

**Evaluatie examens Nederlands
vmbo, havo en vwo 2018
na overleg met het College
voor Toetsen en Examens**

**LEVENDE
Talen**

NEDERLANDS

2018

INHOUDSOPGAVE

1.	ALGEMEEN	2
1.1	Teksten en vragen.....	2
1.2	Antwoordmodel.....	3
1.3	De N-term	4
1.4	Mogelijke aantrekkelijk invalshoeken voor examenvragen	4
1.5	Communicatie	5
2.	VMBO-EXAMEN NEDERLANDS 2018.....	6
2.1	Praktisch.....	6
2.2	Inhoudelijk	6
3.	HAVO-EXAMEN NEDERLANDS 2018	7
3.1	Praktisch.....	7
3.2	Inhoudelijk	8
3.2.1	Vraagstelling	8
3.2.2	Andere kwesties.....	8
4.	VWO-EXAMEN NEDERLANDS 2018.....	8
4.1	Praktisch.....	8
4.2	Inhoudelijk	9
4.2.1	Hoofdgedachtevragen	9
4.2.2	Verschillende interpretaties van woorden en begrippen.....	9
4.2.3	Meerkeuzevragen	10

Deze evaluatie is opgesteld op basis van de ervaringen van leraren (kenbaar gemaakt via facebook en docentbijeenkomsten en analyse van de kwaliteiten van de toets (Toets- en itemanalyse) door leden van de werkgroep Examens van LT Nederlands. Na de evaluatie van de drie examens bieden we een aantal opties voor toekomstige examenvragen, die aansluiten op het Advies Examen Nederlands dat SBN met Nederlands Nu samenstelde. Op basis van het gesprek met CvTE/Cito op 19-11-2018 is de tekst van de evaluatie op enkele punten verduidelijkt en aangepast.

1. ALGEMEEN

1.1 Teksten en vragen

- 1) Voor alle examens gold in 2018 dat de sets open vragen en de set Analyse en Interpreteren de beste voorspellers waren van het examenresultaat. Vraag is of er niet nog meer op deze vraagvorm en deze relevante vaardigheid ingezet zou kunnen worden? Uit de toets- en itemanalyses blijkt dat de ca. 20 vragen in de categorie Analyse & Interpretatie even betrouwbaar is als de gehele toets van bijna 40 vragen. Zie ook punt 2.
- 2) Op enkele bijzonderheden na werden de teksten en de vragen gewaardeerd; de psychometrische kwaliteit was in het algemeen op vraagniveau in orde. Blijft het niet problematisch dat de betrouwbaarheid van de hele toets toch voor verbetering vatbaar blijft? Is de grens bereikt qua verbetering omdat er zo'n conglomeraat aan vaardigheden getoetst wordt/moet worden vanwege de eindtermen? In dat geval zou Levende Talen Nederlands wel een lans willen breken voor een toetsmatrijs die uitsluitend op tekstbegrip is toegespitst, in het bijzonder op de categorie Analyse & interpretatie, in overeenstemming met het Advies Examen Nederlands dat het SBN met Nederlands Nu uitbracht zomer 2018.
- 3) De overkoepelende vragen worden gewaardeerd, en lijken psychometrisch een voltreffer.

Uit de reactie CvTE blijkt dat de examenmakers gehouden zijn aan de toetsmatrijs zoals die nu is opgesteld, waardoor het examen een mix van vaardigheden toetst, en het examen als geheel niet veel betrouwbaarder kan worden dan nu het geval is (ca. .60). Als Levende Talen wil inzetten op een sterkere nadruk op Analyse en Interpretatie, of op een CSE Nederlands dat uitsluitend tekstbegrip toetst, is overleg met het Ministerie de aangewezen weg.

1.2 Antwoordmodel

Het antwoordmodel blijft een punt van zorg, ongemak en wrevel bij docenten. Heel veel punten worden door collega's ingebracht wegens onzekerheid over waar de grenzen liggen: wat mag nog wel en niet meer goed gerekend worden. Dit wordt bij collega's onder meer veroorzaakt door (in willekeurige volgorde) (1) beperkte examenervaring, (2) beperkt overleg op school (maar één docent op de examenklas(sen) van een schooltype of weinig/geen overleg binnen sectie door allerlei oorzaken) en (3) ervaringen met eerste en tweede correctoren in voorgaande jaren. Hoe kunnen we hier aan werken? (Cursus examen nakijken? Gezamenlijk nakijken in de regio? Collega's zijn op zoek naar zekerheid, vaak ingegeven door nog beperkte ervaring of door ervaringen uit het verleden. We zien wel mogelijkheden om daarin verbetering aan te brengen.

