

Format bij Beoordelingskader Eindtoetsen

IDnummer	21.011
Naam Toets	AMN Eindtoets 2022
Aanvrager	AMN
Beoordelaars	Gert Gelderblom Gea Spaans Cees Glas
Datum beoordeling	20210831
Datum toevoeging Expertgroep	20210908

Inhoudsopgave

Inhoudsopgave.....	2
Algemene indruk van de Eindtoets	4
3.1 Onderwijskundige Inhoud	5
3.1.1 Validiteit toets.....	5
3.1.2 Handleiding.....	5
3.1.3 Gezamenlijk Anker	6
3.1.4 Taal.....	7
3.1.4.1 Lezen	7
Zakelijke teksten	7
Fictionele, narratieve en literaire teksten	8
3.1.4.2 Taalverzorging.....	8
3.1.5.1 Verdeling toetsopgaven over domeinen	10
3.1.5.2 Verdeling toetsopgaven over onderdelen	10
3.1.5.3 Opgaven met en zonder context	11
3.1.5.4 Gebruik kladpapier.....	11
3.1.5.5 Gebruik rekenmachine	11
3.1.5.6 Toetsopgaven op niveau 1S en 1F (S/F).....	12
3.2 Organisatorische aspecten.....	13
3.2.1 Leerlingrapport	13
3.2.3 Afname eindtoets.....	15
3.3 Psychometrische aspecten	16
3.3.1 Steekproef pretest.....	16
3.3.2 Toets-samenstelling n.a.v. pretest.....	17
3.3.3 Kwaliteit operationele eindtoets.....	18
3.3.4 Normering	19
3.4 Beveiligingsaspecten	20
3.4.1 Risicoanalyse en beveiligingsmaatregelen	20
3.4.2 Incidentenprotocol.....	20

3.5 Verzamelstaat

Algemene indruk van de Eindtoets

Eindtoets AMN 2022

De AMN Eindtoets is een digitale adaptieve toets en toetst de verplichte onderdelen. De AMN Eindtoets geeft een niveau-advies voor het vervolgonderwijs en geeft een score op de referentieniveaus van Meijerink. De toets is digitaal en de verwerking van de resultaten is gedigitaliseerd.

Het Referentiekader Taal en Rekenen is uitgangspunt voor de ontwikkeling van de Eindtoets. Alleen de vereiste domeinen zijn opgenomen in de toets.

Voor Taal: Lezen (niveaus onder 1F, 1F en 2F) en Taalverzorging (niveaus 1F, 2F en 3F) zonder grammatica.

Voor Rekenen: Getallen, Verhoudingen, Meten & Meetkunde en Verbanden (niveaus onder 1F, 1F, 1S en 2F).

Opgaven in de toets worden aangepast op de ingeschatte vaardigheid van een leerling. Er zijn items op verschillende referentieniveaus opgenomen.

Er wordt gewerkt met een itembank. Ieder jaar wordt minimaal 20% van de items ververs.

In de verantwoording zijn bij de inhoudelijke beschrijving AMN Eindtoets de toetsmatrijzen opgenomen. In de bijlage is bij ieder item de koppeling weergegeven tussen code, item, goede antwoord, niveau, soort tekst voor Lezen en taak uit referentiekader. Bij Taalverzorging biedt dit inzicht in de verschillende niveaus en taak uit Referentiekader. Voor Rekenen is rekening gehouden met de volgende aspecten: verschillende niveaus, de representatieve afspiegeling van domeinen, items per domein, subdomeinen, de verhouding context en contextloze items en taak uit het Referentiekader.

Al eerder is via een pre-test inhoudelijk gereageerd op items voor de Eindtoets 2022.

Voor taal zijn de leesteksten passend bij de leeftijd van de kinderen en de onderwerpen actueel en/of passend bij hun belevingswereld. De vormgeving van de teksten blijft, net als in de vorige toetsen, sober qua opmaak. Ook is ervoor gekozen geen illustraties op te nemen. De keuze wordt mede onderbouwd vanuit het argument dat op deze manier rekening wordt gehouden met kinderen met speciale onderwijsbehoeften die de toets ook maken.

