



Regeling van het College voor toetsen en examens van 3 februari 2025, CvTE-25.00472, houdende wijziging van de Regeling beoordelingskader doorstroomtoets PO

Het College voor toetsen en examens,

Gelet op artikel 3a, eerste lid, onderdeel g, van de Wet College voor toetsen en examens;

Gezien de goedkeuring van de Staatssecretaris van Onderwijs, Cultuur en Wetenschap, gegeven op datum 19 februari 2025, nummer 50621269,

Besluit:

ARTIKEL I

De regeling beoordelingskader doorstroomtoets PO wordt als volgt gewijzigd:

De bijlage wordt vervangen door de bijlage bij deze regeling.

ARTIKEL II INWERKINGTREDING

Deze regeling treedt in werking met ingang van de dag na de datum van uitgifte van de Staatscourant waarin de regeling wordt geplaatst.

Deze regeling zal met de toelichting in de Staatscourant worden geplaatst.

*Het College voor toetsen en examens,
J.H. van der Vegt*

BIJLAGE 1 BEOORDELINGSKADER VOOR DE DOORSTROOMTOETS

Bijlage behorende bij artikel 1 van de Regeling beoordelingskader doorstroomtoets PO.

1. Inleiding

1.1 Begrippen en definities

Aanbieder: aanbieder van een doorstroomtoets.

Absolute normen: standaarden waaraan de vaardigheidsniveaus van leerlingen gerelateerd kunnen worden. Deze standaarden worden vastgesteld op basis van inhoudelijke en/of toetsspecifieke criteria.

Afnemer: een afnemer van een toets zoals een leerkracht, intern-begeleider, directeur of observator.

Algoritme (beslisregel): systeem van regels, logische of wiskundige handelingen, die als stappenplan gelden en kunnen worden uitgevoerd om tot een bepaald resultaat te komen.

A-parameter: geeft bij een 2PL itemresponsmodel het discriminerend (onderscheidend) vermogen van een item weer. Hoe hoger het onderscheidend vermogen, hoe beter het item onderscheid maakt tussen lage en hoge vaardigheidsniveaus. In de praktijk worden de waarden relatief – d.w.z. tussen de items onderling – bekeken: hoe hoger de waarde van een bepaald item, hoe hoger het onderscheidend vermogen ten opzichte van de andere items.

A-waarde: is de proportie leerlingen die, of het percentage leerlingen dat, bij een meerkeuzevraag de desbetreffende afleider als antwoord heeft gekozen.

Beoordelaarsovereenstemming: de mate waarin verschillende beoordelaars tot identieke beoordelingen komen van (bijvoorbeeld) het werk van leerlingen. Twee veel gebruikte maten voor beoordelaarsovereenstemming zijn het percentage overeenstemming tussen twee beoordelaars en Cohen's Kappa.

Betrouwbaarheid: de mate waarin de toetsscores vrij zijn van toevallige meetfouten.

B-parameter: verwijst over het algemeen naar de parameter in itemresponsmodellen die betrekking heeft op de moeilijkheidsgraad van een item.

CAT: computergestuurde adaptieve toets op itemniveau, waarbij een item wordt gekozen die qua moeilijkheid het beste aansluit op de vaardigheid zoals ingeschat op basis van de antwoorden op de eerdere opgaven.

Cesuur: minimale prestatie (gerepresenteerd als vaardigheid of toetsscore) die net indicatief is voor het halen van een (referentie)niveau en toetsadviescategorieën.

Constructvaliditeit: eigenschap die een doorstroomtoets heeft als kan worden aangetoond dat de toets het door de constructeur beoogde kenmerk van de leerling (onderliggende trek, vaardigheid) meet. Deze wordt per kenmerk binnen een doorstroomtoets bepaald.

Dichotome scoring: wijze van scoren (punten toekennen) waarbij slechts twee waarderingen van het antwoord worden onderscheiden: bijvoorbeeld 'goed' of 'fout'. Een meerkeuzevraag wordt dichotoom gescoord (0 = niet correct, 1 = correct).

Differentieel item functioneren (DIF): is een verschijnsel dat items verschillend functioneren voor bepaalde (sub)groepen leerlingen, zelfs wanneer deze (sub)groepen leerlingen een vergelijkbaar prestatieniveau hebben. Het gaat hierbij over de vraagonzuiverheid, dat wil zeggen wanneer de verschillen in de prestaties op het item tussen groepen niet geheel verklaard kunnen worden door verschillen in de (latente) vaardigheid die aan de opgave ten grondslag ligt.

Discriminerend vermogen: mate waarin op grond van de scores op een item hoog- en laagscorende leerlingen (op de hele toets) onderscheiden kunnen worden.

Domein: verzameling van onderling samenhangende kennis en vaardigheden waarover getoetst wordt. Zo bestaat het terrein Rekenen uit de vier domeinen Getallen, Verhoudingen, Meten en Meetkunde en Verbanden.

Effectgrootte: statistische maat om de grootte van een effect – bijvoorbeeld het verschil in de gemiddelde rekenprestaties tussen jongens en meisjes – weer te geven. Er wordt een onderscheid gemaakt tussen niet gestandaardiseerde (ruwe) effectgroottes en gestandaardiseerde effectgroottes. De bekendste gestandaardiseerde effectmaat is Cohen's d. Deze maat geeft het verschil tussen twee gemiddelden in het aantal standaarddeviaties. Ook de correlatie kan worden gezien als een gestandaardiseerde maat voor de effectgrootte.

Gezamenlijk anker: verzameling van items die in alle doorstroomtoetsen wordt opgenomen om verschillen in de moeilijkheidsgraad tussen de betreffende doorstroomtoetsen vast te kunnen stellen.

Inhoudsgebied: vaardigheid die in het onderwijsaanbod in het primair onderwijs is opgenomen. Zoals vaardigheden waarover referentieniveaus kunnen worden gerapporteerd (rekenen, lezen, taalverzorging) of andere vaardigheden waarop op hoofdniveau wordt gerapporteerd.

Inhoudsvaliditeit: eigenschap dat de operationalisering van de inhoud van de doorstroomtoets een representatieve weergave is van het te meten referentieniveau.

Itemkalibratie: schatten (berekenen) van de psychometrische eigenschappen (parameters) van een verzameling items met het doel om deze items op dezelfde schaal te brengen.

Itemtype: vraagvorm, zoals meerkeuze, open, dictee.



Jaarset: de selectie van gezamenlijke ankeritems uit de ankerbank die in een bepaald afnamejaar wordt afgenomen.

Leerdoel: beoogde resultaat van onderwijsleeractiviteiten. Als het beoogde leerresultaat wordt geformuleerd in termen van uiterlijke condities en waarneembaar leerlinggedrag spreekt men van 'concrete leerdoelen' of 'gedragsdoelen'. Als het beoogde leerresultaat wordt geformuleerd in termen van cognitief gedrag of van cognitieve structuren, spreekt men van cognitieve leerdoelen.

Lokale betrouwbaarheid: mate waarin de toetsscores op een specifiek deel van de meetschaal vrij zijn van toevallige meetfouten.

Module: set van items in een multistage toets (MST), vaak van vergelijkbare moeilijkheid.

Moeilijkheidsgraad: in hoeverre leerlingen uit de doelpopulatie in staat zijn om de vraag of vragen goed te beantwoorden. De moeilijkheid wordt meestal uitgedrukt met een p-waarde.

Multistage toets (MST): adaptieve doorstroomtoets op moduleniveau. Een MST lijkt op een CAT, maar in plaats dat items een voor een worden geselecteerd, worden er sets van items afgenomen in modules. Nadat een leerling een set heeft afgerond, wordt een nieuwe inschatting van de vaardigheid gemaakt en op basis daarvan wordt een nieuwe set gekozen. Een MST kan een digitale toets zijn, of één op papier.

Normeren: toekennen van een waardering aan een score. Bij de landelijke (toetsoverstijgende) normering van de doorstroomtoets betreft dit het vaststellen van de cesuren van de referentieniveaus en de toetsadviescategorieën.

Onderwijstype: specificatie van onderwijssoort op basis van de mate van extra ondersteuning. Onderscheiden worden regulier basisonderwijs (BO), speciaal onderwijs (SO) en speciaal basisonderwijs (SBO).

Plan, Do, Check, Act (PDCA)-cyclus: is een model om de continue verbetering en vernieuwing in een organisatie te sturen. Het cyclisch karakter zorgt ervoor dat de kwaliteitsverbetering continu onder de aandacht is.

Po: primair onderwijs.

Polytome scoring: scoren van een antwoord waarbij er meer dan twee onderscheidingen zijn in de waardering van het antwoord. Aan het antwoord op een vraag kunnen bijvoorbeeld 0, 1 of 2 punten worden toegekend. Bij open vragen hanteert men vaak polytome scoring.

Populatie: verzameling van vooraf gedefinieerde eenheden waarover men met behulp van statistisch onderzoek uitspraken wil doen.

Pretesten: afnemen van een aantal opgaven bij een groep leerlingen, meestal met het doel de eigenschappen van de doorstroomtoets te onderzoeken. Deze groep leerlingen is zoveel mogelijk vergelijkbaar met de populatie waarvoor de doorstroomtoets oorspronkelijk is ontwikkeld.

P-waarde: is een getal tussen 0 en 1 waarmee de moeilijkheidsgraad van een opgave kan worden weergegeven. De p-waarde wordt berekend door de gemiddelde score op een opgave te delen door de maximaal haalbare score op die opgave. De p-waarde voor een meerkeuzevraag is gelijk aan het deel van leerlingen die het juiste antwoord heeft gekozen.

Referentiekader: beschrijving van wat leerlingen op verschillende momenten in hun schoolloopbaan op het gebied van Nederlandse taal en Rekenen moeten kennen en kunnen (zie ook: Besluit referentieniveaus Nederlandse taal en Rekenen).

Referentieniveaus: beschrijvingen van beheersingsdoelen voor Nederlandse taal en Rekenen die aangeven wat leerlingen moeten kennen en kunnen op verschillende momenten tijdens een schoolloopbaan. De referentieniveaus zijn wettelijk verankerd in de Wet referentieniveaus Nederlandse taal en Rekenen.

Rit-waarde: correlatie tussen het item en de totaalscore op alle items. De item-testcorrelatie is een maat voor het discriminerend vermogen van items. De item-testcorrelatie valt hoger uit dan de item-restcorrelatie omdat hier de itemscore in de totaalscore van een toets wordt meegeteld. Daarom wordt aangeraden om de item-restcorrelatie te gebruiken.

Rir-waarde: correlatie tussen het item en de totaalscore op alle andere items. De item-restcorrelatie is een maat voor het discriminerend vermogen van een item.

Schaal: reeks getallen die volgens een bepaald voorschrift gekoppeld worden aan waarnemingen.

Steekproef: verzameling leerlingen met specifieke kenmerken die op aselecte wijze uit de te onderzoeken populatie zijn getrokken.

Stopcriterium: aanpassen van de toetslengte. Een doorstroomtoets kan voor wat betreft het aantal toetsopgaven adaptief zijn in toetslengte. Het aanpassen van de toetslengte wordt in het geval van een computergestuurde adaptieve toets op itemniveau (CAT) geautomatiseerd afgehandeld. Na iedere opgave, wordt dan gekeken of de toets al gestopt kan worden. Stoppen gebeurt, zodra de vaardigheid van een leerling met de vereiste precisie te schatten is.

Subdomein: verzameling van onderling samenhangende kennis en vaardigheden waarover getoetst wordt. Zo bestaat het domein lezen uit de subdomeinen zakelijke en fictionele teksten.

Taaktype/taakuitvoering: classificering of combinatie van classificaties van de onderverdeling van een inhoudsgebied, zoals begrijpen, interpreteren, context loos, morfologische spelling.

Terrein: centraal onderdeel dat in de doorstroomtoets aan bod dient te komen. Het referentiekader onderscheidt twee centrale terreinen: het terrein Nederlandse taal en het terrein Rekenen. Beide terreinen worden uitgewerkt in onderliggende (sub)domeinen.



Toetsadvies: op basis van de behaalde resultaten op de doorstroomtoets berekende best passende niveau voor het volgen van aansluitend voorgezet onderwijs voor de leerling.

(Toets)item: is synoniem voor (toets)opgave.

(Toets)opgave: is een onderdeel van een toets en bevat altijd een stimulus (een middel of prikkel om een bepaald van een kandidaat uit te lokken) waarop een respons (het uitgelokte gedrag) mogelijk moet zijn.

Toetsreglement: document waarin de wijze van afnemen wordt geregeld.

Toetsterm: concrete, eenduidige en meetbare operationalisering van het betreffende leerdoel zoals beschreven in het bijpassende referentieniveau Nederlandse taal en Rekenen.

Zaai-opgave: een toetsopgave die nog niet eerder is gepretest en in een doorstroomtoets wordt afgenomen om data over te verzamelen ten behoeve van parameterschattingen.

1.2 Reikwijdte

Sinds het schooljaar 2014–2015 zijn scholen in het primair onderwijs verplicht in groep 8 een eindtoets Nederlandse taal en Rekenen af te nemen. Per 1 januari 2023 is de eindtoets vervangen door de doorstroomtoets. De Wet doorstroomtoets po schrijft voor dat het College voor Toetsen en Examens (CvTE) als wettelijke taak heeft om doorstroomtoetsen te erkennen. Doorstroomtoetsen die het CvTE erkent, worden voor een periode van vier jaar toegelaten tot het primair onderwijs.

Het CvTE maakt voor het erkennen van een doorstroomtoets gebruik van een adviseur. De adviseur baseert diens advies op het beoordelingskader vastgesteld met deze regeling.

Op basis van het advies van de adviseur stelt het CvTE tevens jaarlijks vast of de erkende doorstroomtoets voor dat jaar voldoet aan de criteria van de Regeling beoordelingskader doorstroomtoetsen po.

Bij het beoordelen van extra kennisgebieden, buiten de verplichte en optionele domeinen van de terreinen Nederlandse taal en Rekenen, oordeelt het CvTE of ook dat deel van de toets inhoudelijk valide en betrouwbaar is en of het van een deugdelijke normering is voorzien. Voor de beoordeling van dit deel worden de toetsbare onderdelen van dit beoordelingskader toegepast.

Het CvTE beoordeelt ook de door de overheid aangeboden calamiteitentoets¹. Vanwege de bijzondere aard van de calamiteitentoets, bestaat de jaarlijkse vaststelling slechts uit een inhoudelijke check op actualiteiten en de kalibratie. Indien de calamiteitentoets in enig jaar moet worden ingezet, wordt een nieuwe toets voor het daarop volgende jaar ontwikkeld. Deze dient dan opnieuw erkend te worden door het CvTE.

1.3 Aan te leveren documentatie

Alvorens het CvTE een aanvraag in behandeling neemt, moet de aanvrager de navolgende gegevens en bescheiden verstrekken:

- Een door het CvTE beschikbaar gestelde model voor een leeswijzer met daarin uiteengezet in welk aangeleverd document en hoofdstuk een kwaliteitscriterium uit het beoordelingskader wordt verantwoord;
- Toets met bijbehorende ondersteunende materialen voor de gebruiker (inclusief beschikbare versie(s) voor leerlingen met ondersteuningsbehoeften);
- Itemlijst met itemgegevens (volgnummer, itemcode, (sub)domein, itemtype, sleutel en/of scoringsregels, kenmerk van de taakuitvoering, referentieniveau, ankertype (intern of gezamenlijk), p-waarde, rit/rir, alpha, beta) en indien van toepassing, aangeven of en wanneer de items bij een vroegtijdig onderwijskundige beoordeling zijn beoordeeld, en of items al eerder zijn afgenomen in operationele setting;
- Praktische handleiding voor leerkrachten;
- Wetenschappelijke en inhoudelijke verantwoording, inclusief de psychometrische aspecten van de doorstroomtoets, zoals beschreven in hoofdstuk 5 van deze regeling en zoals beschreven in de Regeling Beoordelingsnormen Doorstroomtoetsen po.
- Toetsreglement van de betreffende toets;
- Informatie over de toetsinhoud en de afnamevorm;
- Voorbeeld leerlingrapportage en eventueel andere rapportages;
- Toetsveiligheidsplan;
- Actieplan hoe om te gaan met diefstal van toetsinhoud;
- Fraudepreventieplan;
- Indien van toepassing: Cd's, dvd's en/of usbsticks en log-ins en wachtwoorden;
- Indien van toepassing: Antwoordformulieren;

¹ https://www.eerstekamer.nl/behandeling/20210907/gewijzigd_amendement_van_het_lid/document3/f=/vImnj6ccx4n2.pdf



- Indien van toepassing: ISO-certificaat inclusief een verklaring van toepasbaarheid (c.q. de scope van het certificaat);
- Indien van toepassing: een calamiteitenplan.

Let op: bovenstaande opsomming is niet uitputtend en een indicatie voor de aan te leveren documentatie. Met het invullen van de leeswijzer maakt een aanbieder helder in welke documentatie de kwaliteitscriteria precies worden verantwoord.

Een aanbieder heeft de mogelijkheid om voor verschillende wijzen van afname (bijv. digitaal en papier) een erkenning of vaststelling aan te vragen (naast de optie om verschillende wijzen van afname in één aanvraag in te dienen bij het CvTE). Indien een aanbieder hiervoor kiest, dient de aanbieder ook daadwerkelijk meerdere aanvragen in te dienen, in separate mappen op de terminal server, met separaat ingevulde leeswijzers. Allebei de aanvragen krijgen dan ook separaat een erkenning of vaststelling. Deze aanvragen dienen echter volledig onafhankelijk van elkaar te kunnen worden beoordeeld. Dat wil zeggen dat de ene afnamemodus bijvoorbeeld niet de terugvaloptie kan zijn van de andere afnamemodus. Dit zou namelijk betekenen dat als de ene afnamemodus niet wordt erkend of vastgesteld, dit ervoor zorgt dat door de afhankelijkheid de andere afnamemodus ook niet kan worden erkend of vastgesteld.

1.4 Het aanleveren van items

Om de items in een doorstroomtoets efficiënt en gestructureerd te kunnen beoordelen, zijn er eisen aan het leveren van items aan het CvTE:

In het geval van een papieren toets:

- Toetsboeken (digitaal aanleveren volstaat) inclusief ankeropgaven en eventuele zaai-opgaven.

In het geval van een lineaire digitale toets:

- Inlogmogelijkheid als leerling en als administrator (scrollen zonder antwoorden in te vullen, toets onderbreken en op later tijdstip te hervatten);
- Mogelijkheid om items per inhoudsgebied te bekijken;
- Mogelijkheid alle items te kunnen inzien inclusief ankeropgaven en eventuele zaai-opgaven.

In het geval van een MST:

- Inlogmogelijkheid als leerling en als administrator (scrollen zonder antwoorden in te vullen, toets onderbreken en op later tijdstip te hervatten);
- Mogelijkheid om items per inhoudsgebied te bekijken;
- Mogelijkheid om items per module te bekijken;
- Mogelijkheid om items in (bij voorkeur verschillende) toetspaden te bekijken;
- Mogelijkheid alle items te kunnen inzien inclusief ankeropgaven en eventuele zaai-opgaven.