- 4) Misschien moet het correctiemodel niet zozeer *vormen* van goede antwoorden bieden, die als uitwerking gepresenteerd worden van de vaste aanloop 'een kern van een goed antwoord bevat ...'. Zo'n aanbod leidt er wellicht toe dat de beoordeling dan vooral een zoektocht is naar de mate waarin antwoorden van kandidaten overeenkomen met een of meer voorbeelden van goede antwoorden. Misschien moet de startzin niet zijn: De kern van een goed antwoord *bevat*, maar eerder 'uit een antwoord moet blijken dat ...'? Als het antwoordmodel inderdaad een *omschrijving* van de kern van een goed antwoord zou bieden, is het misschien zo dat men minder een op een formuleringen gaat vergelijken. Het zou interessant zijn om te onderzoeken wat het effect van de formulering van het antwoordmodel is op het beoordelingsproces en de opvattingen over beoordelen van docent-beoordelaars. De problematiek neemt denkkelijk toe met het niveau van de vraag: hoe opener de vraag, hoe complexer het antwoordproces en het beoordelingsproces. Een mogelijkheid om tekstschalen te benutten bij het beoordelen van antwoorden op open vragen waarin niet zozeer de dichotomie van goed/fout overheerst, maar de relatieve kwaliteit (zwakker/sterker) zou wellicht ook kunnen helpen.

Het CvTE zal een onderzoeksproject entameren (Van den Bergh, Schuurs & SBN) om na te gaan of een andere verwoording van de kern van een goed antwoord inderdaad mogelijk is, en in hoeverre die leidt tot andere beoordelingsprocessen dan de huidige formulering. Wellicht kan op kleine schaal het antwoordmodel in deze richting opschuiven (voor het eerst in het tweede tijdvak 2020).

- 5) Levende Talen Nederlands zal proberen het verslag n.a.v. de docentbijeenkomst anders, en meer in de richting van bovenstaande inhoud en vorm te geven, en proberen te vermijden dat het een opsomming van antwoorden is, en gaat dienen als een tweede correctiemodel.

- 6) Levende Talen Nederlands denkt graag mee of we samenkomsten van docenten kunnen organiseren waarin minder ervaren docenten getraind worden in het beoordelen van antwoorden op open vragen, en beoordelingsessies voor docenten die weinig steun vinden in hun sectie.
- 7) Levende Talen Nederlands heeft een plan opgesteld om de docentbijeenkomsten n.a.v. de examens een ander karakter te geven en legt dat plan graag voor aan de medewerkers van het CvTE.

1.3 De N-term

- 8) Levende Talen Nederlands zou graag weten of het mogelijk gaat zijn dat leerlingen voor het SCE Nederlands een (erg) hoog cijfer kunnen krijgen. Er is een voorbeeld van een VMBO leerling die een meerkeuzevraag fout had, twee komma's vergat, en in de samenvattingsopdracht een naam niet opnam. Op basis van de N-term kwam deze leerling niet hoger dan een 8,8.

Er is binnen het CvTE een discussie gaande over de N-term. Een van de voorstellen is om de N-term te baseren op de range van behaalde scores, en niet op de range van potentiële scores. Als de maximale score 62 is, wordt de N-term daarop gebaseerd. Die zou ook gebaseerd kunnen worden op de hoogste score behaald in een cohort (58): een score van 58 betekent dan; een 10 op de lijst. Levende Talen Nederlands brengt deze mogelijkheid onder de aandacht van het Ministerie van OC&W.

1.4 Mogelijke aantrekkelijk invalshoeken voor examenvragen

Op basis van het VWO-examen leggen we een paar mogelijkheden voor die ons inziens passen bij het Advies Examens Nederlands dat wij zomer 2018 samenstelden.

- 1) Bij de teksten over nut en noodzaak van filosofie zou feitelijke informatie over het aantal universiteiten met een studie filosofie, het aantal masterstudenten, arbeidsmarktgegevens van filosofen etc. goede combinatiemogelijkheden kunnen bieden.
- 2) Een optie kan zijn de kandidaat te laten formuleren welke gegevens ze nodig zouden hebben om het een en ander te beoordelen.
- 3) Een tekst als die van Hertzberger biedt mogelijkheden voor stijl- en retorica-analyse.
- 4) Vraag 21 biedt een mogelijkheid tot een syntheseopdracht.