Zowel op leerlingniveau als op groepsniveau kan een rapportage worden gegenereerd. Het leerlingrapport geeft inzicht in het toetsadvies en hoe de leerling scoort op de referentieniveaus. De rapportage kan nog meer aan kracht winnen als ook de detailscores zichtbaar worden voor taal en rekenen. Dus bijvoorbeeld voor rekenen de scores op getallen, verhoudingen, meten & meetkunde en verbanden. Deze laatste opmerking is ook al eerder gemaakt. Hetzelfde geldt ook voor de groeps- en managementrapportage.

Voor detailopmerkingen bij een aantal items voor taal zie de bijlage.

3.1 Onderwijskundige Inhoud

3.1.1 Validiteit toets

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
V1	Heeft de toetsaanbieder op enigerlei wijze inzichtelijk gemaakt hoe de toetsopgaven zijn gekoppeld aan de domeinen en de onderdelen (hiervoor kan gebruik gemaakt worden van Bijlage 1 (taal) en Bijlage 2 (rekenen). Een andere vorm (bijvoorbeeld d.m.v. toetsmatrijzen) is ook mogelijk)?	ja	Via toetsmatrijzen wordt inzicht gegeven hoe de items gekoppeld zijn aan de domeinen van het Referentiekader voor Taal en Rekenen. Ook bieden de toetsmatrijzen inzicht in de verdeling van de items over de domeinen en inhoudsonderdelen en de niveaubepaling (onder 1F, 1F, 1S, 2F en 3F)
V2	Dragen de items in de toets bij aan de validiteit van de toets (hierbij gaat het om aspecten al relevantie, objectiviteit en efficiëntie van de items)?	ja	
Oordeel	Om als toets toegelaten te kunnen worden, dienen vraag V1 en V2 met JA te worden beantwoord.	voldoende	

3.1.2 Handleiding

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
H1	Is er een handleiding aanwezig met daarin een verantwoording van de gemaakte keuzes ten aanzien van de toetsinhoud, de afnamevorm, de cesuur, de psychometrische aspecten en de kwaliteitszorg itemconstructie van de toets?	ja	Er is een verantwoordingsdocument dat inzicht geeft in o.a. de uitgangspunten, het normeringstraject, inhoudelijke beschrijving van de AMN Eindtoets, betrouwbaarheid en validiteit.
Oordeel	Om als toets toegelaten te kunnen worden, dient vraag H1 met JA te worden beantwoord.	voldoende	

3.1.3 Gezamenlijk Anker

<i>Code</i>	<i>Vragen</i>	<i>Mogelijke antwoorden / oordelen</i>	<i>eventuele toelichting</i>
GA1	Worden de ankeritems voor de kalibratie van het gezamenlijke anker bij tenminste 1000 leerlingen in een high-stakes situatie afgenomen?	ja	
GA2	Is bij de opname van de ankeritems rekening gehouden met de daarover gemaakte afspraken?	ja	
Oordeel	Om als toets toegelaten te kunnen worden, dienen vragen GA1 en GA2 met JA te worden beantwoord.	voldoende	