In het geval van een CAT:

- Inlogmogelijkheid als leerling en als administrator (scrollen zonder antwoorden in te vullen, toets onderbreken en op later tijdstip te hervatten);
- Mogelijkheid om items per inhoudsgebied te bekijken;
- Mogelijkheid om items in (bij voorkeur verschillende) toetspaden te bekijken;
- Mogelijkheid alle items te kunnen inzien inclusief ankeropgaven en eventuele zaai-opgaven

In het geval van aparte varianten van de doorstroomtoets voor leerlingen met specifieke ondersteuningsbehoeften:

- Mogelijkheid om een aantal voorbeelditems te bekijken waarbij een duidelijk beeld wordt gegeven van de aanpassing, zoals items met gebarentaal of aanpassing van de kleuren en/of contrast.

2. Format Regeling beoordelingskader doorstroomtoetsen po

Het beoordelingsformat bevat de volgende drie inhoudelijke onderdelen, die enkele onderliggende thema's bevatten:

- H.3. Onderwijskundige aspecten
- H.4. Organisatorische aspecten
- H.5. Psychometrische aspecten

3. Onderwijskundige aspecten

3.1. Inleiding

In de volgende paragrafen worden de vakinhoudelijke kwaliteitseisen uitgewerkt voor de wettelijk

verplichte domeinen Lezen en Begrippenlijst en Taalverzorging, daaronder niet begrepen het subdomein Begrippenlijst, van het terrein Nederlandse taal en de wettelijk verplichte domeinen Getallen, Verhoudingen, Meten en Meetkunde en Verbanden van het terrein Rekenen én de optionele domeinen Schrijven en Mondelinge taalvaardigheid en het eveneens optionele subdomein Begrippenlijst van het terrein Nederlandse taal, in lijn met het Toetsbesluit PO, de Wet doorstroomtoetsen po en zoals vastgesteld in het Besluit referentieniveaus Nederlandse taal en Rekenen. Voor alle wettelijk verplichte en optionele (sub)domeinen zijn in deze regeling kwaliteitseisen geformuleerd.

Naast de wettelijk verplichte én optionele terreinen en (sub)domeinen staat het de toetsaanbieder vrij om extra kennisgebieden aan de doorstroomtoets toe te voegen, bijv. aardrijkskunde. De score die de leerling op deze extra kennisgebieden haalt, mag de toetsaanbieder toevoegen aan het leerlingrapport. Deze extra kennisgebieden tellen niet mee voor het berekende toetsadvies en tellen ook niet mee voor de berekende score op de referentieniveaus. Alle toetsopgaven van de wettelijk verplichte en optionele (sub)domeinen en terreinen, én de extra kennisgebieden worden inhoudelijk door een adviseur van het CvTE beoordeeld. Zie in dit kader de kwaliteitseisen in paragraaf 3.2. Dit beoordelingskader schrijft niet voor wat de minimale en maximale toetslengte in aantal toetsvragen van een doorstroomtoets dient te zijn. Eveneens worden er geen richtlijnen gegeven over de verhouding van het aantal toetsvragen tussen de wettelijk verplichte en optionele (sub)domeinen en terreinen, én extra kennisgebieden. Verder worden er ook geen richtlijnen gegeven voor het aantal toetsvragen per referentieniveau 1F/2F/1S. Dit om de aanbieders de mogelijkheid te geven om zich van elkaar te onderscheiden en volgens de eigen zienswijze een compleet beeld van de leerlingen te kunnen geven.

3.2. Inhoudsvaliditeit doorstroomtoets

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
V1	Zijn alle toetsopgaven (inclusief ankeropgaven en eventuele zaai-opgaven) van de gehele toets inzichtelijk aangeleverd?	ja/nee
V2	Is de toetsmatrijs een adequate representatie van het meetdoel voor in ieder geval het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én de, indien van toepassing, optionele (sub)domeinen van Nederlandse taal?	ja/nee
V3	Voldoen alle toetsopgaven – inclusief eventuele zaai-opgaven – van enerzijds het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én anderzijds de, indien van toepassing, optionele (sub)domeinen van Nederlandse taal in de toets aan de kwaliteitscriteria voor de inhoudsvaliditeit (zijnde de relevantie van de inhoud van de toets) van toetsopgaven?	ja/nee
V4	Zijn alle toetsopgaven – inclusief eventuele zaai-opgaven – van het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én de, indien van toepassing, optionele (sub)domeinen van Nederlandse taal in de toets correct geconstrueerd?	ja/nee

Toelichting V1: De aanbieder wordt gevraagd om alle toetsopgaven inzichtelijk aan te leveren. Daaronder vallen alle items van de doorstroomtoets zelf én ankeropgaven en eventuele zaai-opgaven.

Toelichting V2: De toetsmatrijs is voor wat betreft het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én de, indien van toepassing, optionele (sub)domeinen van Nederlandse taal een adequate representatie van het meetdoel. Dat betekent eveneens dat er via de toetsmatrijs wordt voldaan aan de eisen uit de Toetswijzer PO.

Er is sprake van een adequate representatie wanneer de toetstermen het meetdoel representeren. Dit blijkt uit het gegeven dat:

- de leerdoelen zijn geoperationaliseerd in observeerbare en meetbare toetstermen die ieder voor zich aansluiten bij de domeinbeschrijvingen en indeling van de referentieniveaus voor het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én de, indien van toepassing, optionele (sub)domeinen van Nederlandse taal;
- de toetsmatrijs in ieder geval bevat: (1) het aantal vragen met bijbehorende scorepunten per vraag, (2) de toetsvorm en/of het type vragen, (3) de verdeling over de (sub)domeinen, (4) de verdeling van de verschillende vraag- en teksttypes in de toets, (5) de verdeling over de referentieniveaus, (6) de toegestane hulpmiddelen, en (7) de geschatte toetsduur;
- in het geval van een CAT dat de aanbieder verantwoordt hoe het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én de,

indien van toepassing, optionele (sub)domeinen van Nederlandse taal op representatieve wijze voor alle individuele leerlingen aan bod komen.

Toelichting V3: De vijf kwaliteitscriteria relevantie, objectiviteit, efficiëntie, specificiteit en neutraliteit staan beschreven in de Checklist voor het beoordelen van de kwaliteit van observatie-categorieën en toetsopgaven. Alle toetsopgaven dienen hieraan te voldoen – inclusief eventuele zaai-opgaven – van in ieder geval het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én de, indien van toepassing, optionele (sub)domeinen van Nederlandse taal.

Indien in de doorstroomtoets ook optionele productieve vaardigheden worden getoetst, levert de aanbieder een beoordelaarsschema in, aangevuld met informatie over de beoordelaarsovereenstemming, een en ander conform de voorschriften uit de Toetswijzer PO.

Toelichting V4: De constructievoorschriften voor toetsvragen staan beschreven in de Checklist voor het beoordelen van de kwaliteit van observatie-categorieën en toetsopgaven. Alle toetsopgaven dienen hieraan te voldoen – inclusief eventuele zaai-opgaven – van in ieder geval het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én de, indien van toepassing, optionele (sub)domeinen van Nederlandse taal.

Ongewenste gevoeligheid van bijvoorbeeld opgaven, teksten en luisterfragmenten (denk aan reclame) wordt ondervangen met de kwaliteitseis Neutraliteit in de Checklist voor het beoordelen van de kwaliteit van observatie-categorieën en toetsopgaven. Het CvTE beoordeelt niet of er rechtmatig gebruik wordt gemaakt van auteursrechtelijk materiaal.

3.3. Verantwoording

Kwaliteitseis

Beslisregel: Om aan deze categorie kwaliteitseis te voldoen, dient de vraag met JA te worden beantwoord.

Code	Vraag	Mogelijke antwoorden
H1	Is er (een set van) verantwoordingsdocumenten aangeboden aan het CvTE met daarin voor wat betreft in ieder geval het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én de, indien van toepassing, optionele (sub)domeinen van Nederlandse taal, een verantwoording van de gemaakte keuzes ten aanzien van de toetsinhoud, de afnamevorm, de psychometrische aspecten en de kwaliteitszorg rondom de itemconstructie van de toets?	ja/nee

Toelichting H1: Er wordt gecontroleerd of de verantwoordingsdocumenten voor wat betreft het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen én de, indien van toepassing, optionele (sub)domeinen van Nederlandse taal, inhoudelijk voldoen aan de wettelijke kwaliteitseisen uit het Toetsbesluit PO en aan de inhoudelijke eisen van de Toetswijzer PO. Meer specifiek dient de door het CvTE te beoordelen set verantwoordingsdocumenten te bestaan uit de (separate) documenten aansluitend bij hoofdstuk 1.3.

In het kader van transparantie, dient in de (set van) verantwoordingsdocumenten terug te vinden te zijn welke terreinen en (sub)domeinen wel en niet meetellen voor het berekenen van het toetsadvies. Alle documenten dient de toetsaanbieder jaarlijks te updaten.

3.4. Terrein Nederlandse taal

In iedere doorstroomtoets worden voor wat betreft het terrein Nederlandse taal de wettelijk verplichte domeinen Lezen (§ 3.4.1) en Begrippenlijst en Taalverzorging, daaronder niet begrepen het subdomein Begrippenlijst, (§ 3.4.2) getoetst. Aanvullend hierop kunnen optioneel één of meerdere van de domeinen Schrijven (§ 3.4.3) en Mondelinge taalvaardigheid (§ 3.4.4) en het subdomein Begrippenlijst (§ 3.4.5) worden getoetst. De gestelde vakinhoudelijke kwaliteitseisen voor deze vijf onderdelen en de vereiste verdeling van de toetsopgaven over de (sub)domeinen in de toetsmatrijs en de toetssamenstelling zijn opgenomen in de Toetswijzer PO.

3.4.1. Wettelijk verplicht domein: Lezen

Beslisregel:

Om aan deze categorie kwaliteitseisen te voldoen, gelden de volgende beslisregels:

1. Taken: van L1.1 t/m L1.3 dienen tenminste twee van de drie vragen met JA te worden beantwoord en dient L4 met JA te worden beantwoord.



2. Tekstkenmerken: vraag L2 en L5 dienen beide met JA te worden beantwoord.
3. Kenmerken van de taakuitvoering: alle vragen (L3.1 t/m L3.7a en L6.1 t/m L6.4a) dienen met JA of N.V.T. te worden beantwoord.

3.4.1.1. Subdomein Zakelijke Teksten

Kwaliteitseisen

Taken:

Code	Vragen	Mogelijke antwoorden
L1.1	Lezen de leerlingen minimaal één informatieve tekst en maken ze daar opdrachten bij?	ja/nee
L1.2	Lezen de leerlingen minimaal één instructieve tekst en maken ze daar opdrachten bij?	ja/nee
L1.3	Lezen de leerlingen minimaal één betogende tekst en maken ze daar opdrachten bij?	ja/nee

Tekstkenmerken:

Code	Vragen	Mogelijke antwoorden
L2	Voldoen de teksten die de leerlingen lezen aan de tekstkenmerken inhoud (thema), vorm (taalgebruik en structuur) en bedoeling (doel), zoals gespecificeerd in Checklist 1 van deze regeling en zoals nader omschreven in het Referentiekader?	ja/nee

Kenmerken van de taakuitvoering:

Code	Vragen	Mogelijke antwoorden
L3.1	Wordt in de toets als kenmerk van de taakuitvoering techniek en woordenschat getoetst?	ja/nee
L3.2	Wordt in de toets als kenmerk van de taakuitvoering begrijpen getoetst?	ja/nee
L3.3	Wordt in de toets als kenmerk van de taakuitvoering interpreteren getoetst?	ja/nee
L3.4	Wordt in de toets als kenmerk van de taakuitvoering samenvatten getoetst?	ja/nee
L3.5	Wordt in de toets als kenmerk van de taakuitvoering opzoeken getoetst?	ja/nee
L3.6	Wordt in de toets als kenmerk van de taakuitvoering evalueren getoetst (conform de Toetswijzer PO is het onderwerp 'evalueren' een optioneel en niet verplicht onderdeel van de toets)?	ja/nee/n.v.t.
L3.7	Worden in de toets de verschillende kenmerken van de taakuitvoering evenwichtig getoetst binnen het subdomein zakelijke teksten?	ja/nee
L3.7a	In geval van een digitale adaptieve toets (CAT of MST), verantwoordt de toetsaanbieder, door middel van het opleveren van enkele (voorbeeld)toetspaden, op basis van welke beslisregel (het algoritme in het geval van een CAT en het module-design in geval van een MST) er wordt gegarandeerd dat er een evenwichtige verdeling is tussen opgaven rond zakelijke teksten?	ja/nee/n.v.t.

Toelichting L1.1 t/m L1.3: Tenminste twee van de drie taken dienen in de doorstroomtoets te worden opgenomen.

Toelichting L3.7: Evenwichtig wil zeggen dat de in L3.1 tot en met L3.5 (en eventueel L3.6) genoemde kenmerken gezamenlijk een representatieve weergave zijn van de inhoud van het betreffende subdomein van het terrein Nederlandse taal.

Toelichting L3.7a: Naast het algoritme of module-design dient de werking van de beslisregel te worden aangetoond door middel van het opleveren van enkele (voorbeeld)toetspaden, zoals die door de beslisregel worden gegenereerd.

Toelichting L1 t/m L3: Hierbij zijn de beschrijvingen op niveau 1F en 2F in het Referentiekader (2010, p. 12–14) leidend.

3.4.1.2. Subdomein Fictionele, narratieve en literaire teksten

Kwaliteitseisen

Taken:

Code	Vragen	Mogelijke antwoorden
L4	Lezen de leerlingen minimaal één fictionele, narratieve en/of literaire tekst en maken ze daar opdrachten bij?	ja/nee

Tekstkenmerken:

Code	Vragen	Mogelijke antwoorden
L5	Voldoen de teksten die de leerlingen lezen aan de tekstkenmerken inhoud (thema), vorm (taalgebruik en structuur) en bedoeling (doel), zoals gespecificeerd in Checklist 1 van deze regeling en zoals nader omschreven in het Referentiekader?	ja/nee

Kenmerken van de taakuitvoering:

Code	Vragen	Mogelijke antwoorden
L6.1	Wordt in de toets als kenmerk van de taakuitvoering begrepen getoetst?	ja/nee
L6.2	Wordt in de toets als kenmerk van de taakuitvoering interpreteren getoetst?	ja/nee
L6.3	Optioneel: Wordt in de toets als kenmerk van de taakuitvoering evalueren getoetst (conform de Toetswijzer PO is het onderwerp 'evalueren' een optioneel en niet verplicht onderdeel van de toets)?	ja/nee/n.v.t.
L6.4	Worden in de toets de verschillende kenmerken van de taakuitvoering evenwichtig getoetst binnen het subdomein fictionele, narratieve en literaire teksten?	ja/nee
L6.4a	In geval van een digitale adaptieve toets (CAT of MST), verantwoordt de toetsaanbieder, door middel van het opleveren van enkele (voorbeeld)toetspaden, op basis van welke beslisregel (het algoritme in het geval van een CAT en het module-design in geval van een MST) er wordt gegarandeerd dat er een evenwichtige verdeling is tussen opgaven rond fictionele, narratieve en literaire teksten?	ja/nee/n.v.t.

Toelichting L6.4: Evenwichtig wil zeggen dat de in L6.1 en L6.2 (en eventueel L6.3) genoemde kenmerken gezamenlijk een representatieve weergave zijn van de inhoud van het betreffende subdomein van het terrein Nederlandse taal.

Toelichting L6.4a: Naast het algoritme of module-design dient de werking van de beslisregel te worden aangetoond door middel van het opleveren van enkele (voorbeeld)toetspaden, zoals die door de beslisregel worden gegenereerd.

3.4.2. Wettelijk verplicht subdomein: Taalverzorging

Kwaliteitseisen

De leerlingen worden getoetst op verschillende taalverzorgingscategorieën, zoals gespecificeerd in TV1 t/m TV3.

Beslisregel:

Om aan deze categorie kwaliteitseisen te voldoen, dienen de vragen TV1.1 t/m TV1.7a met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
TV1	Worden de leerlingen getoetst op verschillende categorieën van spellingsproblemen en -regels?	ja/nee
TV1.a	Worden de leerlingen getoetst op regels voor lettergreepgrenzen?	ja/nee
TV1.b	Worden de leerlingen getoetst op morfologische spelling?	ja/nee
TV1.c	Worden de leerlingen getoetst op regels voor werkwoordspelling?	ja/nee
TV1.d	Worden de leerlingen getoetst op overige regels?	ja/nee
TV1.e	Worden de leerlingen getoetst op interpunctie (hoofdletters en leestekens)?	ja/nee
TV2	Verantwoordt de toetsaanbieder dat er een evenwichtige verdeling is tussen opgaven rond niet-werkwoordspelling, werkwoordspelling en interpunctie?	ja/nee
TV3	In geval van een digitale adaptieve toets (CAT of MST), verantwoordt de toetsaanbieder, door middel van het opleveren van enkele (voorbeeld)toetspaden, op basis van welke beslisregel (het algoritme in het geval van een CAT en het module-design in geval van een MST) er wordt gegarandeerd dat er een evenwichtige verdeling is tussen opgaven rond niet-werkwoordspelling, werkwoordspelling en interpunctie?	ja/nee/n.v.t.



Toelichting TV2: Evenwichtig wil zeggen dat de in TV1.a tot en met TV1.e genoemde categorieën van spellingsregels gezamenlijk een representatieve weergave zijn van de inhoud van het betreffende subdomein van het terrein Nederlandse taal.

Toelichting TV3: Naast het algoritme of module-design dient de werking van de beslisregel te worden aangetoond door middel van het opleveren van enkele (voorbeeld)toetspaden, zoals die door de beslisregel worden gegenereerd.

3.4.3. Optioneel domein: Schrijven

Kwaliteitseisen schrijven via indirecte meting

Beslisregel schrijven indirecte meting:

Om aan deze categorie kwaliteitseisen te voldoen, gelden de volgende beslisregels:

1. Taken: van S1.1 t/m S1.3 dienen tenminste twee van de drie vragen met JA te worden beantwoord.
2. Kenmerken van de taakuitvoering: de vragen bij S2 dienen met JA te worden beantwoord.

Taken:

Code	Vragen	Mogelijke antwoorden
S1.1	Lezen de leerlingen tekstsoorten uit de schrijftakencategorie: Correspondentie (om te reviseren of te beoordelen)?	ja/nee
S1.2	Lezen de leerlingen tekstsoorten uit de schrijftakencategorie: Formulieren, berichten, advertenties, aantekeningen (om te reviseren of te beoordelen)?	ja/nee
S1.3	Lezen de leerlingen tekstsoorten uit de schrijftakencategorie: Verslagen, werkstukken, samenvattingen, artikelen (om te reviseren of te beoordelen)?	ja/nee

Kenmerken van de taakuitvoering:

Code	Vragen	Mogelijke antwoorden
S2.1	Maken leerlingen opgaven waarin ze teksten reviseren of beoordelen op het aspect: samenhang?	ja/nee
S2.2	Maken leerlingen opgaven waarin ze teksten reviseren of beoordelen op het aspect: afstemming op doel?	ja/nee
S2.3	Maken leerlingen opgaven waarin ze teksten reviseren of beoordelen op het aspect: afstemming op publiek?	ja/nee
S2.4	Maken leerlingen opgaven waarin ze teksten reviseren of beoordelen op het aspect: woordgebruik en woordenschat?	ja/nee

Toelichting S1 en S2: Hierbij zijn de beschrijvingen op niveau 1F en 2F in het Referentiekader (2010) leidend.