Het CvTE ziet mogelijkheden om de eerste optie te exploreren in komende examens (niet: 2019, dat examen ligt vast). Met optie 3 zijn al verkenningen gedaan, maar men heeft nog geen item van dien aard kunnen arresteren in de vaststellingscommissie: er kleven steeds bezwaren aan. Het opnemen van een syntheseopdracht als blijk van begrip van de inhoudelijke contrasten en

overeenkomsten tussen twee bronnen, vindt het CvTE voorsnog niet wenselijk, omdat de formuleervaardigheid dan geïntroduceerd wordt in de meting van tekstbegrip. Het standpunt van Levende Talen Nederlands (zie Advies Examen Nederlands, 2018) is dat formuleervaardigheid niet losgezien zou moeten worden van de meting van tekstbegrip: formuleervaardigheid speelt nu in de meting van open vragen ook een rol. Het zou in Nederland niet toegestaan moeten worden dat leerlingen op het hoogste niveau van het voortgezet onderwijs een eindexamen mogen afleggen zonder zich te bekommeren om de kwaliteit van de formulering van het antwoord. De discussie hierover wordt voortgezet.

1.5 Communicatie

- 1) Dit jaar waren er door omstandigheden geen collega's van CvTE en Cito aanwezig op de havo-besprekingen, terwijl we juist de afgelopen jaren positieve ervaringen opdeden met die aanwezigheid, wegens de korte lijntjes. De vragen die dit jaar rezen tijdens de docentbijeenkomsten, werden evenwel per mail gedeeld en snel beantwoord.
- 2) Het is in de praktijk gebleken dat niet alle collega's op de hoogte waren van de aanvullingen. Wellicht dat het persoonlijk account dat iedereen nu kan krijgen via aanmelding van de examencoördinator hier voor verbetering zorgt.
- 3) Niet iedereen blijkt op de hoogte van het bestand waarin de aanwijzingen staan rondom de interpretatie van de regels rondom taalfouten (link opnemen in het cv?). Aan dit soort nakijkaanwijzingen is behoefte bij de collega's: duidelijk heeft men om welke reden dan ook deze bevestiging nodig. Wellicht kunnen er toevoegingen op dit gebied komen in de vakspecifieke afspraken in het cv. Daarmee worden bepaalde discussies tussen eerste en tweede corrector voorkomen.

Het CvTE zal nagaan of het plaatsen van zo'n link naar het bestand mogelijk is.

- 4) Er was wrevel na de VMBO-bijeenkomst bij de voorzitter van de docentbijeenkomst over de beslissing van het CvTE om de discrepantie tussen de inhoud van de situatieschets en de opdracht die leerlingen zou hebben kunnen verwarren niet als fout in het examen te erkennen.

Het CvTE neemt contact op met de Voorzitter van die bijeenkomst voor tekst en uitleg.

2. VMBO-EXAMEN NEDERLANDS 2018

- 1) Het examen was voor de leerlingen die werken op een laptop / computer redelijk goed te maken binnen de tijd. Leerlingen die de opdrachten moesten uitschrijven, hebben moeite gehad alles binnen de tijd af te krijgen.
- 2) De teksten waren passend en interessant.
- 3) Fijn dat nu gekozen is voor het schrijven van een zakelijke e-mail.
- 4) Met een lage n-term van 0,3 is het bijna onmogelijk een 9 of hoger te scoren voor dit examen. Dergelijke cijfers komen bij andere vakken nog wel eens voor. Waarom is de n-term toch altijd zo laag?
- 5) De uitwerkbijlage is heel praktisch en duidelijk. Er is voldoende ruimte voor het formuleren van de antwoorden. Ook is duidelijk zichtbaar wanneer de leerlingen meerdere redenen / antwoorden moeten geven. Daarover kan geen enkele discussie ontstaan bij correctoren.

2.1 *Praktisch*

- 1) Voor veel leerlingen zou het een uitkomst zijn als alle examenboekjes zouden worden voorzien van nietjes. Op het tafeltje liggen: drie boekjes met losse bladen, een woordenboek, schrijfgerei, kladblaadjes, een glas drinken.... etc. Dat is te chaotisch voor sommige leerlingen.
- 2) Citeer de zin (vr. 18, tijdvak 2). Volgens het CV mag ook een zinsgedeelte goed worden gerekend. Dat levert verwarring op.
- 3) Schrijfopdracht, tijdvak 2. In het uitwerkingsboekje staat: Beste meneer Knol. Deze aanhef heeft bij veel correctoren verbazing opgeroepen. Wij leren de leerlingen dat immers niet aan.