3.1.4 Taal

3.1.4.1 Lezen

Zakelijke teksten

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
L1.1	Lezen de leerlingen informatieve teksten en maken ze daar opdrachten bij?	ja	Er is een evenwichtige verdeling tussen informatieve, instructieve en betogende teksten.
L1.2	Lezen de leerlingen instructieve teksten en maken ze daar opdrachten bij?	ja	
L1.3	Lezen de leerlingen betogende teksten en maken ze daar opdrachten bij?	ja	
L2	Voldoen de teksten die de leerlingen lezen aan de tekstkenmerken zoals beschreven in het Referentiekader?	ja	
L3.1	Wordt in de toets als kenmerk van de taakuitvoering techniek en woordenschat getoetst?	ja	
L3.2	Wordt in de toets als kenmerk van de taakuitvoering begrijpen getoetst?	ja	Gezien ook het hoofdoel van het lezen van teksten, hebben items gericht op begrijpen en interpreteren de nadruk.
L3.3	Wordt in de toets als kenmerk van de taakuitvoering interpreteren getoetst?	ja	
L3.4	Wordt in de toets als kenmerk van de taakuitvoering samenvatten getoetst?	ja	
L3.5	Wordt in de toets als kenmerk van de taakuitvoering opzoeken getoetst?	ja	Het onderdeel wordt apart en niet adaptief aangeboden omdat het op deze manier de vaardigheid van de leerling inschat waarop de leestekst wordt geselecteerd. Deze items worden ieder jaar 100% ververst.
L3.6	Worden in de toets de verschillende kenmerken van de taakuitvoering evenwichtig getoetst binnen het subdomein zakelijke teksten?	ja	
Oordeel	Om voor dit subdomein een voldoende te kunnen krijgen, gelden de volgende beslisseregels: <ol style="list-style-type: none"> 1. Taken: Van L1.1 tot en met L1.3 dienen tenminste twee van de drie vragen met JA te worden beantwoord (taak minimaal één keer aanwezig). 2. Tekstkenmerken: vraag L2 dient met JA te worden beantwoord. 3. Kenmerken van de taakuitvoering: alle vragen (L3.1 tot en met L3.6) dienen met JA te worden beantwoord. 	Voldoende	De onderwerpen zijn goed gekozen qua inhoud, actualiteit en belevingswereld van de kinderen. Er worden geen illustraties gebruikt en de lay-out is sober.

Fictionele, narratieve en literaire teksten

L4	Lezen de leerlingen fictionele, narratieve en literaire teksten en maken ze daar opdrachten bij?	ja	Er wordt geen gebruik gemaakt van gedichten. De fictionele teksten zijn aantrekkelijk om te lezen omdat er vaak een spannend moment wordt beschreven en dat houdt de aandacht vast.
L5	Voldoen de teksten die de leerlingen lezen aan de tekstkenmerken zoals beschreven in het Referentiekader?	ja	
L6.1	Wordt in de toets als kenmerk van de taakuitvoering begripen getoetst?	ja	
L6.2	Wordt in de toets als kenmerk van de taakuitvoering interpreteren getoetst?	ja	
L6.3	Worden in de toets de verschillende kenmerken van de taakuitvoering evenwichtig getoetst binnen het subdomein fictionele, narratieve en literaire teksten?	ja	
Oordeel	Om voor dit subdomein een voldoende te kunnen krijgen, gelden de volgende beslisregels: <ol style="list-style-type: none"> 1. <u>Taken</u>: L4 met JA te worden beantwoord (eveneens taak minimaal één keer aanwezig). 2. <u>Tekstkenmerken</u>: vraag L5 dient met JA te worden beantwoord. 3. <u>Kenmerken van de taakuitvoering</u>: alle vragen (L6.1 tot en met L6.3) dienen met JA te worden beantwoord. 	voldoende	

3.1.4.2 Taalverzorging

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
T1.1	Worden de leerlingen getoetst op verschillende categorieën van spellingsproblemen en - regels?	ja	
T1.2	Worden de leerlingen getoetst op regels voor lettergreepgrenzen?	ja	
T1.3	Worden de leerlingen getoetst op regels voor woordgrenzen?	ja	Alleen 3F conform Referentiekader
T1.4	Worden de leerlingen getoetst op morfologische spelling?	ja	
T1.5	Worden de leerlingen getoetst op regels voor werkwoordspelling?	ja	
T1.6	Worden de leerlingen getoetst op overige regels?	ja	Alleen 3F
T1.7	Worden de leerlingen getoetst op leestekens?	ja	
T1.8	Is er een evenwichtige verdeling tussen opgaven rond niet-werkwoordspelling, werkwoordspelling en interpunctie (dit zal moeten blijken uit de inhoudelijke verantwoording)?	ja	
Oordeel	Om voor dit subdomein een voldoende te kunnen krijgen, dienen alle vragen (T1.1 tot en met T1.8) met JA te worden beantwoord.	voldoende	In de Praktische handleiding wordt aangegeven dat er met twee vraagvormen wordt gewerkt: de vraagvorm waarin het vetgedrukte woord verkeerd is gespeld en de vraagvorm waarin het woord