Kwaliteitseisen schrijven directe meting

Beslisregel schrijven directe meting:

Om aan deze categorie kwaliteitseisen te voldoen, gelden de volgende beslisregels:

1. Taken: De vragen bij S3.1 t/m S3.3 dienen alle met JA te worden beantwoord (minimaal één taak per categorie).
2. Kenmerken van de taakuitvoering: de vragen bij S4 dienen met JA te worden beantwoord.

Taken:

Code	Vragen	Mogelijke antwoorden
S3.1	Schrijven de leerlingen een tekst uit de schrijftakencategorie: Correspondentie?	ja/nee
S3.2	Schrijven de leerlingen een tekst uit de schrijftakencategorie: Formulieren, berichten, advertenties en/of aantekeningen?	ja/nee
S3.3	Schrijven de leerlingen een tekst uit de schrijftakencategorie: Verslagen, werkstukken, samenvattingen en/of artikelen?	ja/nee

Kenmerken van de taakuitvoering:

Code	Vragen	Mogelijke antwoorden
S4.1	Worden de teksten die de leerlingen schrijven beoordeeld op het aspect: samenhang?	ja/nee
S4.2	Worden de teksten die de leerlingen schrijven beoordeeld op het aspect: afstemming op doel?	ja/nee
S4.3	Worden de teksten die de leerlingen schrijven beoordeeld op het aspect: afstemming op publiek?	ja/nee
S4.4	Worden de teksten die de leerlingen schrijven beoordeeld op het aspect: woordgebruik en woordenschat?	ja/nee

Toelichting S3.1 t/m S3.3: Minimaal één taak per categorie in de doorstroomtoets dient te worden opgenomen.

Toelichting S3 en S4: Hierbij zijn de beschrijvingen op niveau 1F en 2F in het Referentiekader (2010) leidend.

3.4.4. Optioneel domein: Mondelinge taalvaardigheid

Kwaliteitseisen mondelinge taalvaardigheid via indirecte meting

A. Subdomein Gesprekken

Kwaliteitseisen

Beslisregel mondelinge taalvaardigheid indirecte meting, subdomein Gesprekken:

Om aan deze categorie kwaliteitseisen te voldoen, gelden de volgende beslisregels:

1. Taken: van M1.1 en M1.2 dient één van de twee vragen met JA te worden beantwoord (taak is minimaal één keer aanwezig).
2. Kenmerken van de taakuitvoering: de vragen bij M2 dienen met JA te worden beantwoord.

Taken:

Code	Vragen	Mogelijke antwoorden
M1.1	Luisteren en beoordelen de leerlingen gesprekken uit de takencategorie: discussie en overleg?	ja/nee
M1.2	Luisteren en beoordelen de leerlingen gesprekken uit de takencategorie: informatie uitwisselen?	ja/nee

Kenmerken van de taakuitvoering:

Code	Vragen	Mogelijke antwoorden
M2.1	Maken leerlingen opgaven waarin ze gesprekken beoordelen op het aspect: beurten nemen en bijdragen aan samenhang?	ja/nee
M2.2	Maken leerlingen opgaven waarin ze gesprekken beoordelen op het aspect: afstemming op doel?	ja/nee
M2.3	Maken leerlingen opgaven waarin ze gesprekken beoordelen op het aspect: afstemming op de gespreks-partners?	ja/nee
M2.4	Maken leerlingen opgaven waarin ze gesprekken beoordelen op het aspect: woordgebruik en woorden-schat?	ja/nee

Toelichting M1.1 en M1.2: Tenminste één van de twee taken dient in de doorstroomtoets te worden opgenomen.

Toelichting M1 en M2: Hierbij zijn de beschrijvingen op niveau 1F en 2F in het Referentiekader (2010) leidend.

B. Subdomein Spreken

Kwaliteitseisen

Beslisregel mondelinge taalvaardigheid indirecte meting, subdomein Spreken:

Om aan deze categorie kwaliteitseisen te voldoen, gelden de volgende beslisregels:

1. Taken: vraag M3 dient met JA te worden beantwoord.
2. Kenmerken van de taakuitvoering: de vragen bij M4 dienen met JA te worden beantwoord.

Taken:



Code	Vragen	Mogelijke antwoorden
M3	Luisteren en beoordelen de leerlingen een monoloog?	ja/nee

Kenmerken van de taakuitvoering:

Code	Vragen	Mogelijke antwoorden
M4.1	Maken leerlingen opgaven waarin ze een monoloog beoordelen op het aspect: samenhang?	ja/nee
M4.2	Maken leerlingen opgaven waarin ze een monoloog beoordelen op het aspect: afstemming op doel?	ja/nee
M4.3	Maken leerlingen opgaven waarin ze een monoloog beoordelen op het aspect: afstemming op publiek?	ja/nee
M4.4	Maken leerlingen opgaven waarin ze een monoloog beoordelen op het aspect: woordgebruik en woordenschat?	ja/nee

Toelichting M3 en M4: Hierbij zijn de beschrijvingen op niveau 1F en 2F in het Referentiekader (2010) leidend.

Kwaliteitseisen mondelinge taalvaardigheid via directe meting

A. Subdomein Gesprekken

Kwaliteitseisen

Beslisregel mondelinge taalvaardigheid directe meting, subdomein Gesprekken:

Om aan deze categorie kwaliteitseisen te voldoen, gelden de volgende beslisregels:

1. Taken: van M5.1 en M5.2 dient één van de twee vragen met JA te worden beantwoord (taak is minimaal één keer aanwezig).
2. Kenmerken van de taakuitvoering: de vragen bij M6 dienen met JA te worden beantwoord.

Taken:

Code	Vragen	Mogelijke antwoorden
M5.1	Nemen de leerlingen deel aan gesprekken uit de takencategorie: discussie en overleg?	ja/nee
M5.2	Nemen de leerlingen deel aan gesprekken uit de takencategorie: informatie uitwisselen?	ja/nee

Kenmerken van de taakuitvoering:

Code	Vragen	Mogelijke antwoorden
M6.1	Worden de gesprekken die leerlingen voeren beoordeeld op het aspect: beurten nemen en bijdragen aan samenhang?	ja/nee
M6.2	Worden de gesprekken die leerlingen voeren beoordeeld op het aspect: afstemming op doel?	ja/nee
M6.3	Worden de gesprekken die leerlingen voeren beoordeeld op het aspect: afstemming op de gesprekspartners?	ja/nee
M6.4	Worden de gesprekken die leerlingen voeren beoordeeld op het aspect: woordgebruik en woordenschat?	ja/nee

Toelichting M5.1 en M5.2: Tenminste één van de twee taken dient in de doorstroomtoets te zijn opgenomen.

Toelichting M5 en M6: Hierbij zijn de beschrijvingen op niveau 1F en 2F in het Referentiekader (2010) leidend.

B. Subdomein Luisteren

Kwaliteitseisen

Beslisregel mondelinge taalvaardigheid indirecte meting, subdomein Luisteren:

Om aan deze categorie kwaliteitseisen te voldoen, gelden de volgende beslisregels:

1. Taken: van M7.1 t/m M7.3 dienen twee van de drie vragen met JA te worden beantwoord (taak is minimaal één keer aanwezig).
2. Kenmerken van de taakuitvoering: de vragen bij M8 dienen met JA te worden beantwoord.



Taken:

Code	Vragen	Mogelijke antwoorden
M7.1	Luisteren de leerlingen naar instructies?	ja/nee
M7.2	Luisteren de leerlingen als lid van een live publiek?	ja/nee
M7.3	Luisteren de leerlingen naar radio en/of televisie en/of naar gesproken tekst op internet en/of naar overige door de aanbieder gehanteerde communicatiemiddelen?	ja/nee

Kenmerken van de taakuitvoering:

Code	Vragen	Mogelijke antwoorden
M8.1	Wordt tijdens het luisteren getoetst of leerlingen de luistertekst kunnen begrijpen?	ja/nee
M8.2	Wordt tijdens het luisteren getoetst of leerlingen de luistertekst kunnen interpreteren?	ja/nee
M8.3	Wordt tijdens het luisteren getoetst of leerlingen de luistertekst kunnen samenvatten?	ja/nee

Toelichting M7.1 t/m M7.3: Twee van de drie taken dienen in de doorstroomtoets te worden opgenomen.

Toelichting bij M7 en M8: Hierbij zijn de beschrijvingen op niveau 1F en 2F in het Referentiekader (2010) leidend.

C. Subdomein Spreken

Kwaliteitseisen

Beslisregel mondelinge taalvaardigheid indirecte meting, subdomein Spreken:

Om aan deze categorie kwaliteitseisen te voldoen, gelden de volgende beslisregels:

1. Taken: vraag M9 dient met JA te worden beantwoord.
2. Kenmerken van de taakuitvoering: de vragen bij M10 dienen met JA te worden beantwoord.

Taken:

Code	Vragen	Mogelijke antwoorden
M9	Houden de leerlingen een monoloog?	ja/nee

Kenmerken van de taakuitvoering:

Code	Vragen	Mogelijke antwoorden
M10.1	Wordt de monoloog die de leerlingen houden beoordeeld op het aspect: samenhang?	ja/nee
M10.2	Wordt de monoloog die de leerlingen houden beoordeeld op het aspect: afstemming op doel?	ja/nee
M10.3	Wordt de monoloog die de leerlingen houden beoordeeld op het aspect: afstemming op publiek?	ja/nee
M10.4	Wordt de monoloog die de leerlingen houden beoordeeld op het aspect: woordgebruik en woordenschat?	ja/nee

Toelichting M9 en M10: Hierbij zijn de beschrijvingen op niveau 1F en 2F in het Referentiekader (2010) leidend.

3.4.5. Optioneel subdomein: Begrippenlijst

Kwaliteitseis

Beslisregel Begrippenlijst:

Om aan deze categorie kwaliteitseis te voldoen, dient vraag B1 met JA te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
B1	Worden de leerlingen getoetst met opgaven die betrekking hebben op de begrippenlijst?	ja/nee



Toelichting B1: Hierbij zijn de beschrijvingen op niveau 1F en 2F in het Referentiekader (2010, p. 17–19) leidend.

3.5. Terrein Rekenen

Het Referentiekader onderscheidt voor het terrein Rekenen vier wettelijk verplichte domeinen, te weten Getallen (g), Verhoudingen (vh), Meten en meetkunde (m/mk) en Verbanden (vb). In de doorstroomtoets dienen alle domeinen getoetst te worden. Dit betekent dat alle doorstroomtoetsen moeten voldoen aan de gestelde kwaliteitseisen aan het terrein Rekenen. De gestelde kwaliteitseisen voor de vier domeinen zijn opgenomen in de Toetswijzer PO.

Een individuele rekenopgave kan betrekking hebben op meerdere domeinen van het terrein Rekenen. In dat geval dient de aanbieder dit duidelijk te vermelden, bijvoorbeeld door te kiezen voor het domein dat het beste past bij het beoogde toetsdoel van de betreffende opgave. Tevens dient de aanbieder dan toe te lichten op welke wijze er wordt voldaan aan de voorgeschreven verdeling van toetsopgaven over de domeinen van het terrein Rekenen.

3.5.1. Verdeling toetsopgaven over domeinen in toetsmatrijs en toetssamenstelling

De doorstroomtoets bevat opgaven uit alle domeinen uit het referentiekader Rekenen. Het betreft de domeinen Getallen (g), Verhoudingen (vh), Meten en meetkunde (m/mk) en Verbanden (vb).

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

De opgaven dienen in de toets als volgt over de domeinen te zijn verdeeld:

Code	Vragen	Mogelijke antwoorden
D.G	Ligt het percentage toetsopgaven uit het domein Getallen tussen de 30% en 40% van het totaal aantal toetsopgaven van het terrein Rekenen?	ja/nee
D.Vh	Ligt het percentage toetsopgaven uit het domein Verhoudingen tussen de 20% en 30% van het totaal aantal toetsopgaven van het terrein Rekenen?	ja/nee
D.M/mk	Ligt het percentage toetsopgaven uit het domein Meten & meetkunde tussen de 20% en 30% van het totaal aantal toetsopgaven van het terrein Rekenen?	ja/nee
D.Vb	Ligt het percentage toetsopgaven uit het domein Verbanden tussen de 15% en 20% van het totaal aantal toetsopgaven van het terrein Rekenen?	ja/nee
D.1	In geval van een digitale adaptieve toets (CAT of MST), verantwoordt de toetsaanbieder, door middel van het opleveren van enkele (voorbeeld)toetspaden, op basis van welke beslisregel (het algoritme in het geval van een CAT en het module-design in geval van een MST) de vereiste verdeling van toetsopgaven over de domeinen?	ja/nee/n.v.t.

Toelichting D.G t/m D.Vb: De gestelde percentages toetsopgaven per domein gelden voor zowel papieren toetsen, lineaire digitale toetsen, adaptieve toetsen op itemniveau (CAT) en adaptieve toetsen op moduleniveau (MST).

Toelichting D.1: Naast het algoritme of module-design dient de werking van de beslisregel te worden aangetoond door middel van het opleveren van enkele (voorbeeld)toetspaden, zoals die door de beslisregel worden gegenereerd.

3.5.2. Verdeling toetsopgaven over onderdelen

Elk domein is opgebouwd uit drie onderdelen (zie ook Checklist 3 van deze regeling):

- Notatie, taal en betekenis, waarbij het gaat om de uitspraak, schrijfwijze en betekenis van getallen, symbolen en relaties en om het gebruik van wiskundetaal.
- Met elkaar in verband brengen, waarbij het gaat om het verband tussen begrippen, notaties, getallen en dagelijks spraakgebruik.
- Gebruiken, waarbij het er om gaat rekenkundige vaardigheden in te zetten bij het oplossen van problemen.

De toets bevat opgaven uit alle onderdelen.

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

De relatieve (procentuele) verdeling² van opgaven over de onderscheiden onderdelen is als volgt:

Code	Vragen	Mogelijke antwoorden
O.A	Is iedere toetsopgave, met uitzondering van onderdeel A, onder één onderdeel geplaatst ¹ ?	ja/nee
O.B	Is het percentage toetsopgaven uit onderdeel B ten minste 20 procent van het totaal aantal toetsopgaven van het terrein Rekenen?	ja/nee
O.C	Is het percentage toetsopgaven uit onderdeel C ten minste 30 procent van het totaal aantal toetsopgaven van het terrein Rekenen?	ja/nee
O.BC	Is het percentage toetsopgaven uit onderdeel C hoger dan het percentage toetsopgaven uit onderdeel B?	ja/nee
O.1	In geval van een digitale adaptieve toets (CAT of MST), verantwoordt de toetsaanbieder, door middel van het opleveren van enkele (voorbeeld)toetspaden, op basis van welke beslisregel (het algoritme in het geval van een CAT en het module-design in geval van een MST) de vereiste verdeling van toetsopgaven over de onderdelen?	ja/nee/n.v.t.

¹ Voor onderdeel A is geen minimum vereist, omdat deze inhouden ook kunnen worden gebruikt en toegepast bij de inhouden van B en C.

Toelichting O.A: De toetsopgaven die onder onderdeel A geplaatst zijn, kunnen ook worden gebruikt en toegepast bij de inhouden van B en C. Vandaar dat hier de uitzondering van het slechts plaatsen bij één onderdeel niet geldt en er ook geen percentage vereist wordt.

Toelichting O.A t/m O.BC: De gestelde percentages toetsopgaven per domein gelden voor zowel papieren toetsen, lineaire digitale toetsen, computergestuurde adaptieve toetsen op itemniveau (CAT) en adaptieve toetsen op moduleniveau (MST). Naast het algoritme of module-design is door aanbieders van adaptieve toetsen (CAT en MST) de werking van de beslisregel aan te tonen door middel van het opleveren van enkele toetspaden.

Toelichting O.1: Naast het algoritme of module-design dient de werking van de beslisregel te worden aangetoond door middel van het opleveren van enkele (voorbeeld)toetspaden die voldoen aan de vereiste verdeling van toetsopgaven over de onderdelen, zoals die door de beslisregel worden gegenereerd.

3.5.3. Opgaven met en zonder context

In het terrein Rekenen van een doorstroomtoets moeten zowel opgaven met context als opgaven zonder context (zogenoemde 'kale opgaven') worden opgenomen, een en ander conform de Toetswijzer PO.

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
C1	Is het percentage toetsopgaven van het type contextopgaven voor het domein Getallen ten minste 30 procent van het totaal aantal toetsopgaven van het domein Getallen?	ja/nee
C2	Is het percentage toetsopgaven van het type contextloze opgaven voor het domein Getallen ten minste 20 procent van het totaal aantal toetsopgaven van het domein Getallen?	ja/nee
C3	Is het percentage toetsopgaven van het type contextopgaven voor het domein Verhoudingen ten minste 30 procent van het totaal aantal toetsopgaven van het domein Verhoudingen?	ja/nee
C4	Is het percentage toetsopgaven van het type contextloze opgaven voor het domein Verhoudingen ten minste 20 procent van het totaal aantal toetsopgaven van het domein Verhoudingen?	ja/nee
C5	In geval van een digitale adaptieve toets (CAT of MST), verantwoordt de toetsaanbieder, door middel van het opleveren van enkele (voorbeeld)toetspaden, op basis van welke beslisregel (het algoritme in het geval van een CAT en het module-design in geval van een MST) de vereiste procentuele verdelingen zoals in C1 t/m C4 geschetst?	ja/nee/n.v.t.

² Gebruik hierbij checklist 2 van deze regeling.



Toelichting C1 t/m C4: De gestelde percentages toetsopgaven per domein gelden voor zowel papieren toetsen, lineaire digitale toetsen, computergestuurde adaptieve toetsen op itemniveau (CAT) en adaptieve toetsen op moduleniveau (MST).

Toelichting C5: Naast het algoritme of module-design dient de werking van de beslisregel te worden aangetoond door middel van het opleveren van enkele (voorbeeld)toetspaden met de vereiste percentuele verdelingen zoals in C1 t/m C4 geschetst, zoals die door de beslisregel worden gegene-reerd.

3.5.4. Gebruik kladpapier

Kwaliteitseis

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
K1	Mogen de leerlingen bij het beantwoorden van de (digitale) toetsopgaven fysiek en/of niet programmeerbaar kladpapier gebruiken?	ja/nee
K2	Wordt in de handleiding voor leraren en/of het toetsreglement aangegeven dat zij leerlingen nadrukkelijk dienen te wijzen op het toegestane gebruik van fysiek en/of niet programmeerbaar kladpapier?	ja/nee

3.5.5. Gebruik rekenmachine

Kwaliteitseis

Beslisregel: Om aan deze categorie kwaliteitseis te voldoen, dient deze vraag met JA te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
R1	Mogen de leerlingen bij maximaal 20% van de (digitale) toetsopgaven een niet-grafische en/of niet-programmeerbare rekenmachine gebruiken?	ja/nee

Toelichting: Gebruik van een niet-grafische en/of niet-programmeerbare rekenmachine is bij maximaal 20% van de toetsopgaven toegestaan. Gebruik van een grafische en/of programmeerbare rekenma-chine is niet toegestaan.

