VIII, nr. 127.

¹⁶ Kamerstukken II 2016/17, 31289, nr. 343.