			<p>correct is gespeld. N.a.v. eerdere beoordelingen is de opmerking gemaakt dat bij fout geschreven vetgedrukte woorden er een kans is dat dit beeld blijft hangen bij leeszwakke kinderen. De toetsaanbieder had toen aangegeven dat hier in de toekomst rekening mee zou worden gehouden. Dat is, ook gezien het aantal items hierover, nog niet terug te zien.</p>
--	--	--	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

3.1.5 Rekenen

3.1.5.1 Verdeling toetsopgaven over domeinen

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
D.G	Ligt het percentage toetsopgaven uit het domein Getallen tussen de 30% en 40%?	ja	
D.Vh	Ligt het percentage toetsopgaven uit het domein Verhoudingen tussen de 20% en 30%?	ja	
D.M/mk	Ligt het percentage toetsopgaven uit het domein uit het domein Meten & meetkunde tussen de 20% en 30%?	ja	
D.Vb	Ligt het percentage toetsopgaven uit het domein Verbanden tussen de 15% en 20%?	ja	
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen alle vragen met JA te worden beantwoord.	voldoende	

3.1.5.2 Verdeling toetsopgaven over onderdelen

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
O.D	Is iedere opgave onder één onderdeel geplaatst?	ja	
O.B	Is het percentage toetsopgaven uit onderdeel B ten minste 20 procent?	ja	
O.C	Is het percentage toetsopgaven uit onderdeel C ten minste 30 procent?	ja	
O.BC	Is het percentage toetsopgaven uit onderdeel C hoger dan het percentage toetsopgaven uit onderdeel B?	ja	
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen alle vragen met JA te worden beantwoord.	voldoende	

3.1.5.3 Opgaven met en zonder context

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
C1	Bevat de toets voor het domein Getallen ten minste 30 procent contextopgaven?	ja	
C2	Bevat de toets voor het domein Getallen ten minste 20 procent contextloze opgaven?	ja	
C3	Bevat de toets voor het domein Verhoudingen ten minste 30 procent contextopgaven?	ja	
C4	Bevat de toets voor het domein Verhoudingen ten minste 20 procent contextloze opgaven?	ja	
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen alle vragen met JA te worden beantwoord.	voldoende	

3.1.5.4 Gebruik kladpapier

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
K1	Mogen de leerlingen bij ten minste 80 procent van de toetsopgaven kladpapier gebruiken?	ja	De leerlingen mogen gebruik maken van kladpapier.
K2	Wordt in de handleiding voor leraren aangegeven dat zij leerlingen nadrukkelijk wijzen op het toegestane gebruik van kladpapier?	ja	
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen beide vragen met JA te worden beantwoord.	voldoende	

3.1.5.5 Gebruik rekenmachine

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
R1	Mogen de leerlingen bij 0% tot maximaal 20% van de toetsopgaven de rekenmachine gebruiken?	ja	Het gebruik van de rekenmachine is niet toegestaan.
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dient deze vraag met JA te worden beantwoord.	voldoende	

3.1.5.6 Toetsopgaven op niveau 1S en 1F (S/F)

<i>Code</i>	<i>Vragen</i>	<i>Mogelijke antwoorden / oordelen</i>	<i>eventuele toelichting</i>
S/F1	Zijn de opgaven van de toets breed samengesteld qua moeilijkheid?	ja	Er zijn opgaven opgenomen van de niveaus <1F, 1F, 1S en 2F.
S/F2	Bevat de toets ten minste de inhoud van 1F en 1S?	ja	
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen beide vragen met JA te worden beantwoord.	voldoende	