3.6. Toetsopgaven Nederlandse taal en Rekenen op niveau 1F en 2F/1S

In het document *Referentiekader taal en Rekenen* (SLO, 2011) is vastgelegd wat leerlingen eind groep 8 van de basisschool moeten kennen en kunnen als het gaat om basisvaardigheden voor Nederlandse taal en Rekenen. Deze kennis en vaardigheden worden in het referentiekader gespecificeerd in een aantal referentieniveaus. De verschillende niveaus worden beschreven in twee kwaliteiten: fundamen-tele kwaliteit (F) en streefkwaliteit (S). De niveaus zijn cumulatief. Een leerling moet op een hoger niveau alle vaardigheden beheersen die op een lager niveau genoemd worden. Deze worden niet telkens herhaald. De beschrijvingen van de niveaus voor Nederlandse taal en Rekenen verschillen. Dit komt doordat de vakken verschillen.

Voor Nederlandse taal zijn vier niveaus beschreven: 1F, 2F, 3F en 4F en voor Rekenen zijn zes niveaus beschreven: 1F, 2F, 3F en 1S, 2S en 3S. Voor het po zijn voor het terrein Nederlandse taal de niveaus 1F en 2F wettelijk vastgelegd. Voor het terrein Rekenen zijn dit voor het po de niveaus 1F en 1S.

Dit betekent dat, in lijn met de Toetswijzer PO, de doorstroomtoets, ongeacht de toetsvorm, voor de terreinen Nederlandse taal en Rekenen toetsopgaven moet bevatten die de inhoud van de niveaus 1F (fundamentele kwaliteit) en 2F / 1S (streefkwaliteit) meten.

Kwaliteitseis

Beslisregel: Om aan deze categorie kwaliteitseis te voldoen, dient vraag TN1 met JA te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
TN1	Zijn de opgaven van de toets voor de terreinen Nederlandse taal en Rekenen evenwichtig samengesteld qua referentieniveaus 1F en 2F / 1S?	ja/nee

Toelichting TN1: Dit aspect dient uit de toetsmatrijs naar voren te komen. Dat wil zeggen dat in de toetsmatrijs de verhouding 1F en 2F / 1S vragen is gespecificeerd en dat in zowel de toetsmatrijs als in de varianten van de doorstroomtoets de verhouding 1F en 2F / 1S vragen evenwichtig is. Evenwichtig wil zeggen dat het aantal en de verhouding 1F en 2F / 1S vragen gezamenlijk inhoudelijk de referentieniveaus 1F en 2F/1S van de terreinen Nederlandse taal en Rekenen representeren, een en ander conform de voorschriften uit de Toetswijzer PO.

Een CAT dient voor elke individuele leerling voor de terreinen Nederlandse taal en Rekenen een uitspraak over het behaalde niveau 1F en/of 2F / 1S te geven. In het geval van een CAT is er verantwoord hoe het 1F en/of 2F/1S niveau van de terreinen Nederlandse taal en Rekenen op representatieve wijze voor alle individuele leerlingen aan bod komt. De aanbieder kan dit verantwoorden aan de hand van de toetsmatrijs, aangevuld met de procesbeschrijving van hoe het algoritme werkt, aangevuld met enkele mogelijke itemsets inclusief de daarin aangeboden toetsopgaven en inclusief een toelichting op het daarbij geldende stopcriterium.

Elke routing binnen het MST design dient voor de terreinen Nederlandse taal en Rekenen een uitspraak over het behaalde niveau 1F en/of 2F / 1S te geven. In het geval van een MST is er verantwoord hoe het 1F en/of 2F/1S niveau van de terreinen Nederlandse taal en Rekenen op een bij de moeilijkheidsgraad van de betreffende route passende wijze voor alle individuele leerlingen aan bod komt. De aanbieder kan dit verantwoorden aan de hand van de toetsmatrijs, aangevuld met de beschrijving van de samenstelling van het MST design, aangevuld met enkele toetspaden uit het design inclusief de daarin aangeboden toetsopgaven.

4. Organisatorische aspecten

4.1. Leerlingrapport

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
LR1	Stelt de toetsaanbieder voor elke individuele leerling die de doorstroomtoets aflegt een leerlingrapport op?	ja/nee
LR2	Bevat het leerlingrapport het toetsadvies, uitgedrukt in één van de toetsadviescategorieën?	ja/nee
LR3	Geeft het leerlingrapport aan welk referentieniveau Lezen de leerling beheerst?	ja/nee
LR4	Geeft het leerlingrapport aan welk referentieniveau Taalverzorging de leerling beheerst?	ja/nee
LR5	Geeft het leerlingrapport aan welk referentieniveau Rekenen de leerling beheerst?	ja/nee
LR6	Bevat het leerlingrapport een deugdelijke toelichting bij het toetsadvies en de beheersing van de referentieniveaus van de leerling die (ook) voor leerlingen, ouders, verzorgers, voogden en afnemers begrijpelijk is?	ja/nee
LR7	Zijn de andere uit te geven rapporten naast het leerlingrapport, en/of niet verplichte gegevens op het leerlingrapport, deugdelijk en helder voor de beoogde gebruiker?	ja/nee/n.v.t.

Toelichting LR1: De toetsaanbieder geeft één of meerdere (model)rapportages ter beoordeling. Wanneer er meer dan één variant van het leerlingrapport bestaat, moeten al deze varianten aangeboden worden, met een uitleg wanneer welke variant wordt gebruikt. Aangezien alle varianten aan de eisen moeten voldoen, worden alle voorbeeldrapporten beoordeeld op de kwaliteitseisen LR2 tot en met LR6. Mochten er rapporten ontbreken (d.w.z. LR1 = nee), dan krijgt de toetsaanbieder een onvoldoende voor 'leerlingrapport'. Wanneer het leerlingrapport wel wordt aangeboden, worden de kwaliteitseisen LR2 tot en met LR6 beoordeeld op basis van die versie(s) van het leerlingrapport.

Toelichting LR2: De toetsaanbieder dient in het leerlingrapport een toetsadvies te geven. De aanbieder hanteert hierbij de door de Rijksoverheid vastgestelde toetsadviescategorieën. Dit toetsadvies is gebaseerd op in ieder geval de resultaten van het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen. Aanvullend mag de toetsaanbieder in de bepaling van het toetsadvies ook één of meerdere van de optionele (sub)domeinen van

Nederlandse taal meenemen die in dit kader in Hoofdstuk 3 beschreven zijn. In de wetenschappelijke verantwoording moet zijn aangegeven op basis van metingen van welke (sub)domeinen het toetsadvies tot stand is gekomen.

Toelichting LR3: In het leerlingrapport moet aangegeven zijn op welk referentieniveau de leerling gepresteerd heeft op het domein Lezen. In het leerlingrapport moet in ieder geval aangegeven worden in welk van de volgende drie categorieën de leerlingprestatie past: onder niveau 1F, op niveau 1F, of op streefniveau 2F. Hoewel in de theorie de categorieën elkaar niet uitsluiten, en de leerling ook op een hoger niveau zou kunnen presteren dan het maximaal gegeven niveau, is in deze regeling vastgesteld dat in de rapportage een leerlingprestatie in één (niet meer en niet minder) van deze drie categorieën geclassificeerd wordt.

Toelichting LR4: In het leerlingrapport moet aangegeven zijn op welk referentieniveau de leerling gepresteerd heeft op het subdomein Taalverzorging. In het leerlingrapport moet in ieder geval aangegeven worden in welk van de volgende drie categorieën de leerlingprestatie past: onder niveau 1F, op niveau 1F, of op streefniveau 2F. Hoewel in de theorie de categorieën elkaar niet uitsluiten, en de leerling ook op een hoger niveau zou kunnen presteren dan het maximaal gegeven niveau, is in deze regeling vastgesteld dat in de rapportage een leerlingprestatie in één (niet meer en niet minder) van deze drie categorieën geclassificeerd wordt.

Toelichting LR5: In het leerlingrapport moet aangegeven zijn op welk referentieniveau de leerling gepresteerd heeft op het terrein Rekenen. In het leerlingrapport moet in ieder geval aangegeven worden in welk van de volgende drie categorieën de leerlingprestatie past: onder niveau 1F, op niveau 1F, of op streefniveau 1S. Hoewel in de theorie de categorieën elkaar niet uitsluiten, en de leerling ook op een hoger niveau zou kunnen presteren dan het maximaal gegeven niveau, is in deze regeling vastgesteld dat in de rapportage een leerlingprestatie in één (niet meer en niet minder) van deze drie categorieën geclassificeerd wordt.

Toelichting LR6: Bij het rapport dient een toegankelijke toelichting te worden geboden, waarin begrijpelijke handvatten gegeven worden voor de interpretatie van het toetsadvies en de beheersing van de referentieniveaus van de leerling door diverse betrokkenen. De aanbieder toont aan op welke wijze de rapportage rekening houdt met de verschillende doelgroepen, te weten leerlingen, ouders, verzorgers, voogden en afnemers.

Toelichting LR7: De toetsaanbieder geeft één of meerdere (model)rapportages zoals bijvoorbeeld een groepsrapportage, schoolrapportage of bovenschoolse rapportage ter beoordeling. Bij de uitgegeven rapporten dient een toegankelijke toelichting te worden geboden aan de beoogde gebruiker, waarin begrijpelijke handvatten gegeven worden voor de interpretatie van gerapporteerde gegevens. De aanbieder toont aan op welke wijze de rapportage rekening houdt met de verschillende doelgroepen, te weten leerkrachten, intern begeleiders, en directeurs of vertegenwoordigers van schoolbesturen of samenwerkingsverbanden.

4.2 Leerlingen

De doorstroomtoets is in principe geschikt voor alle leerlingen, ook voor leerlingen met een specifieke ondersteuningsbehoefte, met uitzondering van leerlingen die onder de ontheffingsgronden vallen.

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA te worden beantwoord. In de toelichting wordt bij enkele kwaliteitseisen aangegeven waar een doorverwijzing naar www.rijksoverheid.nl voor een toets van een aanbieder die voorziet in die behoefte voldoende kan zijn.

Code	Vragen	Mogelijke antwoorden
LL1	Kunnen alle leerlingen, ook met specifieke ondersteuningsbehoeften (uitgezonderd leerlingen die vallen onder de ontheffingsgronden), deelnemen aan de toets?	ja/nee
LL1.a	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met een visuele beperking, zoals leerlingen die slechtziend, blind of kleurenblind zijn?	ja/nee
LL1.b	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor dove/slechthorende leerlingen?	ja/nee
LL1.c	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met een taalontwikkelingsstoornis (TOS)?	ja/nee
LL1.d	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met dyslexie?	ja/nee

Code	Vragen	Mogelijke antwoorden
LL1.e	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met ernstige rekenproblemen of dyscalculie?	ja/nee
LL1.f	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met een motorische beperking?	ja/nee
LL1.g	Is de toets aantoonbaar geschikt, al dan niet door extra ondersteuning, voor leerlingen met ondersteuningsbehoeften voortkomend uit chronische/langdurige ziekte, een psychische aandoening en/of gedragsproblemen of voortkomend uit hier niet nader gespecificeerde beperkingen?	ja/nee
LL2	Staat in de handleiding en/of toetsreglement voor de afnemer bij de toets in volledigheid beschreven op welke wijze de toets rekening houdt met leerlingen met ondersteuningsbehoeften?	ja/nee
LL2.a	Staat omschreven in de handleiding en/of toetsreglement welke hulpmiddelen zijn te gebruiken tijdens de toetsafname?	ja/nee
LL2.b	Staat omschreven in de handleiding en/of toetsreglement welke aanpassingen in de afnamecondities mogelijk zijn?	ja/nee
LL2.c	Is de toets plaatsonafhankelijk af te nemen?	ja/nee
LL2.d	Wordt beschreven wat de mogelijkheden zijn als de geboden ondersteuning niet toereikend is?	ja/nee

Toelichting LL1: Er is automatisch voldaan aan kwaliteitseis LL1 wanneer is voldaan aan de kwaliteitseisen LL1a t/m LL1g. Er is aantoonbaar aandacht besteed aan alle in LL1.a t/m LL1.g genoemde doelgroepen en hun ondersteuningsbehoeften.

Dit houdt het volgende in:

- De aanbieder neemt bij het maken van de toets zoveel mogelijk elementen op die zorgen dat de toets voor alle leerlingen (digitaal) toegankelijk is. Waar mogelijk worden de richtlijnen uit de WCAG 2.1 toegepast. Bij de itemconstructie wordt generiek minstens op de volgende wijze rekening gehouden met (digitale) toegankelijkheid: richtlijnen voor kleurgebruik en contrast bij items, richtlijnen met betrekking tot een ondubbelzinnige lay-out en tekstuele vormgeving, richtlijnen met betrekking tot figuurlijk taalgebruik (als dat geen toetsdoel is) en richtlijnen met betrekking tot het gebruik van ik, gebiedende wijs en vraagzinnen.
- De toets kent geen tijdslimiet en het is mogelijk om pauzes te nemen.
- De toets voorziet in een voorleesfunctie. De wijze waarop kan per toets(soort) verschillen.

Waar de bovengenoemde standaard onvoldoende tegemoet komt aan de doelgroepen genoemd onder LL1.a t/m LL1.g is aantoonbaar geprobeerd de belemmeringen die deze leerlingen ervaren bij het maken van de toets weg te nemen door de mogelijkheid tot inzet van hulpmiddelen of ondersteuning, aanpassing van de afnamecondities of door het aanbieden van aangepaste versies.

Toelichting LL1.a: De toets wordt door de aanbieder (op bestelling) geleverd in een format dat zich leent voor de hulpmiddelen die door leerlingen met een visuele beperking worden gebruikt en met inachtneming van de toetseisen inhoudelijk geschikt zijn gemaakt voor de leerlingen met een visuele beperking. De toetsaanbieder specificeert dit per type visuele beperking (kleurenblind, slechtziend, blind).

Leerlingen die kleurenblind zijn. Bij de constructie van de toets is rekening gehouden met leerlingen die kleurenblind zijn. Dit kan onder andere door kleurgebruik alleen ondersteunend te laten zijn. Daarnaast is er bijvoorbeeld de mogelijkheid tot het aanbieden van een zwart/wit-variant. Relevant voor de normering is dat de aanbieder er rekening mee houdt dat het ontbreken van een kleuren(afbelding) geen invloed heeft op de moeilijkheid van het item. Dit houdt in dat de inhoud en de structuur van het item duidelijk en begrijpelijk moeten zijn, ongeacht het gebruik van kleur.

Bij twijfel mag de leerling gebruik maken van de volgende procedure: voor kleurenblinde leerlingen kan een opzoekhulp worden ingezet die op verzoek van de leerling de kleur van een door de leerling aangewezen vlakdeel benoemt, of een door de leerling aangewezen kleur aanwijst in toets of hulpmiddel.

(Zeer) slechtziende en blinde leerlingen

Aanbieders houden bij het maken van de toets rekening met (digitale) toegankelijkheid voor leerlingen met een visuele beperking, middels bijvoorbeeld contrast, mogelijkheid tot inzoomen en mogelijkheid tot het gebruik van hulpmiddelen, waaronder audio-ondersteuning en gebruik van een brailleleesregel

Een aanbieder kan voor(zeer) slechtziende en blinde leerlingen aanvullende aanpassingen voor de toegankelijkheid van hun toets aan te bieden. Bijvoorbeeld door gebruik te maken van tekenboeken of het aanbieden van een braillevariant. Het uitgangspunt is dat zo min mogelijk wordt afgeweken van de



reguliere toetsen. Daar waar gekozen wordt om de toets aan te passen, verantwoordt de aanbieder deze aanpassingen.

Het aanbieden van een brailleversie van de doorstroomtoets is facultatief. Wel moeten toetsaanbieders leerlingen met die behoefte in de verantwoordingsdocumenten en handleiding voor de afnemer doorverwijzen naar www.rijksoverheid.nl voor een toets van een aanbieder die voorziet in die behoefte.

Afwijkende beoordeling calamiteitentoets: Van de calamiteitentoets moet een versie beschikbaar zijn in braille, in zwart-wit en voorzien zijn van een vorm van audio-ondersteuning om te voldoen aan LL1.a.

Toelichting LL1.b: De aanbieder levert bijvoorbeeld (op bestelling) een aangepaste toets indien de toets gebruikmaakt van functioneel geluid (zoals bijvoorbeeld bij dictee-opgaven). Indien een aanbieder hier niet in kan voorzien, dient deze de school in de verantwoordingsdocumenten en handleiding voor de afnemer te verwijzen naar www.rijksoverheid.nl waar is aangegeven welke toets van weke aanbieder voorziet in die behoefte.

Toelichting LL1.c: De aanbieder geeft aan hoe rekening is gehouden met leerlingen met een taalontwikkelingsstoornis (TOS). Hierbij kan gedacht worden aan de mogelijkheid tot het inzetten van de voorleesfunctie en mogelijke aanpassingen in de afnamecondities, zoals het werken met een afdekblad en het markeren van tekst.

Toelichting LL1.d: De aanbieder geeft aan hoe bij de lay-out en typografie met deze doelgroep rekening wordt gehouden. Denk hierbij bijvoorbeeld aan een schreefloos lettertype met minimale puntgrootte 12, afstand tussen letters, woorden, regels en alinea's, gebruik van witruimte en een eenduidige, rustige lay-out.

Als een leerling door dyslexie moeite heeft met het lezen van teksten op beeldscherm, moet de leerling gebruik kunnen maken van een door de aanbieder geleverde voorleesfunctie. Bij een papieren versie moet de aanbieder (op vraag) een versie leveren die geschikt is voor voorleessoftware. De wijze waarop kan per toets(soort) verschillen.

Andere voorbeelden van hulpmiddelen ter ondersteuning van deze doelgroep zijn de beschikbaarheid van een zoomfunctie en een markeerstift.

Toelichting bij LL1.e: Voor leerlingen met dyscalculie wordt geen aangepaste toets geleverd. In de afnamecondities kan (bijvoorbeeld door in de planning meer pauzes of spreiding van de toetsmomenten op te nemen) rekening worden gehouden met deze ondersteuningsbehoefte. De aanbieder benoemt de eventuele mogelijkheden in diens verantwoordingsdocumenten. Indien de aanbieder het gebruik van een rekenkaart als hulpmiddel toestaat, dient de aanbieder te verantwoorden hoe dit hulpmiddel de toetsdoelstellingen beïnvloedt. Daarbij wordt toegelicht welke impact op de toetsresultaten verwacht wordt bij gebruik van de rekenkaart en onder welke voorwaarden deze tijdens de toets gebruikt mag worden.