2.2 *Inhoudelijk*

Het antwoordmodel voor de schrijfopdracht leverde zoals altijd veel discussies op.

- 1) Inhoud: in hoeverre mag de leerling een eigen draai geven aan de inhoud? Is het verplicht dat de leerling zich houdt aan de informatie die hij krijgt op bladzijde 8 of mag eigen inbreng / creativiteit ook goed gerekend worden? De informatie op bladzijde 8 en 9 correspondeert niet op alle punten met elkaar. Er is hier sprake van een fout. Helaas is deze fout niet hersteld.
- 2) Taalgebruik: de verdeling in formuleringsfouten en spelfouten is niet voor alle correctoren heel helder. Misschien is een verdeling in taalfouten en interpunctiefouten het overwegen waard? Daarnaast hebben de correctoren behoefte aan duidelijke regels. Voorbeeld: is een komma voor voegwoorden verplicht of niet? Waar horen witregels? Moeten getallen tot en met tien voluit worden geschreven of tot en met twintig? Etc.

Het CvTE zal bekijken of een tweedeling (interpunctie-taalfouten) het probleem kan oplossen en geen nieuwe problemen za introduceren. Er zullen geen nadere regels gesteld worden door het CvTE. Het CvTE volgt de gangbare prescripties van de Nederlandse taal zoals het Groene Boekje, maar die bieden lang niet altijd een eenduidige regel. Leraren Nederlands weten dat: sommige regels die zij leerlingen leren kunnen dan wel didactisch verantwoord zijn, taalkundigen zullen vaak meer dan een optie toestaan. Dat zou bij het nakijken van examenwerk ook zo moeten zijn.

- 3) Conventies. Correctoren zijn blij met de volgende opmerking op het CV: Als de kandidaat gebruikmaakt van conventies die niet passen bij de tekstsoort 1 scorepunt aftrekken.

Samenvatting

- 1) De tekst is boeiend. Het aantal woorden dat de leerlingen mogen gebruiken, is goed gekozen.
- 2) Het CV is helder en levert voldoende houvast bij het corrigeren.
- 3) Wanneer een leerling geen enkele naam overneemt in zijn samenvatting, kost dat 4 punten op een totaal van 51. Dat is veel te streng gerekend met een n-term van 0.3.

3. HAVO-EXAMEN NEDERLANDS 2018

- 1) Het examen was goed te maken in de tijd en de teksten waren passend.
- 2) Complimenten voor de snelle en adequate reactie van het CvTE, waardoor er aanvullingen kwamen op het cv bij de vragen 11, 12, 13, 14 en 25. Tegelijkertijd waren er dit jaar veel aanvullingen en kwam de laatste aanvulling pas op 28 mei, zodat veel tweede correctoren hiermee werden geconfronteerd.
- 3) Het examen bevatte beslist een aantal mooie vragen met goede resultaten op de tia-lijst.

3.1 Praktisch

- 1) Het is in de praktijk gebleken dat niet alle collega's op de hoogte waren van de aanvullingen. Wellicht dat het persoonlijk account dat iedereen nu kan krijgen via aanmelding door de examencoördinator hier voor verbetering zorgt. Advies: maak meer reclame voor dit persoonlijke account.
- 2) Relatief gezien bevatte het havo-examen meer open vragen dan het vwo-examen, dus hadden kandidaten meer 'gelegenheid' tot het maken van taalfouten. Hier zou meer afstemming mogen zijn tussen havo en vwo over de gewenste verdeling van vraagtypes.

3.2 Inhoudelijk

3.2.1 Vraagstelling

Vraag 9. De nummering in het schema zorgde voor onnodige verwarring. De opdracht luidt 'Vat Martins redenering uit alinea 14 samen door onderstaand schema aan te vullen', maar er wordt niet/nauwelijks samengevat in de antwoorden in het CV. Citaten zijn toegestaan volgens de septembermailing. Correctoren waren in verwarring.