3.2 Organisatorische aspecten

3.2.1 Leerlingrapport

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
LR1	Stelt de toetsaanbieder voor elke individuele leerling een leerlingrapport op?	ja	
LR2	Bevat het leerlingrapport het resultaat?	ja	De totaalscore van de leerling staat opgenomen en is gekoppeld aan een schooltype. Ook is zichtbaar gemaakt waar de leerling staat met de scores op de referentieniveaus t.o.v. andere leerlingen die op het betreffende niveau scores als de leerling.
LR3	Geeft het leerlingrapport een advies omtrent het te volgen vervolgonderwijs conform de afspraken met OCW?	ja	Door de totaalscore te koppelen aan schooltypen wordt het advies zichtbaar. De positie van de leerling op de referentieniveaus wordt zichtbaar gemaakt door een pinnetje (in de vorm van een popje). Het vraagt van leraren naar ouders toe wel toelichting hoe de positie op de referentieniveaus inhoudelijk kan worden geïnterpreteerd en waarom het van meerwaarde is om dit zichtbaar te maken.
LR4.1	Geeft het leerlingrapport aan welk eindniveau de leerling heeft behaald (1F en 2F) voor taalverzorging?	ja	Het zou van meerwaarde zijn, zowel voor rekenen als taal, als de detailscores zichtbaar zouden zijn op de domeinen. Dit geeft ook handvatten waar al dan niet extra aandacht aan besteed zou kunnen/moeten worden in het vervolgonderwijs.
LR4.2	Geeft het leerlingrapport aan welk eindniveau de leerling heeft behaald (1F en 2F) voor lezen?	ja	
LR4.3	Geeft het leerlingrapport aan welk eindniveau de leerling heeft behaald (1F en 1S) voor rekenen?	ja	
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen alle vragen met JA te worden beantwoord.	voldoende	Naast een leerlingrapport kan een Excel-format worden ingezien waarin de resultaten van groepen leerlingen gezamenlijk worden weergegeven. Het is ook mogelijk een managementrapportage in te zien die tabellen en grafieken toont waarin de scores van de school op de domeinen en de toetsadviezen zichtbaar worden gemaakt.

3.2.2 Leerlingen

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
LL1	Kunnen alle leerlingen, ook met specifieke ondersteuningsbehoeften (de leerlingen die onder de uitzonderingsgronden vallen daargelaten), deelnemen aan de toets?	ja	Er is geen tijdslimiet voor het maken van de toets. De drie onderdelen (Lezen, Taalverzorging en Rekenen) kunnen na elkaar of op verschillende momenten worden afgenomen.
A	Wordt op enigerlei wijze ⁷ voorzien in een aangepaste afname voor leerlingen met dyslexie?	ja	Aangepaste afname door: aparte toets met extra auditieve ondersteuning, instelling beeldschermen, geschikte lettertypes, geschikt kleurgebruik en kleurcontrasten. Voor leerlingen met dyscalculie wordt aangegeven dat er aanpassingen kunnen worden gedaan t.a.v. schermresolutie, inzetten van zo groot mogelijke beeldschermen en inlassen van langere pauzes.
B	Wordt op enigerlei wijze ⁷ voorzien in een aangepaste afname voor leerlingen met een visuele beperking (slechtzienden ¹)?	ja	Schermsresolutie kan worden ingesteld.
C	Wordt op enigerlei wijze ⁷ voorzien in een aangepaste afname voor leerlingen met een auditieve beperking?	nee	Dit is niet nodig voor dit type toetsen.
D	Wordt op enigerlei wijze ⁷ voorzien in een aangepaste afname voor leerlingen met cognitief zwakke leerling (vermoedelijke uitstroom naar praktijkonderwijs of VSO)?	nee	Kinderen maken een adaptieve toets dus wordt aangepast aan hun niveau.
E	Wordt in de handleiding beschreven dat en hoe de school kan komen tot een adequate aanpassing in overleg met de aanbieder als het aanbod van toetsen en de aanpassingen bij de afname bij een beperking (of bijvoorbeeld meervoudige beperkingen) niet leiden tot een adequate aanpassing? ⁷	ja	In de Praktische handleiding wordt aangegeven dat wordt aangeraden om voor die leerlingen gebruik te maken van de Centrale Eindtoets.
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dient tenminste vraag LL1 met JA te worden beantwoord. Vragen A tot en met E vormen een checklist voor de onderwijsinhoudelijke experts, zodat zij LL1 kunnen beantwoorden.	voldoende	

¹ Blinde leerlingen kunnen indien nodig gebruik maken van de braille versie van de centrale eindtoets.