Toelichting bij LL1.f: De aanbieder beschrijft hoe de toets toegankelijk is gemaakt voor leerlingen met motorische beperkingen. Er kan bij digitale toetsen bijvoorbeeld worden gedacht aan het bedienbaar maken van functionaliteiten met het toetsenbord in plaats van de muis en bij schriftelijke toetsen aan de mogelijkheid tot het typen van antwoorden op een antwoordblad. Indien het voor een leerling desondanks niet mogelijk is de fysieke handelingen uit te voeren die nodig zijn om de toets te maken, wordt in de handleiding beschreven dat de leerling dan gebruik kan maken van een hulp ter ondersteuning, zoals een schrijfhulp, die antwoorden noteert op aanwijzing van de leerling.

Afwijkende beoordeling calamiteitentoets: kwaliteitseis LL1.f is niet van toepassing voor de calamiteitentoets.

Toelichting bij LL1.g: De aanbieder toont aan welke ondersteuning wordt geboden voor andere beperkingen zoals hier genoemd. Voor bijvoorbeeld leerlingen met ondersteuningsbehoeften voortkomend uit ziekte, psychische aandoeningen en/of gedragsproblemen geeft de aanbieder aan welke aanpassingen mogelijk zijn in de wijze van afnemen en van welke hulpmiddelen zij gebruik kunnen maken tijdens de afname. De aanbieder benoemt daarbij expliciet welke aanpassingen en hulpmiddelen voor deze doelgroepen extra aandacht behoeven, waarbij gedacht kan worden aan het bieden van voldoende tijd, een prikkelarme werkomgeving of individuele afname, extra structureren van wijze van afname en regels die daarbij gelden.



Afwijkende beoordeling calamiteitentoets: kwaliteitseis LL1.g is niet van toepassing voor de calamiteitentoets.

Toelichting LL2: De vorm van de ondersteuning staat beschreven in de handleiding en/of toetsreglement, waaronder de wijze waarop rekening gehouden wordt met doelgroepen in de itemconstructie, door het aanbod van varianten van de toets, de beschikbaarheid van hulpmiddelen en/of via de afnamecondities.

De aanbieder schrijft voor dat bij alle toets(vorm)en alle onderdelen de hulpmiddelen zijn toegestaan binnen de gegeven kaders van de vigerende regelingen. In de Toetswijzer PO staat van een aantal hulpmiddelen voorgeschreven dat deze altijd zijn toegestaan of wanneer al dan niet deze zijn toegestaan. Aanvullend kan een aanbieder andere hulpmiddelen toestaan.

De aanbieder wijst de afnemer er in de handleiding en/of toetsreglement op dat er geen tijdslimiet bestaat en dat de afnemer zelf kan besluiten wat een passende afnameduur (spreiding van de toetsonderdelen over de dag, pauzes e.d.) is bij leerlingen met ondersteuningsbehoeften. Daarnaast wijst de aanbieder de afnemer op de mogelijkheid van de voorleesfunctie voor alle leerlingen die op basis van hun ondersteuningsbehoefte hier gebruik van moeten kunnen maken.

Toelichting LL2.a: Hulpmiddelen kunnen in twee categorieën opgesplitst worden. Enerzijds gaat het om hulpmiddelen die al in de toets(omgeving) zijn opgenomen of ingebouwd. Anderzijds gaat het om de mogelijkheid tot het inzetten van hulpmiddelen van de school of leerling zelf. Bij deze tweede categorie is het belangrijk dat de toets toegankelijk is voor deze hulpmiddelen. De aanbieder beschrijft beide categorieën in de handleiding voor de afnemer en/of toetsreglement.

Toelichting LL2.b: De aanbieder geeft aan welke aanpassingen in de afnamecondities mogelijk zijn en beschrijft hier ten minste de mogelijkheden voor het bieden van voldoende tijd (langer dan de gegeven tijdsindicatie), de inzet van pauzes en de mogelijkheid tot individuele afname. Voor veel aanpassingen in de afnamecondities geldt dat deze voor alle leerlingen ingezet zouden kunnen worden. Voor leerlingen met ondersteuningsbehoeften vragen deze afnamecondities echter extra aandacht. De aanbieder beschrijft de mogelijke aanpassingen in de handleiding voor de afnemer en/of toetsreglement en wijst de afnemer erop dat de afnamecondities voor leerlingen met ondersteuningsbehoeften extra aandacht vragen.

Toelichting LL2.c: Wanneer een leerling aantoonbaar niet in de gelegenheid is om de toets op school te maken, kan de toets op een andere locatie worden afgenomen. Denk hierbij aan leerlingen die in het ziekenhuis liggen, of thuiszittende leerlingen. De aanbieder beschrijft in de handleiding welke (technische) randvoorwaarden hieraan verbonden zijn en hoe de school afname op een andere locatie kan organiseren.

Toelichting LL2.d: De aanbieder toont aan wat er aan maatwerk mogelijk is indien het beschreven aanbod aan hulpmiddelen, toetsvarianten en aanpassingen van afnamecondities niet toereikend zijn voor een leerling met specifieke ondersteuningsbehoeften.

4.3 Afname doorstroomtoets

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
A1	Is er een mogelijkheid voor een tweede afnamemoment binnen de wettelijke gestelde termijn voor leerlingen die bij de eerste afname zijn verhinderd (indien mogelijk met een inhoudelijk en psychometrisch vergelijkbare alternatieve variant)?	ja/nee/n.v.t.
A2.1	Staat in het toetsreglement/de verantwoordingsdocumenten vermeld hoe de directeur de leerlingen voor de toets moet aanmelden?	ja/nee
A2.2	Staat in het toetsreglement/de verantwoordingsdocumenten vermeld welke hulpmiddelen leerlingen mogen gebruiken?	ja/nee
A2.3	Wordt in het toetsreglement/de verantwoordingsdocumenten de voorwaarden aan de (wijze van) geheimhouding van de toetsopgaven door de toetsaanbieder beschreven?	ja/nee
A2.4	Wordt in het toetsreglement/de verantwoordingsdocumenten de voorwaarden aan de (wijze van) geheimhouding van de toetsopgaven door de deelnemende scholen beschreven?	ja/nee
A2.5	Wordt in het toetsreglement/de verantwoordingsdocumenten de voorwaarden aan de wijze van waarop toezicht op de leerlingen wordt gehouden beschreven?	ja/nee

Code	Vragen	Mogelijke antwoorden
A2.6	Wordt in het toetsreglement/de verantwoordingsdocumenten vermeld op welke wijze de toetsopgaven aan het bevoegd gezag ter beschikking worden gesteld?	ja/nee
A2.7	Wordt in het toetsreglement/de verantwoordingsdocumenten de eisen beschreven waar de ICT omgeving (computerconfiguratie) op de school aan moet voldoen om de digitale toets af te kunnen nemen?	ja/nee/n.v.t.
A3	Wordt in het toetsreglement/de verantwoordingsdocumenten vermeld op welke wijze de afnamecondities moeten bijdragen aan optimale prestaties van de leerlingen?	ja/nee

Toelichting A1:

In de verantwoordingsdocumenten en het toetsreglement is beschreven hoe leerlingen, die bij de eerste afname waren verhinderd, alsnog bij een tweede afnamemoment binnen de wettelijk gestelde termijn een doorstroomtoets kunnen maken. Indien mogelijk is dit een inhoudelijk en psychometrisch vergelijkbare alternatieve doorstroomtoets.

Afwijkende beoordeling calamiteitentoets: kwaliteitseis A1 is niet van toepassing voor de calamiteiten-toets.

Toelichting A2.1 t/m A2.6:

In de verantwoordingsdocumenten voor de afnemer is de volgende informatie gegeven:

- A2.1: Hoe dient de schooldirecteur (of andere verantwoordelijke) de leerlingen voor de toets aan te melden?
- A2.2: Welke hulpmiddelen mogen de leerlingen gebruiken?
- A2.3: Welke voorwaarden worden er gesteld aan de (wijze van) geheimhouding van de toetsopgaven door de toetsaanbieder? De toetsaanbieder dient in te gaan op de borging dat er geen itemexposure van ongewenste omvang kan plaatsvinden.
- A2.4: Welke voorwaarden worden er gesteld aan de (wijze van) geheimhouding van de toetsopgaven door de schooldirecteur (of andere verantwoordelijke) en de afnemers?
- A2.5: Welke voorwaarden zijn er verbonden aan de wijze waarop toezicht op de leerlingen moet worden gehouden?
- A2.6: Op welke wijze worden de toetsopgaven aan het bevoegd gezag ter beschikking gesteld?
- A2.7: Wat zijn de eisen waar de ICT omgeving (computerconfiguratie) op de school aan moet voldoen om de digitale toets af te kunnen nemen?

Bij het evalueren van deze vragen kijkt de adviseur niet alleen of deze informatie er is, maar ook of deze duidelijk is en geen tegenstrijdigheden bevat. Mocht achteraf blijken dat in de praktijk de door de aanbieders verstrekte informatie en communicatie dusdanig verwarrend en/of tegenstrijdig zijn dat dit een negatieve impact heeft op de afname en/of de resultaten van de leerlingen, dan kan dit in de jaarlijkse evaluatie van de toets een negatief oordeel opleveren.

Toelichting A3:

Worden in het toetsreglement/de verantwoordingsdocumenten vermeld op welke wijze de afnamecondities moeten bijdragen aan optimale prestaties van de leerlingen? Denk aan voldoende tijd om de toets te kunnen maken en een goede voorbereiding van leerlingen op de afname van de toets.

4.4 Beveiligingsaspecten

Kwaliteitseisen

Indien de aanbieder ISO 27001 is gecertificeerd, volstaat het om dit certificaat inclusief een verklaring van toepasbaarheid (c.q. de scope van het certificaat) aan te leveren. Indien voorgaande wordt aangeleverd, zijn onderstaande kwaliteitseisen voldoende verantwoord.

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
RB1	Heeft de toetsaanbieder een toetsveiligheidsplan opgesteld met daarin opgenomen een risicoanalyse en een PDCA cyclus voor het beschermen van de beschikbaarheid, integriteit en vertrouwelijkheid van informatie (BIV)?	ja/nee
RB1.1	Zijn de functies en verantwoordelijkheden van de functionarissen van de toetsaanbieder die betrokken zijn bij de inhoud en de beveiliging van de doorstroomtoets beschreven?	ja/nee
RB1.2	Hebben de functionarissen van de toetsaanbieder die betrokken zijn bij de inhoud en de beveiliging van de doorstroomtoets aantoonbaar een geheimhoudingsverklaring ondertekend?	ja/nee

Code	Vragen	Mogelijke antwoorden
RB1.3	Beschermt de toetsaanbieder het toetsconstructieproces door middel van een PDCA cyclus van beveiligingsprocedures, inclusief aantoonbare maatregelen ter voorkoming van het voortijdig uitlekken van de toetsinhoud?	ja/nee
RB1.4	Is voor het toetsafnameproces een PDCA cyclus van beveiligingsmaatregelen geïmplementeerd, inclusief aantoonbare maatregelen ter voorkoming van het voortijdig uitlekken van de toetsinhoud?	ja/nee
RB1.5	Zijn de fysieke en elektronische locaties waar de toetsaanbieder examens, systemen en materialen bewaart, aantoonbaar voldoende beveiligd?	ja/nee
RB1.6	Is de digitale en fysieke informatie omtrent de examenketen zowel tijdens opslag als verzending aantoonbaar voldoende beveiligd?	ja/nee
RB1.7	Verzorgt de toetsaanbieder periodiek training over beveiligingsbewustheid aan haar functionarissen die betrokken zijn bij de inhoud en de beveiliging van de doorstroomtoets?	ja/nee
RB2	Waarborgt de toetsaanbieder, met behulp van passende technische en organisatorische beveiligingsmaatregelen, de privacy en persoonsgegevens van de afnemer, leerling en ouder conform de Algemene Verordening Gegevensbescherming (AVG)?	ja/nee
RB3	Heeft de toetsaanbieder een gedegen en compleet actieplan voor het omgaan met diefstal van (delen van) de toetsinhoud en voor het doen verwijderen van internetpagina's die (delen van) de toetsinhoud onthullen?	ja/nee
RB4	Heeft de toetsaanbieder een gedegen en compleet fraudepreventieplan opgesteld met daarin opgenomen een PDCA cyclus voor de beveiliging van het toetsproces en de toetsinhoud?	ja/nee
RB4.1	Zijn er gedegen en complete procedures voor het ontdekken en evalueren van verdachte toetsresultaten en wijzigingen in kenmerken van toetsonderdelen en toetsscores in de loop van de tijd?	ja/nee
RB4.2	Controleert de toetsaanbieder regelmatig het internet en andere media op het bekend raken van (delen) van de inhoud van de toets of andere niet-openbare informatie over de doorstroomtoets?	ja/nee
RB5	Voorziet de toetsaanbieder in een terugvaloptie, die in geval van kleine incidenten op leerling-, klas- of schoolniveau volgens een bijbehorend calamiteitenplan kan worden ingezet?	ja/nee/n.v.t.

Toelichting RB1 t/m RB4.2:

De aanbieder moet overzichtelijk per kwaliteitseis aangeven hoe daar aan is voldaan. Een voorbeeld daarvan is de verwijzing naar een verwerkersovereenkomst, waar op de kwaliteitseisen RB1 t/m RB4.2 wordt ingegaan. Daarbij dient bewijslast te worden overlegd van de bij Bl.2, Bl.3 en Bl.4 gevraagde documenten, processen, contracten en protocollen.

De verantwoordelijkheid voor het goed gebruik van de toets en de daarmee verzamelde informatie in de dagelijkse praktijk ligt bij de school / de eindgebruiker zelf.

Toelichting RB5: Het doel van de terugvaloptie is dat in het geval van een incident of onregelmatigheid de toetsaanbieder tijdig een toetsversie beschikbaar heeft zodat de leerling toch een doorstroomtoets van de initiële toetsaanbieder kan afnemen tijdens de afname- of inhaalperiode.

De toetsaanbieder beschrijft in een bijbehorend calamiteitenplan de procedure ter voorbereiding op incidenten en calamiteiten. Hierin dient in ieder geval vastgelegd te worden:

- Hoe de bij de calamiteiten aangeboden toetsversies van de doorstroomtoets eruitzien: bijvoorbeeld een itembank waaruit kan worden gekozen voor verschillende digitale versies of verschillende versies van een papieren toets. Ook dient het risico dat iteminhoud openbaar komt te worden afgedekt.
- Hoe scholen incidenten bij de toetsaanbieder kunnen melden en welke procedures klaarliggen om hiermee om te gaan. En hoe eventuele incidenten intern beoordeeld worden.
- De aanpak hoe er richting het onderwijsveld gecommuniceerd wordt over eventuele incidenten (d.w.z. via welke kanalen en welke inhoud).
- Dat de benodigde gegevensoverdracht wordt gerealiseerd in het geval dat de calamiteitentoets wordt ingezet.
- Hoe meldingen gearchiveerd worden, in het kader van evaluatie (en indien mogelijk openbaarheid).

Toetsaanbieders zijn bovendien zelf verantwoordelijk voor het in de gaten houden van signalen vanuit scholen, het onderwijsveld en sociale media dat hun doorstroomtoets mogelijk niet aan de voorwaarden voor doorgang vinden voldoet.

Afwijkende beoordeling calamiteitentoets: kwaliteitseis RB5 is niet van toepassing voor de calamiteitentoets.

5. Psychometrische aspecten

5.1. Inleiding

De toetsaanbieder dient nieuw geconstrueerde items te pretesten. Dit om de psychometrische kwaliteit van de items te schatten, om van daaruit te kunnen bepalen welke items mogen worden gebruikt bij de definitieve toetssamenstelling.

De psychometrische kwaliteitsbeoordeling van de doorstroomtoets door de adviseur in opdracht van het CvTE bestaat uit de volgende stappen:

1. Pretestprocedure voor het valideren van nieuw geconstrueerde items, uit te voeren door de toetsaanbieder en bestaande uit:
 - a. Steekproefkader en samenstelling steekproef (paragraaf 5.2.1);
 - b. Kalibratie en kwaliteit van items (paragraaf 5.2.2);
 - c. Toetssamenstelling na pretest (paragraaf 5.2.3).
2. Landelijke item gerelateerde (toetsoverstijgende) normering, uit te voeren door een adviseur in opdracht van het CvTE, zoals beschreven in de separate Regeling Beoordelingsnormen Doorstroomtoetsen po, en bestaande uit:
 - a. Het door de toetsaanbieder aanleveren van de ruwe itemscores, waarna een adviseur deze controleert en analyseert;
 - b. Kalibratie van de gezamenlijk ankeritems;
 - c. Vaststellen van cesuren van referentieniveaus en toetsadviezen;
 - d. Het door de toetsaanbieder vaststellen van de behaalde referentieniveaus en het toetsadvies per leerling.
3. Beoordeling van de kwaliteit en het functioneren van de doorstroomtoets (rapportage achteraf) zoals beschreven in de eisen KF1 t/m KF5, paragraaf 5.5, te verwerken tot één rapportage.

In het geval van een MST, geldt dat voor wat betreft de psychometrische kwaliteitseisen uit hoofdstuk 5 het CvTE per kwaliteitseis mag beoordelen of en welke van de eisen voor een lineaire doorstroomtoets en/of en welke van de eisen voor een computergestuurde adaptieve doorstroomtoets op itemniveau (CAT) van toepassing zijn.

De psychometrische kwaliteitsbeoordeling van de calamiteitentoets wijkt af van bovengenoemd kader. De calamiteitentoets bevat geen items uit het gezamenlijk anker. Ook de normering van de calamiteitentoets wijkt af van de reguliere normering. Dit is nader uitgewerkt in de Regeling Beoordelingsnormen doorstroomtoetsen po.

5.2 Pretestprocedure

De pretestprocedure heeft tot doel om de bruikbaarheid van items te bepalen om vervolgens vast te kunnen stellen welke items in aanmerking komen voor opname in de definitief samen te stellen varianten van de doorstroomtoets.

De pretest kan op twee manieren worden uitgevoerd:

1. Scenario 1 (eerste afname van items los van de operationele afname):

De toetsaanbieder biedt scholen een vrijwillige proeftoets aan, waarin de nieuw geconstrueerde items zijn opgenomen.

 - Doel A (basis-kwaliteitscontrole):

In de meeste gevallen zal het doel van de proeftoets zijn om de kwaliteit en werking van de items in de praktijk te onderzoeken. Dit betreft zowel het – in een minder high-stakes situatie – onderzoeken van de moeilijkheidsgraad en het onderscheidend vermogen van de items, als het onderzoeken van de kwaliteit van de afleiders. Na de proeftoets selecteert de toetsaanbieder de items die in aanmerking komen voor gebruik in een operationele setting. De gegevens worden alleen gebruikt om items te selecteren. Voor parameterschattingen na de afname voor de vaardigheidsschattingen ten behoeve van het leerlingrapport worden deze data niet gebruikt. Voor dit eerste scenario gelden minder stringente eisen voor de steekproefgrootte en het steekproefkader.
 - Doel B (additionele dataverzameling ten behoeve van itemparameterschattingen):

Het is ook mogelijk een pretest te gebruiken om data te verzamelen die wel gebruikt wordt voor itemparameterschattingen ten behoeve van de vaardigheidsschattingen tijdens de afname. Dat is vooral relevant wanneer de afname niet zo ingericht kan zijn dat van te voren bepaald kan worden dat de opgaven door voldoende leerlingen³ gemaakt zullen worden. Bij een MST of een lineaire toets is dat vaak wel mogelijk, echter bij een CAT kan dat lastiger zijn.