3.2.3 Afname eindtoets

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
A1	Is het eerste afnamemoment voorzien in het tijdvak 15 april – 15 mei?	ja	Tussen 15 april 2022 en 29 april 2022.
A2	Is er een mogelijkheid voor een tweede afnamemoment (voor leerlingen die de eerste afname zijn verhinderd)?	ja	Herkansingsperiode loopt tot juni 2022. In ieder geval tussen 16 april en 1 juni 2022.
A3.1	Staat in het TR/de handleiding vermeld hoe de directeur de leerlingen voor de toets moet aanmelden?	ja	Via het Leerling administratiesysteem
A3.2	Staat in het TR/de handleiding vermeld welke hulpmiddelen leerlingen mogen gebruiken?	ja	Alleen potlood/pen en kladpapier
A3.3	Worden in het TR/de handleiding de voorwaarden aan de (wijze van) geheimhouding van de toetsopgaven door de schooldirecteur beschreven?	ja	
A3.4	Wordt in het TR/de handleiding de geheimhouding van de toetsopgaven door de toetsaanbieder geregeld?	ja	In het hoofdstuk Geheimhouding van de Praktische handleiding en in het Toetsreglement staat aangegeven hoe de school/directeur met de digitale toets en toetsgegevens moet omgaan.
A3.5	Worden in het TR/de handleiding de voorwaarden aan de (wijze van) openbaarmaking van de toetsopgaven door de toetsaanbieder beschreven?	ja	Er wordt aangegeven dat de school verantwoordelijk is voor het beschermen van inhoudelijke informatie over de toets en toetsopgaven. Toetsopgaven mogen niet openbaar worden gemaakt.
A3.6	Worden in het TR/de handleiding de voorwaarden aan de (wijze van) waarop toezicht op de leerlingen wordt gehouden beschreven?	ja	De rol van de toetsleider en de taken die daarbij horen wordt duidelijk beschreven/
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen alle vragen met JA te worden beantwoord.	voldoende	Voor de ouders is er een informatiebrief beschikbaar over de inhoud en het doel van de eindtoets.

3.3 Psychometrische aspecten

3.3.1 Steekproef pretest

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
N1.1	Is de steekproef groot genoeg en representatief voor de relevante populatie?	ja	<p>De steekproef is, volgens de criteria die daarvoor gelden groot genoeg om zowel het 1PLM als het 2PLM te schatten.</p> <p>Bij de informatie paragraaf 3.3 van het document AMN Eindtoets 2022 - verantwoording over de samenstelling van de steekproef gegeven zijn echter wel kanttekeningen te plaatsen. De steekproef blijkt weliswaar niet representatief voor de landelijke populatie, maar omdat de uiteindelijke normering plaatsvindt via het gezamenlijke anker, is representativiteit t.o.v. de landelijke populatie niet essentieel. Om de eindtoets te positioneren is het overigens wel te waarderen dat er informatie over de landelijke dekking gegeven wordt.</p> <p>De functie van de pre-test is om vast te stellen of de items goed functioneren, om de itemparameters voor de itemselectie in de CAT te schatten en om een voorlopige normering vast te stellen. Een relevante eis aan de steekproef is daarom dat die steekproef adequaat is voor de doelgroep, c.q. de groep leerlingen die de eindtoets maakt. Daarom is het in het vervolg aan te bevelen om in Tabel 4 en Tabel 6 ook de gegevens van een of meer operationele afnamen van de eindtoets op te nemen. Overigens, in het document Evaluatie: vergelijking pre-test en officiële afname de benodigde informatie voor handen in Tabel 17.</p> <p>Een andere opmerking heeft betrekking op de statistische toetsingen zoals die in paragraaf 3.3 worden uitgevoerd. Door de grote steekproeven hebben de statistische toetsen een enorme power, en worden alle toetsingen automatisch significant. Zo'n benadering heeft dan ook weinig zin. Een effectgrootte zoals Cramers V, of Phi (Phi is de</p>

			wortel van de ratio van de chi-kwadraat gedeeld door de steekproefgrootte) is meer op zijn plaats.
N1.2	Is het dataverzamelingsdesign van de pretest adequaat?	ja	De steekproef voor de pre-test is opgebouwd uit drie groepen die ofwel deelnamen aan de operationele toets van 2021, ofwel aan de proeftoets 2022, ofwel aan beiden. De integriteit van de data van de proeftoets is via de betrokken leerkrachten en via analyses goed onder controle. De verschillen tussen de drie steekproeven zijn psychometrisch adequaat onderzocht. Er zijn grondige analyses gemaakt van de mate waarin het design gelinkt is.
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen alle vragen met JA te worden beantwoord.	voldoende	

3.3.2 Toets-samenstelling n.a.v. pretest

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
N2.1	Bestaat een niet-adaptieve toets of een multi-stage toets uit 100% nieuwe opgaven (d.w.z. opgaven die niet eerder zijn opgenomen in de eindtoets, met uitzondering van de items van het gezamenlijk anker) voor zowel taal als rekenen? Is de itembank van een CAT tussen de 20% en 30% ververst?	ja	Minimaal 20% van de items in de itembank voor de CAT wordt ververst.
N2.2 (adaptieve toets)	Is er sprake van de adequate verversingsstrategie a. ... voor niet geselecteerde items? b. ... voor items met discrepanties in parameters in het kalibratiedesign en de operationele fase? c. ... voor items die uitgelekt lijken? ... voor items die over jaargangen heen anders functioneren?	ja	Voldoende beargumenteerd en gedocumenteerd. In de tabellen 62, 63, 64 wordt aangegeven welke items verwijderd zijn en wat de reden daarvoor is. De motieven voor verwijdering worden vooral gegeven in termen van psychometrische indices (moeilijkheidsgraad, informatie waarden, etc.). De selectiegraad van de items wordt niet genoemd, maar die selectiegraad zal zeker samenhangen met de itemparameters.

N2.3	Is de geschatte globale betrouwbaarheid van de eindtoets en de geschatte lokale betrouwbaarheid bij de verwachte cesuren voor referentieniveaus en toetsadviezen voldoende?	Ja	De globale en lokale betrouwbaarheden zijn correct geschat en de uitkomsten zijn uitstekend.
N2.4	Is de kwaliteit van de items in termen van hun lokale betrouwbaarheid bij de verwachte cesuren voor referentieniveaus en toetsadviezen voldoende?		n.b. Is niet te bepalen omdat hierover geen informatie wordt verstrekt. In bijlage 2 worden weliswaar psychometrische itemkenmerken getabuleerd, maar IRT itemparameters en informatiewaarden worden niet vermeld. Het verdient aanbeveling om dit in de toekomst wel te doen.
N2.5	Past het gebruikte psychometrische model voldoende, onder andere m.b.t. item-fit en differential item functioning m.b.t. beschikbare achtergrondinformatie zoals geslacht.	ja	Er worden DIF analyses uitgevoerd m.b.t. afnamevorm, dyslexie, en geslacht, en de resultaten in Bijlage 3 zien er goed uit.
N2.6	Klassieke adaptieve toets: Is de itemselectiestrategie adequaat; Multistage test: is de routing-strategie adequaat.		n.b. Er is sprake van een CAT. Het algoritme voor de itemselectie en scoring is in hoofdstuk 4 uitgebreid beschreven.
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen alle vragen met JA te worden beantwoord.	voldoende	

3.3.3 Kwaliteit operationele eindtoets

Code	Vragen	Mogelijke antwoorden / oordelen	eventuele toelichting
N3.1	Was de globale betrouwbaarheid van de eindtoets en de lokale betrouwbaarheid bij de cesuren voor referentieniveaus en toetsadviezen voldoende?	ja	De tabellen 30 en 31 laten zien dat globale en lokale betrouwbaarheden goed waren bij alle cesuurpunten.
N3.2	Was de kwaliteit van de items in termen van hun lokale betrouwbaarheid bij de cesuren voor referentieniveaus en toetsadviezen voldoende?	ja	Paragraaf 4.4 laat zien dat de lokale betrouwbaarheid van de items bij een aantal relevante theta-waarden goed was.
N3.3	Paste het gebruikte psychometrische model voldoende, onder andere m.b.t. item-fit en differential item functioning bij de beschikbare achtergrondinformatie zoals geslacht.	ja	n.b. Er is in het verslag geen informatie over DIF opgenomen. Het verdient aanbeveling om in de toekomst ook DIF m.b.t. urbanisatiegraad te onderzoeken, omdat de populaties in de grote steden, dorpen en het platte land sterk verschillen.