³ Voor de aantallen wat als voldoende gezien wordt, zie *Toelichting N1.1a*

Zeker extreme opgaven worden beperkt geselecteerd door het algoritme. Ook als een CAT of MST relatief weinig respondenten heeft kunnen er relatief veel opgaven zijn met minder dan 600 observaties. Als deze opgaven bij een leerling in de afname zitten, dan moeten er ook voor deze opgaven voldoende gegevens zijn om de parameters te kunnen schatten. Onder voorwaarden die beschreven worden onder het steekproefkader kan de data van de pretest ook gebruikt worden om bij de afname deze opgaven te schalen. Deze voorwaarden zijn strenger dan bij Doel A. Het is ook mogelijk om doel B na te streven bij MST of lineaire toetsen, waarbij dan ook de strengere voorwaarden gelden.

2. Scenario 2 (eerste afname binnen de operationele afname; zaaien, d.w.z.):

Opgaven uitproberen als onderdeel van een werkelijke afname waardoor de opgave in dezelfde omstandigheden gemaakt wordt als waarin ze toegepast moeten worden: De toetsaanbieder zaait de nieuwe items in de operationele versie van de doorstroomtoets. Wanneer uit de pretestprocedure blijkt dat de nieuwe items bruikbaar zijn, mag de toetsaanbieder deze direct mee laten tellen voor de berekening van het toetsadvies en/of de score op de referentieniveaus (doel B). Voor een leerling mag het niet herkenbaar zijn of een vraag wel of niet meetelt. Voor dit tweede scenario gelden strengere eisen voor het minimaal vereiste aantal observaties. Als aan alle eisen voldaan wordt, maar alleen te weinig observaties per opgave waargenomen worden, hoeft dit niet tot een onvoldoende beoordeling te leiden: het is dan mogelijk de resultaten verkregen onder scenario 2 te gebruiken voor een basis-kwaliteitscontrole, mits het aantal observaties minimaal gelijk is als in dat geval vereist is (doel A).

Voor de berekening van het toetsadvies en de berekening van de score op de referentieniveaus mag alleen gebruik gemaakt worden van gekalibreerde (pretest)items die onderwijskundig en psychometrisch goed functioneren. Wanneer de toetsaanbieder er toch voor kiest om minder goed functionerende items mee te laten tellen, dan dient de toetsaanbieder dit te beargumenteren in de wetenschappelijke verantwoording van de doorstroomtoets.

De pretestprocedure bestaat in beide scenario's uit drie stappen: 1) Steekproefkader en samenstelling steekproef vaststellen (§ 5.2.1), 2) Kalibratie en kwaliteit van items vaststellen met de verzamelde pretestdata (§ 5.2.2), en 3) Definitieve doorstroomtoets samenstellen (§ 5.2.3).

5.2.1 Steekproefkader en samenstelling steekproef

De onderwijskundig positief beoordeelde nieuw geconstrueerde items worden gepretest volgens de kwaliteitseisen voor het steekproefkader en de samenstelling van de steekproef.

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
N1.1	Is de gebruikte steekproef van voldoende kwaliteit, gezien het doel van de pretest?	ja/nee
N1.1a	Is de steekproef groot genoeg?	ja/nee
N1.1b	Is de steekproef relevant?	ja/nee
N1.2	Heeft de pretest plaatsgevonden conform het dataverzamelingsdesign; en is het dataverzamelingsdesign van de pretest adequaat?	ja/nee/n.v.t.
N1.3	Is, in het geval van open opgaven in een pretest, de beoordelaarsovereenstemming gegeven?	ja/nee/n.v.t.

Toelichting N1.1:

De beantwoording van deze vraag is afhankelijk van het doel van de steekproef. Hierboven is al verwezen naar de twee scenario's die van toepassing zijn. De vraag naar de kwaliteit van de steekproef is verder opgedeeld naar de grootte en de relevantie van de steekproef. Er is automatisch voldaan aan kwaliteitseis N1.1, wanneer is voldaan aan de kwaliteitseisen N1.1a en N1.1b.

Toelichting N1.1a:

Als onder *scenario 1 doel A* wordt nagestreefd, het bepalen of een opgave van voldoende kwaliteit is om opgenomen te mogen worden in de operationele doorstroomtoets, is het benodigde aantal leerlingen beperkt. Uit een relatief beperkt aantal observaties valt immers te achterhalen of een opgave goed genoeg functioneert om opgenomen te worden in de afname. De focus moet hierbij vooral liggen op het onderscheidend vermogen van de opgave en de werking van de afleiders in het geval van een meerkeuze-opgave. Om de kans te verkleinen dat het onderscheidend vermogen bij de afname negatief blijkt te zijn is het verstandig een redelijke ondergrens van de Rir-waarde of de geschatte a-parameter te kiezen. De aanbieder is hier vrij in, maar dient wel aan te geven welk criterium gehanteerd is (en waarom). Het aantal observaties wordt als voldoende gezien wanneer



iedere opgave door minstens 200 leerlingen gemaakt is. Uiteraard geven meer gegevens meer zekerheid.

Wanneer onder *scenario 1 doel B* wordt nagestreefd, de additionele dataverzameling ten behoeve van itemparameterschattingen, dan dient het aantal observaties hoger te zijn. Op het moment van itemkalibratie t.b.v. de parameterschattingen voor de leerlingrapportage, geldt bij een 2PL itemresponsemodel een minimum van 600 observaties (dezelfde aantallen als gespecificeerd staan onder scenario 2 hieronder).

Onder *scenario 2* worden de opgaven voor het eerst afgenomen direct onder operationele afnameconditie. Deze opgaven moeten van tevoren door de aanbieder zelf of het CvTE via een vroegtijdige onderwijskundige beoordeling onderwijskundig als kwalitatief voldoende beoordeeld zijn. Deze opgaven mogen zich noch onderwijskundig opzicht, noch in psychometrisch opzicht van de andere opgaven in de toets onderscheiden. Omdat leerlingen er tijd en moeite in steken is het terecht als deze items dan meetellen voor de rapportage. In dat geval kunnen deze items tijdens de campagne bijgeschaald worden (doel B).

Wanneer onder scenario 2 doel A wordt nagestreefd, de basis-kwaliteitscontrole van items, is het minimaal vereiste aantal observaties gelijk aan dat bij doel A van scenario 1 (een minimum van 200 leerlingen).

Wanneer onder scenario 2 doel B wordt nagestreefd, en items mee tellen voor de terugrapportage naar de leerling, moeten de itemparameters goed geschat worden. Daartoe moeten er voldoende observaties per item zijn. Voor dichotoom gescoorde items geldt bij een 2PL itemresponsemodel een minimum van 600 observaties (dezelfde aantallen als gespecificeerd staan onder scenario 1, doel B, hierboven). Aanbieders mogen van deze aantallen afwijken, mits dit gedegen en volledig wordt onderbouwd.

Voor alle items die meetellen voor de rapportage dienen de bovengenoemde aantallen behaald te worden. Bij opgaven die bijvoorbeeld in een CAT of MST te weinig door het algoritme geselecteerd worden, worden de aantallen niet altijd tijdens de afname gehaald. In die gevallen mogen de gegevens van de pretest ook gebruikt worden voor de kalibratie om die itemparameters te schatten. De genoemde aantallen observaties per opgave kunnen dan gehaald worden met de som van de pretest-observaties en de afname-observaties. Houd er dan rekening mee dat dan ook aan andere steekproefvragen voldaan moet zijn aangaande representativiteit.

Gekloonde opgaven zijn in beginsel nieuwe opgaven, en worden zo ook beoordeeld. Wanneer er sprake is van (zeer) kleine wijzigingen, zoals de aanpassing van een munteenheid (van euro naar dollar), of de aanpassing van een naam (bijvoorbeeld ten behoeve van inclusie), waarbij de opgave inhoudelijk hetzelfde blijft, dan kunnen dergelijke items als onveranderd gezien worden. De toetsaanbieder dient dan wel duidelijk aan te geven welke (inhoudelijke) aanpassing gedaan is, bij voorkeur met een rationale waaruit blijkt waarom de aanpassing geen impact heeft.

Wanneer de gekloonde opgave in termen van relevantie en/of objectiviteit verandert ten opzichte van het originele item, dan is er wel sprake van een nieuwe opgave. Meer specifiek is er sprake van een nieuwe opgave wanneer:

- de gekloonde opgave een relatie heeft met een ander of gewijzigd leerdoel;
- de gekloonde opgave naar verwachting een andere of gewijzigde moeilijkheidsgraad heeft;
- de gekloonde opgave naar een ander of gewijzigd modelantwoord leidt.

Wanneer eerder al data verzameld zijn bij het item in een andere vorm dient ook bij kleine wijzigingen het gekloonde item een nieuw label te krijgen.

Toelichting N1.1b:

De steekproef dient relevant te zijn voor de toepassing. Dit betekent dat de steekproef van de pretest overeenkomstig dient te zijn met de verwachte populatie die de toets van de aanbieder regulier afneemt. De aanbieder wordt gevraagd om – indien van toegevoegde waarde – effectgroottes aan te leveren ten behoeve van de gemaakte vergelijking.

Voor scenario 2, het zaaien in de afname, is dat relatief makkelijk in te richten door de zaai-items verstandig op scholen in te zetten. Deze steekproefprocedure dient door de aanbieder beschreven te worden. Een ander voordeel bij scenario 2 is dat de afnamecondities per definitie identiek zijn aan reguliere afnamecondities.

In het geval van scenario 1 dient een aparte steekproef getrokken te worden. De eisen aan deze steekproef en de afnamecondities verschillen voor doel A en doel B bij dit scenario. Onder doel A, de

basis-kwaliteitscontrole, is het belangrijk dat de steekproef lijkt op de toets-populatie. Hierbij dient rekening te worden gehouden met de variabelen in onderstaande tabel. Minstens twee niveaus van iedere variabele moeten mee genomen worden. Voor de variabele schoolweging geldt een ondergrens van drie niveaus. Als scenario 1 doel B nagestreefd wordt, dan wordt verwacht dat de afname onder dezelfde condities plaatsvindt als de reguliere afname, een high-stakes afname is hierbij wenselijk. Er dient door de aanbieder beschreven te worden hoe dat zoveel mogelijk bereikt wordt. Tevens dient aangetoond te worden dat de verdeling van de steekproef goed gelijkend is aan de toets-populatie. Daarvoor mag de toets-populatie van het voorgaande afnamejaar als referentie en steekproefkader gebruikt worden. Deze populatie en steekproef dienen dan beschreven te worden op basis van de variabelen en niveaus zoals beschreven in de onderstaande tabel.

Variabele	#N	Welke niveaus (NB: #N = aantal niveaus)
regio	4	noord, zuid, oost, west
stedelijkheid	2	verstedelijkt, landelijk ¹
schoolgrootte	3	aantal leerlingen: < 100, 100–300, > 300 leerlingen
schoolweging	5	gewichtscategorieën: < 23 / 23–26,99 / 27–32,99 / 33–36,99 / ≥ 37

¹ De mate van verstedelijking is terug te voeren op de bij het CBS gebruikelijke indeling naar vijf niveaus, namelijk zeer sterk, sterk, matig, weinig en niet verstedelijkt. De vijfdeling is hier teruggebracht naar een tweedeling in enerzijds verstedelijkt (zeer sterk en sterk) en anderzijds landelijk (matig, weinig en niet verstedelijkt) die in de praktijk goed lijkt te volstaan (c. f. van Bostel en Hemker, 2009). De toetsaanbieder kan ervoor kiezen om van deze indeling af te wijken, mits deze keuze is voldoende onderbouwd.

Onder *scenario 1* wordt een positief antwoord op deze vraag gegeven wanneer aan de volgende vijf voorwaarden is voldaan:

- 1) Op alle in de tabel genoemde variabelen zitten er op minimaal twee niveaus per variabele en op minimaal drie niveaus van de variabele schoolweging gegevens in de steekproef;
- 2) Er is aangegeven hoe de steekproef getrokken is;
- 3) Er is aangegeven onder welke omstandigheden de data verzameld zijn;
- 4) Er is aangegeven hoe de steekproef zich verhoudt tot de landelijke populatiewaarden van leerlingen in groep 8 met betrekking tot de in de tabel genoemde niveaus per variabele;
- 5) Er is aangegeven of en hoeveel S(B)O scholen hebben mee gedaan.

Als bij (3) de omstandigheden afwijken van de werkelijke afname wordt aangegeven welke impact verwacht wordt en hoe daar in de selectie van opgaven voor de operationele afname rekening mee is gehouden.

Onder *scenario 2* wordt een positief antwoord op deze vraag gegeven wanneer aan de volgende twee voorwaarden is voldaan:

- 1) Er is aangetoond dat de populatie voor de zaai-items niet structureel anders is dan de populatie(s) van de rest van de opgaven.
- 2) Bij de observaties is een steekproefkader gegeven (c.q. een design van hoe de toetsboekjes, de CAT toetsversies en/of de MST toetspaden met zaai-items en ankeropgaven worden ingezet) met een beschrijving van de procedure waarmee de steekproef van observaties tot stand is gekomen.

Afwijkende beoordeling calamiteitentoets: De eis dat de steekproef van de pretest overeen dient te komen met de populatie die de toets van de aanbieder regulier afneemt, geldt niet voor de calamiteitentoets. Voor de calamiteitentoets is 'leerlingen uit groep 8' de best passende operationalisering van de beoogde doelgroep.

Toelichting N1.2:

Voor doel A, de basis-kwaliteitscontrole, is deze eis niet van toepassing, omdat bij de normering de (definitieve) gebruikte parameters bepaald worden op basis van de operationele afname. De kwaliteit van de pretest wordt dan alleen bepaald op basis van N1.1.

Wanneer er sprake is van doel B is het van belang dat de parameters per (sub)domein op dezelfde schaal liggen. Hiervoor is het van belang dat er sprake is van een afnamedesign waarbij dit mogelijk is. Wanneer de pretest gedaan wordt door middel van 'zaai-opgaven' in de afname dan wordt aangegeven hoe dat 'zaai-design' er uitziet. Dit design is meegegeven in de technische verantwoording van de pretest en bevat indien van toepassing in ieder geval informatie over:

- het totaal aantal boekjes ('booklets');
 - aantal observaties per boekje;
 - het aantal items dat tussen de opgavenboekjes onderling overlapt;
 - het aantal items, en de kwaliteit van de items, dat de pretest verbindt met de bestaande schaal.
- Met behulp van deze informatie kan de aanbieder laten zien dat het plaatsen van de itemparameters van alle opgaven op dezelfde schaal als waar de meting over gaat mogelijk is, meestal door middel van een verbonden design.

Wanneer de pretestdata zijn verzameld met een CAT is het niet zinvol om het hele onvolledige design weer te geven, omdat iedere leerling dan in principe een unieke verzameling items maakt. Het dient wel duidelijk te zijn op welke gegevens de linking en normering gebaseerd zijn. Bij een MST of een lineaire afname is het meeleveren van het design wel zinvol.

In het geval van een CAT is, net als bij de andere toetsvormen, aangegeven hoeveel zaai-items een kandidaat kan krijgen, en hoe die opgaven toegewezen worden.

Ook moet duidelijk zijn op welke posities de verschillende opgaven ingezet zijn, zodat door de toetsaanbieder op gedegen en complete wijze beargumenteerd kan worden op welke wijze er rekening is gehouden met vermoeidheids- en/of volgorde effecten.

Toelichting N1.3: De kwaliteit van de niet geautomatiseerd gescoorde open opgaven wordt ook beïnvloed door de beoordelaarsovereenstemming. In de pretest moet er bij niet geautomatiseerd gescoorde open opgaven daarom ook beoordelaarsovereenstemmingsonderzoek worden uitgevoerd. De aanbieder toont aan dat de mate van overeenstemming tussen de experts resulteert in een coëfficiënt voor beoordelaarsovereenstemming waarbij de proportie overeenstemming representatief is voor een instrument voor belangrijke beslissingen (high-stakes instrument). In het geval van doel A, de basis-kwaliteitscontrole, moet aangegeven worden op basis van welk criterium items vanwege een lage beoordelaarsovereenstemming niet opgenomen worden, en wanneer opgaven wel opgenomen kunnen worden. Bij doel B moet aangegeven worden wat de impact is op de geschatte parameters.

5.2.2 Itemkalibratie en toetskwaliteit

Na het verzamelen van de observaties voert de toetsaanbieder een kalibratie uit met de verzamelde testdata. In het geval dat opgaven al eerder zijn afgenomen (bij een itembank of bij pretestgegevens) kan er een schaal gekalibreerd worden die op basis van de afnamegegevens bestendig moet worden. Op basis van de afnamegegevens kunnen de (goed functionerende) zaai-items op de schaal geplaatst worden. In de kalibratie schat de toetsaanbieder de itemparameters, zodat de kwaliteit van de gepreteste items kan worden vastgesteld.

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
KA1	Zijn de nieuwe items psychometrisch gezien van voldoende kwaliteit?	ja/nee
KA2	Is er op gedegen en complete wijze beschreven dat de itemparameters van de geselecteerde gekalibreerde items toepasbaar zijn?	ja/nee
KA3	Worden beslisregels voor adaptieve toetsen geëxpliciteerd?	ja/nee/n.v.t.
KA4	Is er een schatting gemaakt van de populatieparameters?	ja/nee
KA5	Zijn de metingen van de individuele leerlingen voldoende betrouwbaar om uitspraken op individueel niveau te kunnen doen?	ja/nee
KA6	Past het psychometrisch model wat gebruikt wordt voor de toetssamenstelling bij de data?	ja/nee
KA7	Is er een indicatie gegeven van het onderscheidend vermogen van de nieuwe items?	ja/nee

Toelichting KA1:

De toetsaanbieder schat van alle gepreteste items de moeilijkheidsgraad (p-waarde of een vergelijkbare parameter), het discriminerend vermogen (Rir-waarde of vergelijkbare parameters) en, in het geval van meerkeuzevragen de kwaliteit van de afleiders (a-waarde, Rar-waarde of vergelijkbare waarden). De aanbieder dient te verantwoorden op basis van welke criteria opgaven toegelaten worden voor de operationele afname.

De toetsaanbieder selecteert vervolgens de items die psychometrisch gezien van voldoende kwaliteit zijn om te kunnen gebruiken in de samen te stellen doorstroomtoets. Dat betekent dat de toetsaanbieder op gedegen en complete wijze onderbouwt dat de geselecteerde items aan minimaal de volgende eisen voldoen:

- de moeilijkheidsgraad is passend voor het meetdoel en de doelgroep (d.w.z. voor lineaire toetsen vergelijkbaar met een p-waarde tussen de gokkans en 0.95 en voor adaptieve toetsen passend bij de moeilijkheidsgraad van het betreffende design en/of toetspad);
- het discriminerend vermogen is vergelijkbaar met een Rir-waarde > 0.15 , rekening houdend met de beperkte vaardigheidsrange van de leerlingen in het geval van een adaptieve toets;

- de kwaliteit van de afleiders is, in het geval van meerkeuzevragen, vergelijkbaar met een a-waarde > 0.05 ;
- de interbeoordelaarsovereenstemming, in het geval van open vragen, vergelijkbaar is met een interbeoordelaarsovereenstemmingscoëfficiënt > 0.80 .

De toetsaanbieder verantwoordt op gedegen en complete wijze de keuze voor de geselecteerde items die in aanmerking komen voor de doorstroomtoets en waarom items die niet aan deze eisen voldoen onder gegeven omstandigheden toch zijn mee genomen.

Toelichting KA2:

De toetsaanbieder beschrijft op gedegen en complete wijze dat de itemparameters van de geselecteerde items toepasbaar zijn. Dit wordt gedaan door (1) aan te tonen dat de itemparameters tussen verschillende afnamen en boekjes niet veranderen, en (2) een DIF analyse (differential item functioning) op itemniveau uit te voeren op tenminste de achtergrondvariabele schoolweging. Wanneer er sprake is van items met DIF, dient de toetsaanbieder aan te tonen dat dit in overeenstemming is met de verwachting op basis van de relevante literatuur. De aanbieder wordt gevraagd de DIF analyse grafisch te onderbouwen en indien mogelijk effectgroottes aan te leveren (bijvoorbeeld het gemiddelde absolute verschil tussen de geobserveerde en verwachte proporties juiste antwoorden, gegeven het IRT-model). Het ontbreken van de genoemde DIF-analyse of een gedegen uitleg kan resulteren in een onvoldoende voor deze kwaliteitseis.

Toelichting KA3 (alleen van toepassing voor adaptieve toetsen):

Een adaptieve toets is een toets die zich aanpast aan het individuele vaardigheidsniveau van de persoon die de toets maakt. In plaats van alle deelnemers dezelfde set vragen voor te leggen, past een adaptieve toets de complexiteit van de vragen aan op basis van de correctheid van de eerdere antwoorden van de deelnemer. Het idee hierachter is dat hierdoor een efficiënte schatting van het vaardigheidsniveau van de leerling kan worden verkregen. Derhalve dienen van dit type doorstroomtoets de beslisregels of de algoritmes voor de samenstelling van de toets te zijn geëxpliciteerd. Voor een CAT dient de toetsaanbieder op gedegen en complete wijze te verantwoorden: (1) hoe de toets wordt gestart, (2) hoe de keuze voor een volgend item en wordt gemaakt, en (3) wanneer de toets wordt beëindigd (maximum aantal items en standard error). Voor een MST dient de toetsaanbieder op gedegen en complete wijze te verantwoorden: (1) hoe de toets wordt gestart, (2) hoe de keuze voor het volgende blokje wordt gemaakt en (3) wanneer de toets wordt beëindigd. Naast dat de beslisregels voor de samenstelling van de toets geëxpliciteerd dienen te zijn, is het belangrijk dat tevens beslisregels geëxpliciteerd zijn om dekking van het inhoudelijke domein te waarborgen per individuele afname. Dit houdt in dat indien er bijvoorbeeld vier verschillende domeinen van rekenen worden bevraagd, de beslisregels van de adaptieve toets ervoor dienen te zorgen dat ook uit alle domeinen voldoende opgaven worden afgenomen. Wanneer of de startprocedure (CAT), of de selectieprocedure (CAT en MST) of de stopprocedure (CAT) niet op gedegen en complete wijze is beschreven, kan dit resulteren in een onvoldoende voor deze kwaliteitseis.

Toelichting KA4:

De aanbieder van de toets dient informatie te verstrekken over de aannames die worden gemaakt over het type verdeling (normaal of anders), evenals het gemiddelde, de standaardafwijking en hun schattingsfout (standard error) op de vaardigheidsverdeling. Daarnaast moet de toetsaanbieder verantwoorden hoe deze schattingen zijn verkregen en aangeven voor welke populatie(s) deze parameters gelden. Deze gegevens dienen in ieder geval gegeven te worden voor de leerlingpopulatie van het reguliere basisonderwijs.

Indien de doorstroomtoets eerder is afgenomen is deze schatting gebaseerd op de daadwerkelijke populatievaardigheid van het voorgaande jaar. Bij een eerste afname is dat niet mogelijk, en dient de toetsaanbieder diens verwachtingen aan te geven zoals ook beschreven bij de jaarlijkse rapportage na de afname (zie hoofdstuk 5.4).

Toelichting KA5:

De meting van de vaardigheid van een leerling moet voldoende betrouwbaar zijn. Dit betekent dat de schattingsfout (standard error) van iedere geschatte vaardigheid beperkt is. Deze schattingsfout zal in de praktijk (waarschijnlijk) niet voor iedere leerling even groot zijn. Daarom dient aangegeven te worden wat de verdeling van de schattingsfouten is. Om de grootte van de schattingsfout op waarde te kunnen schatten zal de standaardafwijking zoals deze gegeven wordt in KA4 als referentie dienen.

De eis van de minimale betrouwbaarheid (de manier om de schattingsfout van de persoonsparameters in te schatten) wordt op drie manieren beoordeeld:

- Is er een gedegen en complete beschrijving van de procedure, waaruit blijkt dat de betrouwbaarheid op de juiste manier is berekend?
- Is de hoogte van de betrouwbaarheid rond de cesuurpunten van de referentieniveaus groter dan 0,70?

- Is de hoogte van de betrouwbaarheid van de schaal waarop de toetsadviezen gegeven worden minimaal 0,80?

Toelichting KA6:

Alle psychometrische modellen werken met aannames en voorwaarden waaraan voldaan moet worden, wil het model bruikbaar zijn. Past het – voor de toetssamenstelling gebruikte model – bij de data? De aanbieder dient de itemfit grafisch te onderbouwen, indien mogelijk op basis van voor het gebruikte model en de toets toepasbare statistieken.

Toelichting KA7:

De aangeboden opgaven moeten voldoende onderscheidend zijn. In het geval van een adaptieve toets met een itembank (CAT) zullen items met weinig onderscheidend vermogen zelden of nooit aangeboden worden. Deze items vormen daarmee een overschatting van de vulling van de bank in aantallen opgaven. In het geval van een lineaire toets of een MST, is het onwenselijk dat een leerling tijd besteed aan opgaven met een extreem laag gewicht.

Om een positief oordeel op deze kwaliteitseis te krijgen wordt aangetoond dat alle opgaven een voldoende onderscheidend vermogen hebben. Dat kan door te laten zien dat het onderscheidend vermogen van de opgaven in de bank of toets, aangetoond door de a-parameter, niet kleiner is dan 10% van het gemiddelde onderscheidend vermogen van de gezamenlijke ankeritems van de meest recent gekalibreerde jaarset. Dit dient berekend te worden voor het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen.

5.2.3. Toetssamenstelling

Na afronding van de kalibratie van de nieuwe items, stelt de toetsaanbieder de definitieve doorstroomtoets samen.

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
T1	Bevat de doorstroomtoets voor de terreinen Nederlandse taal en Rekenen in ieder geval vragen met een moeilijkheidsgraad tussen de cesuur van referentieniveau 1F en 2F/1S?	ja/nee
T2	Bevat de doorstroomtoets de door het CvTE voorgeschreven subset van gezamenlijke ankeritems voor het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen op de in de Regeling Beoordelingsnormen Doorstroomtoetsen po beschreven wijze?	ja/nee/n.v.t.
T3	Bestaat een papieren toets of een lineaire digitale toets jaarlijks uit nieuwe items voor het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen, met uitzondering van de items van het gezamenlijk anker?	ja/nee/n.v.t.
T4	Is er in het geval van een CAT of MST sprake van een adequate verversingsstrategie?	ja/nee/n.v.t.
T5	Heeft de itembank van een CAT separaat voor het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen een omvang van minimaal 6x het aantal items van een maximale operationele toetsversie?	ja/nee/n.v.t.

Toelichting T1: Voor zowel een MST als papieren en digitale lineaire doorstroomtoetsen geldt dat de toets voor de terreinen Nederlandse taal en Rekenen in ieder geval vragen moet bevatten die inhoudelijk aansluiten bij de referentieniveaus 1F en 2F/1S en die een moeilijkheidsgraad tussen de cesuur van referentieniveau 1F en 2F/1S hebben. In het geval van een CAT is de itemselectie geoptimaliseerd rond de opeenvolgende vaardigheidsschattingen.

Toelichting T2: De door het CvTE toegewezen subset van gezamenlijke ankeritems is opgenomen in (de varianten van) de doorstroomtoets. Eventuele (minimale) aanpassingen op deze items t.o.v. het moederitem zijn alleen toegestaan met akkoord en na beoordeling van het CvTE.

De door het CvTE toegewezen subset van gezamenlijke ankeritems moet op een zodanige wijze over de toets worden verspreid, dat de positie in de toets zo min mogelijk invloed op de prestatie van de leerling heeft en de ankeritems qua lay-out en tekstuele vormgeving niet als opvallend afwijkend in de toets herkenbaar zijn, een en ander zoals gespecificeerd in de Regeling Beoordelingsnormen Doorstroomtoetsen po. De aanbieder legt duidelijk uit hoe de ankeritems over de toets zijn verspreid. Omdat de vergelijkbaarheid van de ankeritems tussen verschillende toetsaanbieders gewaarborgd moet worden, is het desondanks belangrijk dat de ankeritems nagenoeg onveranderd in de toets

worden opgenomen. Daarom dient de toetsaanbieder de voorgenomen veranderingen aan de ankeritems uit te leggen en te rechtvaardigen.

Afwijkende beoordeling calamiteitentoets: kwaliteitseis T2 is niet van toepassing voor de calamiteitentoets.

Toelichting T3 (alleen van toepassing voor lineaire toetsen): Een lineaire toets bestaat jaarlijks volledig uit nieuwe gekalibreerde items voor het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen, met uitzondering van de items van het gezamenlijk anker en een intern anker. Als een intern anker wordt gebruikt moet de noodzakelijkheid ervan voldoende worden uitgelegd.

Afwijkende beoordeling calamiteitentoets: kwaliteitseis T3 is niet van toepassing voor de calamiteitentoets.

Toelichting T4 (alleen van toepassing bij al eerder erkende of vastgestelde adaptieve toetsen): In het geval van een MST die bestaat uit meerdere lagen van modules (c.q. routes) of in het geval van een CAT, moet er sprake zijn van een adequate verversingsstrategie. De verversing moet op itemniveau gedocumenteerd zijn. Daarbij geldt het volgende:

- De aanbieder hanteert jaarlijks een adequate verversingsstrategie voor de opgaven van de itembank. De toetsaanbieder geeft op gedegen en complete wijze aan hoeveel procent van de opgaven is verversed en onderbouwt dat de toegepaste vernieuwing voldoende is om overmatige blootstelling van items te voorkomen (exposure control).
- Items die uitgelekt lijken, moeten verwijderd worden. De toetsaanbieder geeft op gedegen en complete wijze aan hoeveel items er vermoedelijk zijn uitgelekt en hoeveel van deze items uit de itembank zijn verwijderd.
- Voor items die over de jaren heen andere itemparameters hebben, moet de toetsaanbieder op gedegen en complete wijze aangeven hoeveel items er veranderende itemparameters hebben en hoeveel van deze items uit de itembank zijn verwijderd.
- Voor items die in het vorige jaar door de algoritme niet geselecteerd zijn, geeft de toetsaanbieder op gedegen en complete wijze aan, waarom deze niet of wel uit de itembank zijn verwijderd.

De punten b en c kunnen onderzocht worden met een statistische toets voor DIF en/of met een statistische toets voor Item Parameter Drift (IPD), mits het aantal observaties de gekozen analyse-methode toelaat. De aanbieder dient de keuze voor de gebruikte methode volledig te onderbouwen.

Afwijkende beoordeling calamiteitentoets: kwaliteitseis T4 is niet van toepassing voor de calamiteitentoets.

5.3 Normering ten behoeve van leerlingrapportage

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
NR1	Is de gekozen normeringsmethode voldoende omschreven?	ja/nee
NR2	Zijn de gerapporteerde scores voldoende nauwkeurig?	ja/nee/n.v.t.

Toelichting NR1: Vanuit de landelijke normering worden aan alle aanbieders 2PL parameterschattingen geleverd. Hoe past de toetsaanbieder deze toe bij diens normering? Zodra de gehanteerde methode afwijkt van de landelijke normering zoals omschreven in de Regeling beoordelingsnormen doorstroomtoetsen PO (bijvoorbeeld een normering via een discrete (standaard)scoreschaal of verschillen in de meegekalibreerde onderwijstypen), dient de toetsaanbieder de gehanteerde normering te verklaren en deze te onderbouwen aan de hand van relevante literatuur.

Toelichting NR2: Indien de uiteindelijke rapportageschaal afwijkt van de latente schaal waarop de cesuren zijn vastgesteld dient de toetsaanbieder de nauwkeurigheid van de gekozen normeringsmethode ten opzichte van een 2PL-vaardigheidsschatter en de landelijk gehanteerde cesuurpunten van de toetsadviezen en referentieniveaus via een tabel weer te geven (verwarringsmatrix) en verantwoordt mogelijk verschillen. Ook dient de toetsaanbieder deze te duiden door middel van bijvoorbeeld de plus-minus 1 niveau-index (Pilliner 1969) of de marginal classification accuracy. De

aanbieder legt uit waarom voor een bepaalde statistische methode en de gehanteerde grenzen is gekozen.

5.4 Gegevenslevering

Deze kwaliteitseisen zijn alleen van toepassing voor al eerder afgenomen toetsen. De beoordeling van de kwaliteit en het functioneren van de calamiteitentoets vindt alleen plaats nadat de operationele toets daadwerkelijk is ingezet. Het CvTE beslist of de calamiteitentoets meedoet aan het Toelatings- en doorstroomonderzoek.

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
GL1	Heeft de toetsaanbieder na de afname van de vorige doorstroomtoets alle benodigde gegevens verstrekt ten behoeve van de landelijke normering?	ja/nee/n.v.t.
GL2	Heeft de toetsaanbieder na de afname van de vorige doorstroomtoets alle benodigde gegevens verstrekt ten behoeve van toelatings- en doorstroomonderzoek?	ja/nee/n.v.t.

Toelichting GL1: De toetsaanbieder heeft na de afname van de vorige doorstroomtoets de (persoons) gegevens verstrekt ten behoeve van de landelijke normering zoals uiteengezet in de Regeling Beoordelingsnormen doorstroomtoetsen po. Tenminste zijn aangeleverd:

- Leerlinggegevens
 - leerlingidentificatiecode
 - populatie specificering
 - schooltype
 - voorlopig schooladvies
 - toetsadvies
- Scoringsgegevens
 - leerlingidentificatiecode
 - item label
 - item score
- Representativiteitsgegevens (aantal leerlingen met verschillende schooladviezen);
- Controlegegevens
 - item label
 - naam (sub)domein
 - aantal observaties
 - p-waarde van het item
- RIO-codes
- Optioneel: lijst van no-score items

Toelichting GL2: De toetsaanbieder heeft na de afname van de vorige doorstroomtoets de (persoons) gegevens verstrekt ten behoeve van toelatings- en doorstroomonderzoek zoals uiteengezet in de Regeling Beoordelingsnormen doorstroomtoetsen po. Tenminste zijn aangeleverd per individuele leerling:

- Pseudoniem
- Toetsscore
- Toetssoort
- Toetsadvies
- Afnamemodus van de doorstroomtoets
- Kalenderjaar van de doorstroomtoets
- Referentieniveaus
- Score op de gewogen latente vaardigheidsschaal
- Scores op de latente vaardigheidsschalen van de losse (sub)domeinen

5.5 Kwaliteit en functioneren van de doorstroomtoets (rapportage achteraf)

De beoordeling van de kwaliteit en het functioneren van de calamiteitentoets vindt alleen plaats nadat de operationele toets daadwerkelijk is ingezet. Het CvTE beslist of de calamiteitentoets meedoet aan het Toelatings- en doorstroomonderzoek.

Kwaliteitseisen

Beslisregel: Om aan deze categorie kwaliteitseisen te voldoen, dienen alle vragen met JA of N.V.T. te worden beantwoord.

Code	Vragen	Mogelijke antwoorden
KF1	Verantwoordt de toetsaanbieder op een gedegen manier of en waarom opgaven met inhoudelijke fouten, psychometrisch disfunctioneren, DIF of te weinig waarnemingen al dan niet zijn meegenomen voor de scorerapportage?	ja/nee/n.v.t.
KF2	Zijn verschillen in percentage referentieniveaus voldoende verantwoord?	ja/nee/n.v.t.
KF3	Zijn verschillen in de verdeling van populatievaardigheden ten opzichte van de verantwoording van het afgelopen jaar en de onderdelen onderling voldoende verantwoord?	ja/nee/n.v.t.
KF4	Zijn analyses met betrekking tot verschillen in parameters van de meest recente operationele afname en eerdere afnames aangeleverd en voldoende onderbouwd?	ja/nee/n.v.t.
KF5	Worden afwijkingen van het landelijke gemiddelde in de relatie tussen toets- en schooladviezen voldoende in kaart gebracht en waar mogelijk voldoende verantwoord?	ja/nee/n.v.t.

Deze kwaliteitseisen zijn alleen van toepassing voor al eerder afgenomen toetsen. De toetsaanbieder verantwoordt jaarlijks na de toetsafname-periode de kwaliteit en het functioneren van de afgenomen doorstroomtoets op basis van de boven genoemde vijf kwaliteitscriteria. Er wordt beoordeeld of de inhoud van de informatie die door de toetsaanbieder is aangeleverd voldoet aan de vooraf gestelde eisen. De resultaten en de kwaliteit van de informatie van het voorafgaande jaar worden gebruikt bij de beoordeling van de doorstroomtoets, dan wel de jaarlijkse her-evaluatie. Als er een nieuwe aanbieder is die dit nog niet op kan leveren naar aanleiding van een vorige afname, spelen deze eisen bij de eerste evaluatie geen rol, maar heeft het wel een groter gewicht bij de jaarlijkse her-evaluatie die volgt na de eerste afname.

Toelichting KF1: De toetsaanbieder dient informatie te verstrekken over de aantallen opgaven per vaardigheid: aantal wel en aantal niet meegenomen voor de scorerapportage naar de leerlingen wegens (1) een inhoudelijke fout, (2) psychometrisch disfunctioneren, (3) DIF (bijvoorbeeld tussen afnames), (4) te weinig waarnemingen (in tabelformaat). Wanneer er zaai-items ingezet worden kunnen deze in een aparte tabel weergegeven worden. Het leveren van een lijst welk item dit betreft wordt aangeraden. De aanbieder wordt gevraagd om de resultaten voldoende verantwoord.

Toelichting KF2: De toetsaanbieder dient informatie te verstrekken over de percentages referentieniveaus van het wettelijk verplichte domein Lezen en subdomein Taalverzorging van het terrein Nederlandse taal en het terrein Rekenen plus een overzicht van de percentages toetsadviezen (de zes categorieën) van dit jaar ten opzichte van het vorige jaar (in tabelformaat). Eventuele verschillen tussen de beide jaren worden door de toetsaanbieder verantwoord.