N3.4	Waren de schattingen van de pretest m.b.t. itemparameters, toets-eigenschappen (percentages in referentie- en toetscategorieën) en populatiekenmerken (vaardigheidsniveau) adequaat?	ja	Tabel 11 geeft de correlatie tussen de itemparameters van de pre-test en de operationele afname. Die correlatie is hoog. Ook zijn er DIF toetsen afgenomen; de resultaten staan in Tabel 13. Het aantal items met een effectgrootte boven de 0.10 is verwaarloosbaar.
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen alle vragen met JA te worden beantwoord.	voldoende	

3.3.4 Normering

<i>Code</i>	<i>Vragen</i>	<i>Mogelijke antwoorden / oordelen</i>	<i>eventuele toelichting</i>
N4	Wordt er informatie verstrekt waaruit blijkt dat de adviezen van de Expertgroep ten aanzien van de normering van de referentieniveaus, zijn opgevolgd?	ja	De normering vond plaats onder regie van de Expertgroep, en de aanwijzingen van de Expertgroep zijn adequaat opgevolgd.
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dient deze vraag met JA te worden beantwoord.	voldoende	

3.4 Beveiligingsaspecten

3.4.1 Risicoanalyse en beveiligingsmaatregelen

<i>Code</i>	<i>Vragen</i>	<i>Mogelijke antwoorden</i>	
RB1	Wordt de risicoanalyse weergegeven die is gemaakt voor alle fasen van het toetsproces in het kader van beveiligingsrisico's?	ja	Dit is uitgewerkt via BIV-classificatie veilig toetsen en een RACI-model.
RB2	Zijn specifieke en relevante beheersmaatregelen geformuleerd voor situaties met een middelmatig of hoog risico?	ja	
RB3	Zijn de beheersmaatregelen die de toetsleider moet ondernemen helder omschreven?	ja	
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dienen alle vragen met JA te worden beantwoord.	Voldoende	

3.4.2 Incidentenprotocol

<i>Code</i>	<i>Vragen</i>	<i>Mogelijke antwoorden</i>	
Ip1	Wordt het incidentenprotocol weergegeven? En is dit incidentenprotocol zodanig opgesteld dat het door alle mogelijke gebruikers te begrijpen en uitvoerbaar is?	ja	In de handleiding is een duidelijk uitgewerkt incidentenprotocol opgenomen.
Oordeel	Om dit thema met een voldoende te kunnen afsluiten, dient de vraag met JA te worden beantwoord.	Voldoende	

3.5 Verzamelstaat

Thema's	Mogelijke oordelen
Onderwijskundige inhoud	
Validiteit toets	voldoende
Handleiding	voldoende
Gezamenlijk Anker	voldoende
<i>Taal</i>	
Lezen	voldoende
Taalverzorging	voldoende
<i>Rekenen</i>	
Verdeling toetsopgaven over domeinen	voldoende
Verdeling toetsopgaven over onderdelen	voldoende
Opgaven met en zonder context	voldoende
Gebruik kladpapier	voldoende
Gebruik rekenmachine	voldoende
Niveaus S en F	voldoende
Organisatorische aspecten	
Leerlingrapport	voldoende
Leerlingen	Voldoende
Afname eindtoets	voldoende
Psychometrische aspecten	
Steekproef pretest	voldoende
Toets-samenstelling n.a.v. pretest	voldoende
Kwaliteit operationele eindtoets	voldoende
Normering	voldoende
Beveiligingsaspecten	
Risicoanalyse en Beheersmaatregelen	voldoende
Incidentenprotocol	voldoende

Zoals eerder vermeld dienen ook aan de gestelde eisen betreffende de overige taaldomeinen te worden voldaan wanneer het schooladvies wordt gebaseerd op meer dan de domeinen lezen en taalverzorging. In dat geval geldt ook onderstaand schema.

Thema's	Mogelijke oordelen
Inhoud	
<i>Taal</i>	
Schrijven	voldoende/onvoldoende
Mondelinge taalverzorging	voldoende/onvoldoende
Begrippenlijst	voldoende/onvoldoende