Toelichting KF3: De toetsaanbieder dient informatie te verstrekken over de verdeling van de populatievaardigheid (gemiddelde en standaardafwijking) van in ieder geval het regulier basisonderwijs inclusief de schattingsfout (standard error) op de vaardigheidsverdeling, op dezelfde schaal als de itemparameters. De verdelingen worden door middel van histogrammen inzichtelijk gemaakt. In deze informatie wordt onderscheid gemaakt tussen de verschillende onderwijstypen. Eventuele verschillen ten opzichte van de verantwoording van het afgelopen jaar en de landelijke normering in het afnamejaar worden door de toetsaanbieder verantwoord. Verschillen tussen de onderdelen worden door de toetsaanbieder verklaard.

Toelichting KF4: Als opgaven eerder zijn afgenomen (bijvoorbeeld in een voorgaand jaar of tijdens een pretest), dient de toetsaanbieder informatie te verstrekken over item-bias met betrekking tot de parameters van de meest recente operationele afname in vergelijking met eerdere afnames, bijvoorbeeld door middel van DIF-analyses.

Toelichting KF5: De toetsaanbieder dient informatie te verstrekken over de relatie tussen de toetsadviezen en de eerder gegeven (voorlopige) schooladviezen voor zover deze bekend zijn (in tabelformaat). Als de relatie tussen toets- en schooladviezen van de aanbieder zelf sterk afwijkt van de landelijke gemiddelde verdeling tussen toets- en schooladviezen, dient de toetsaanbieder mogelijke redenen van de afwijking in kaart te brengen. Deze redenen dienen in ieder geval (maar niet uitsluitend) te worden gezocht op basis van specifieke aspecten van de eigen populatie en de gebruikte toets (zoals afnamevorm, algoritme etc.), en moeten waar mogelijk voldoende worden verantwoord.

6. Vierjaarlijkse erkenning en jaarlijkse vaststelling door het CvTE

6.1. De vierjaarlijkse erkenning

Een adviseur beoordeelt vierjaarlijks in opdracht van het CvTE de onderwijskundige, organisatorische en psychometrische aspecten van de doorstroomtoets. Na de beoordeling stelt een adviseur een advies op. Dit advies gaat naar het CvTE dat, op basis van het ontvangen advies, een definitieve beslissing neemt over het al dan niet erkennen van de betreffende doorstroomtoets. Wanneer de toets is erkend kan deze toets vier achtereenvolgende jaren worden ingezet. Onderdeel van deze uitgebreide beoordeling is een psychometrische controle van de pretestprocedure én van de operationele testprocedure en -resultaten die jaarlijks worden verzameld in de afnameperiode.

Na de eerste toelating van vier jaar, dienen de toetsaanbieders opnieuw een toelatingsaanvraag in bij het CvTE. Voor deze nieuwe toelatingsaanvraag vindt een herbeoordeling plaats. Voor deze herbeoordeling van de onderwijskundige inhoud, organisatorische aspecten en psychometrische aspecten doet het CvTE eveneens een beroep op onafhankelijke onderwijskundige en vakinhoudelijke experts van een adviseur.

6.2. De jaarlijkse vaststelling

Een adviseur beoordeelt jaarlijks aan de hand van ditzelfde beoordelingskader de kwaliteit van de doorstroomtoets. De aanbieder levert hiervoor bij het CvTE alle gevraagde documenten aan die eveneens voor de vierjaarlijkse erkenning ingediend dienen te worden. De aanbieder verstrekt hierbij een leeswijzer. Hierin staat per document aangegeven op welke punten deze afwijkend is van de eerder verstuurde versie ten tijde van de vierjaarlijkse check of jaarlijkse vaststelling. Vervolgens stelt het CvTE jaarlijks vast of de erkende doorstroomtoets nog voldoet aan de criteria van het beoordelingskader. Dit doet het CvTE met de resultaten van de door een adviseur uitgevoerde jaarlijkse vaststelling. Onderdeel van de jaarlijkse vaststelling is de controle achteraf of de toetssamenstellingsprocedure door alle toetsaanbieders correct en consistent was uitgevoerd.

Voor de calamiteitentoets wordt op een aantal punten van het beoordelingskader afgeweken.

Model leeswijzer

Verantwoording Doorstroomtoets

Kwaliteitseis		Niet van toepassing	Beschreven in document: (naam document)	Hoofdstuk + paragraaf	Evt. opmerkingen ¹
Paragraaf kader	Code				
3.2	V1				
	V2				
	V3				
	V4				
3.3	H1				
3.4.1.1	L1.1				
	L1.2				
	L1.3				
	L2				
	L3.1				
	L3.2				
	L3.3				
	L3.4				
	L3.5				
	L3.6				
	L3.7				
	L3.7a				
	L4				
	L5				
	L6.1				
	L6.2				
	L6.3				
	L6.4				



Kwaliteitseis		Niet van toepassing	Beschreven in document: (naam document)	Hoofdstuk + paragraaf	Evt. opmerkingen ¹
Paragraaf kader	Code				
3.4.2	L6.4a				
	TV1				
	TV1.a				
	TV1.b				
	TV1.c				
	TV1.d				
	TV1.e				
3.5.1	TV2				
	TV3				
	D.G				
	D.Vh				
	D.M				
	D.Vb				
	D1				
3.5.2	O.A				
	O.B				
	O.C				
	O.BC				
	O.1				
3.5.3	C1				
	C2				
	C3				
	C4				
	C5				
3.5.4	K1				
	K2				
3.5.5	R1				
3.6	TN1				
4.1	LR1				
	LR2				
	LR3				
	LR4				
	LR5				
	LR6				
	LR7				
4.2	LL1				
	LL1.a				
	LL1.b				
	LL1.c				
	LL1.d				
	LL1.e				
	LL1.f				
	LL1.g				
	LL2				
	LL2.a				
	LL2.b				
	LL2.c				
	LL2.d				
4.3	A1				
	A2.1				
	A2.2				
	A2.3				
	A2.4				
	A2.5				
	A2.6				
	A2.7				



Kwaliteitseis		Niet van toepassing	Beschreven in document: (naam document)	Hoofdstuk + paragraaf	Evt. opmerkingen ¹
Paragraaf kader	Code				
	A3				
4.4	RB1				
	RB1.1				
	RB1.2				
	RB1.3				
	RB1.4				
	RB1.5				
	RB1.6				
	RB1.7				
	RB2				
	RB3				
	RB4				
	RB4.1				
	RB4.2				
	RB5				
5.2.1	N1.1				
	N1.1a				
	N1.1b				
	N1.2				
	N1.3				
5.2.2	KA1				
	KA2				
	KA3				
	KA4				
	KA5				
	KA6				
	KA7				
5.2.3	T1				
	T2				
	T3				
	T4				
	T5				
5.3	NR1				
	NR2				
5.4	GL1				
	GL2				
5.5	KF1				
	KF2				
	KF3				
	KF4				
	KF5				

¹ Denk bijvoorbeeld aan paginanummers. Indien het document geen paginanummers bevat, kan de aanbieder alsnog specificeren waar de informatie gevonden kan worden. Voorbeeld: bij een Excel document aangeven welke rijen en kolommen het betreft.

Optionele domeinen					
Kwaliteitseis		Beschreven in document: (naam document)	Hoofdstuk + paragraaf	Evt. opmerkingen ¹	
3.4.3	S1.1				
	S1.2				
	S1.3				
	S2.1				
	S2.2				
	S2.3				
	S2.4				
	S3.1				



Optionele domeinen				
Kwaliteitseis		Beschreven in document: (naam document)	Hoofdstuk + paragraaf	Evt. opmerkingen ¹
	S3.2			
	S3.3			
	S4.1			
	S4.2			
	S4.3			
	S4.4			
3.4.4	M1.1			
	M1.2			
	M2.1			
	M2.2			
	M2.3			
	M2.4			
	M3			
	M4.1			
	M4.2			
	M4.3			
	M4.4			
	M5.1			
	M5.2			
	M6.1			
	M6.2			
	M6.3			
	M6.4			
	M7.1			
	M7.2			
	M7.3			
	M8.1			
	M8.2			
	M8.3			
	M9			
	M10.1			
	M10.2			
	M10.3			
	M10.4			
3.4.5	B1			

¹ Denk bijvoorbeeld aan paginanummers. Indien het document geen paginanummers bevat, kan de aanbieder alsnog specificeren waar de informatie gevonden kan worden. Voorbeeld: bij een Excel document aangeven welke rijen en kolommen het betreft.

Checklist 1 Toetsopgaven Nederlandse taal per (sub)domein en per onderdeel: formulier voor toetsontwikkelaar en expert

3.4.1.1. Subdomein Zakelijke Teksten

Taken

Code		Tekst(en)
L1.1	Informatieve tekst(en)	
L1.2	Instructieve tekst(en)	
L1.3	Betogende tekst(en)	

Tekstkenmerken

Code		Mogelijke antwoorden	
L2.1	De tekstkenmerken van de teksten.		
Criteria waarop gescoord kan worden ¹		Onvoldoende	Voldoende
Inhoud	Er wordt niet teveel nieuwe informatie gelijktijdig geïntroduceerd.		



Code	Mogelijke antwoorden		
Structuur	De structuur van de teksten is eenvoudig en helder.		
Taalgebruik	De teksten hebben een lage informatiedichtheid.		
	De teksten bestaan voornamelijk uit frequent gebruikte woorden.		
	De teksten zijn niet te lang		
Bedoeling	De informatie in de teksten is herkenbaar geordend; verbanden in de tekst zijn duidelijk aangegeven.		

¹ Globaal oordeel gebaseerd op het totale aantal teksten in de toets. Deze criteria zijn bedoeld als hulpmiddel om tot een eindoordeel van de tekstenmerken van alle in de toets opgenomen teksten te komen.

Kenmerken van de taakuitvoering

Code	Opgaven	
L3.1	Opdrachten waarin kenmerk van de taakuitvoering: techniek en woordenschat wordt getoetst.	
L3.2	Opdrachten waarin kenmerk van de taakuitvoering: begrijpen wordt getoetst.	
L3.3	Opdrachten waarin kenmerk van de taakuitvoering: interpreteren wordt getoetst.	
L3.4	Opdrachten waarin kenmerk van de taakuitvoering: samenvatten wordt getoetst.	
L3.5	Opdrachten waarin kenmerk van de taakuitvoering: opzoeken wordt getoetst.	
L3.6	Opdrachten waarin kenmerk van de taakuitvoering: evalueren wordt getoetst (optioneel).	

3.4.1.2. Subdomein Fictionele, narratieve en literaire teksten

Taken

Code	Tekst(en)	
L4.1	Fictionele, narratieve en literaire tekst(en)	

Tekstkenmerken

Code	Mogelijke antwoorden		
L5.1	De tekstkenmerken van de teksten.		
	<i>Criteria waarop gescoord kan worden</i> ¹	Onvoldoende	Voldoende
Inhoud	Poëzie en liedjes hebben meestal een verhalende inhoud en een emotionele lading.		
Structuur	De structuur van de teksten is eenvoudig en helder.		
Taalgebruik	Het tempo waarin spannende of dramatische gebeurtenissen elkaar opvolgen is hoog.		
Bedoeling	Het verhaal heeft een dramatische verhaallijn waarin de spanning af en toe wordt onderbroken door gedachten of beschrijvingen.		

¹ Globaal oordeel gebaseerd op het totale aantal teksten in de toets. Deze criteria zijn bedoeld als hulpmiddel om tot een eindoordeel van de tekstkenmerken van alle in de toets opgenomen teksten te komen.

Kenmerken van de taakuitvoering

Code	Opgaven	
L6.1	Opdrachten waarin kenmerk van de taakuitvoering: begrijpen wordt getoetst.	
L6.2	Opdrachten waarin kenmerk van de taakuitvoering: interpreteren wordt getoetst.	
L6.3	Opdrachten waarin kenmerk van de taakuitvoering: evalueren wordt getoetst (optioneel).	

3.4.2. Subdomein Taalverzorging

Kenmerken van de taakuitvoering

Code		Opgaven
TV1	Opdrachten waarin verschillende categorieën van spellingsproblemen en -regels worden getoetst.	
TV1.a	Opdrachten waarin regels voor lettergreepgrenzen worden getoetst.	
TV1.b	Opdrachten waarin morfologische spelling wordt getoetst.	
TV1.c	Opdrachten waarin regels voor werkwoordspelling worden getoetst.	
TV1.d	Opdrachten waarin de overige regels worden getoetst.	
TV1.e	Opdrachten waarin het gebruik van interpunctie (leestekens) wordt getoetst.	

Toetsopgaven

Evalueer de kwaliteit van alle toetsopgaven van het wettelijk verplichte domein Lezen en subdomein Taalverzorging (inclusief Dictée) van het terrein Nederlandse taal, en de optionele domeinen Schrijven, Mondelinge taalvaardigheid, en subdomein Begrippenlijst van het terrein Nederlandse taal aan de hand van de kwaliteitseisen en constructievoorschriften uit de Checklist voor het beoordelen van de kwaliteit van observatie-categorieën en toetsopgaven.

Checklist 2 Toetsopgaven Rekenen per domein en per onderdeel: formulier voor toetsontwikkelaar en expert

	Code	A Notatie, taal en betekenis	Code	B Met elkaar in verband brengen	Code	C Gebruiken
Getallen (g)	g.A.	Uitspraak, notatie en betekenis van gehele getallen, breuken en decimale getallen Uitspraak, notatie en betekenis van symbolen en relaties Uitspraak, taal en betekenis van wiskundetaal binnen het domein getallen	g.B.	Getallen en getalsrelaties Structuur en samenhang zien binnen het domein getallen	g.C.	Memoriseren, automatiseren Hoofdrekenen (notaties toegestaan) Hoofdbewerkingen (+, -, x, :) op papier uitvoeren met hele getallen en decimale getallen Bewerkingen met breuken (+, -, x, :) op papier uitvoeren Berekeningen uitvoeren om problemen op te lossen met gehele getallen, breuken en decimale getallen Rekenmachine op een verstandige manier inzetten

Toetsopgave

Verhoudingen (vh)	vh.A.	Uitspraak, notatie en betekenis van breuken, procenten en verhoudingen, Uitspraak, notatie en betekenis van symbolen en relaties Uitspraak, taal en betekenis van wiskundetaal binnen het domein verhoudingen	vh.B.	Verhouding, procent, breuk, decimaal getal, deling, 'deel van' met elkaar in verband brengen	vh.C.	In de context van verhoudingen berekeningen uitvoeren, ook met procenten en verhoudingen
--------------------------	-------	---	-------	--	-------	--

Toetsopgave

Metten & Meetkunde (m /mk)	m/mk.A	Uitspraak, notatie en betekenis van maten voor lengte, oppervlakte, inhoud en gewicht, temperatuur, tijd en geld, Uitspraak, notatie en betekenis van meetkundige symbolen en relaties Meetinstrumenten aflezen	m/mk.B	Meetinstrumenten gebruiken Structuur en samenhang tussen maateenheden Verschillende meetkundige representaties	m/mk.C	Metten en rekenen met maten Rekenen in de meetkunde
---------------------------------------	--------	---	--------	--	--------	--

Toetsopgave

	Code	A Notatie, taal en betekenis	Code	B Met elkaar in verband brengen	Code	C Gebruiken
Verbanden (vb)	vb.A.	Analyseren en interpreteren van informatie uit tabellen, grafische voorstellingen en beschrijvingen Kennis van diagrammen, tabellen en grafieken	vb.B.	Verschillende voorstellingsvormen met elkaar in verband brengen Gegevens verzamelen, ordenen en weergeven Patronen beschrijven	vb.C.	Tabellen, diagrammen en grafieken gebruiken bij het oplossen van problemen Rekenvaardigheden gebruiken binnen het domein verbanden

Toetsopgave

Checklist 3 Verdeling van toetsopgaven Rekenen over domeinen en over onderdelen volgens de Toetswijzer PO

	A Notatie, taal en betekenis [geen minimum]	B Met elkaar in verband brengen [ten minste 20% van de gehele toets]	C Gebruiken [ten minste 30% van de gehele toets]
Getallen 30% – 40% van gehele toets	Uitspraak, notatie en betekenis van gehele getallen, breuken en decimale getallen, van symbolen en relaties Uitspraak, taal en betekenis van wiskundetaal binnen het domein getallen	Getallen en getsrelaties Structuur en samenhang	Memoriseren, automatiseren Hoofdrekenen (notaties toegestaan) Hoofdbewerkingen (+, -, x, :) op papier uitvoeren met hele getallen en decimale getallen Bewerkingen met breuken (+, -, x, :) op papier uitvoeren Berekeningen uitvoeren om problemen op te lossen met gehele getallen, breuken en decimale getallen Rekenmachine op een verstandige manier inzetten
Verhoudingen 20% – 30% van gehele toets	Uitspraak, notatie en betekenis van breuken, procenten en verhoudingen, Uitspraak, notatie en betekenis van symbolen en relaties Uitspraak, taal en betekenis van wiskundetaal binnen het domein verhoudingen	Verhouding, procent, breuk, decimaal getal, deling, 'deel van' met elkaar in verband brengen	In de context van verhoudingen berekeningen uitvoeren, ook met procenten en verhoudingen
Meten & Meetkunde 20% – 30% van gehele toets	Uitspraak, notatie en betekenis van maten voor lengte, oppervlakte, inhoud en gewicht, temperatuur, tijd en geld Uitspraak, notatie en betekenis van meetkundige symbolen en relaties Notatie, taal en betekenis van meetinstrumenten	Meetinstrumenten gebruiken Structuur en samenhang tussen maateenheden Verschillende meetkundige representaties	Metten en rekenen met maten Rekenen in de meetkunde
Verbanden 15% – 20% van totaal aantal toetsopgaven	Analyseren en interpreteren van informatie uit tabellen, grafische voorstellingen en beschrijvingen Kennis van diagrammen, tabellen en grafieken	Verschillende voorstellingsvormen met elkaar in verband brengen Gegevens verzamelen, ordenen en weergeven Patronen beschrijven	Tabellen, diagrammen en grafieken gebruiken bij het oplossen van problemen Rekenvaardigheden gebruiken binnen het domein verbanden



TOELICHTING

Op 1 januari 2023 trad de Regeling beoordelingskader doorstroomtoets PO in werking. Deze Regeling is op 1 maart 2024 gewijzigd na ervaringen met de eerste beoordelingsperiode uit de zomer van 2023. Nu wordt de Regeling opnieuw gewijzigd door ervaringen met de tweede beoordelingsperiode uit de zomer van 2024.

De meest ingrijpende wijzigingen hebben betrekking op de wijze waarop toetsaanbieders items en voorbeeldtoetspaden moeten aanleveren aan het College voor toetsen en examens. Daarnaast zijn enkele kwaliteitseisen uitgewerkt die betrekking hebben op de psychometrie en ondersteuningsbehoeften van leerlingen.

De wijzigingen van het beoordelingskader zijn van zodanige omvang, dat ervoor gekozen is om de bijlage te vervangen.

*Het College voor toetsen en examens,
de voorzitter,
J.H. van der Vegt*