Vraag 12. Lage P (22). De vraag is niet helder gesteld: het oordeel zou in alinea 3 staan volgens de vraagstelling, maar eigenlijk blijkt het oordeel pas duidelijk uit de combinatie van de alinea's 3 en 4. We vragen ons af of leerlingen hier niet benadeeld zijn.

3.2.2 Andere kwesties

Vraag 11: de kandidaten moeten twee elementen kiezen die in de inleiding voorkomen. De vraag had een lage P (25). Deze vraag lijkt niet goed te werken.

Vraag 17: Jammer dat 'oorzaak' en 'reden' tussen haakjes staan; het is een gemiste kans om te toetsen of leerlingen het verschil tussen beide tekstverbanden kennen.

4. VWO-EXAMEN NEDERLANDS 2018

- 1) De omvang van het examen was in orde. Dit werd zeer gewaardeerd.
- 2) Het niveau van de teksten was in orde.
- 3) Docenten waardeerden het dat er drie teksten over één thema gingen, over het algemeen.
- 4) Een aantal mooie vragen, die ook goed lijken te scoren in de Tia-lijst.

4.1 Praktisch

- 1) Planning van het examen: i.v.m. proefcorrectie moet het examen Nederlands op de eerste dagen van de periode worden afgenomen.
- 2) Bij het vwo-examen Nederlands waren veertien van de 35 vragen meerkeuzevragen (40%). Bij het havo-examen was dat 25%: negen meerkeuzevragen van de 36 vragen. Leerlingen die meer open vragen moeten beantwoorden, dus de havoleerlingen, zijn in het nadeel: er zullen niet veel havoleerlingen zijn die minder dan 3 of 4 punten aftrek hebben gekregen voor taal- en spelfouten. Kan op een of andere manier de aftrek wegens taal- en spelfouten worden gerelateerd aan de hoeveelheid vrije tekst (als antwoord op open vragen) die examenkandidaten moeten schrijven? Is dat mogelijk?

4.2 Inhoudelijk

4.2.1 Hoofdgedachtevragen

Vraag 14 – meerkeuzevraag naar de hoofdgedachte – heeft een beduidend lagere rir-waarde dan vraag 27, een open vraag naar de hoofdgedachte. Bij vraag 14 koos 52% voor B (fout) en 47% voor C (goed). We vragen ons af of de antwoorden goed te onderscheiden waren. Het verschil tussen B en C lijkt te berusten op het verschil constaterend – betogend. Tegelijkertijd is de formulering bij B, ‘het verdwijnen van maatschappelijk relevante studies’, niet te lezen als iets dat goed te keuren zou zijn. De studies zijn relevant, dus je wilt niet dat ze verdwijnen. De Rir-waarde (5) duidt erop dat de algemeen-goedscorende kandidaten hier niet goed scoorden.

De vraag is of de mc-vorm hier geschikt is. Een vergelijking van mc-hoofdgedachtevragen bij andere examens en de bijbehorende rir-waarden zou hier extra informatie kunnen bieden. Vraag 27 (open vraag naar de hoofdgedachte) in dit examen heeft met een lage P (27) wel een veel betere rir (13).

Bij havo-vraag 20, meerkeuzevraag naar de hoofdgedachte, is de rir steviger met een waarde van 10, maar ook daar zien we dat vrijwel evenveel leerlingen het goede antwoord B kozen als het foute antwoord D.

4.2.2 Verschillende interpretaties van woorden en begrippen

Vraag 20: de beruchte vraag. Zijn bèta-wetenschappers hier te vervangen door ingenieurs? Onze analyse – *niet van de vraag, maar van de discussie* – is dat we bij de docentbespreking in Utrecht in een tunnelvisie zijn terechtgekomen en dat het goed is dat de docentengemeenschap dit heeft rechtgezet, zodat ingenieurs over het algemeen goed gerekend is. Los van de vraag wie er hier nu op de millimeter gelijk heeft, is de vraag belangrijker of de keuze voor het ene of andere woord écht iets zegt over de leesvaardigheid van de leerling. Een leerling die voor ingenieur kiest heeft de tekst net zo goed begrepen als een leerling die voor bèta-wetenschapper kiest. Is het te voorkomen dat we in volgende jaren weer in dit soort woorddiscussies terechtkomen?

Vraag 21: Het begrip ‘rol’ in de vraagstelling wordt gezien het correctievoorschrift op twee manieren opgevat:

- 1) welke *betekenis* hebben de geesteswetenschappen voor wetenschap en samenleving?
- 2) hoe *groot* de rol is die geesteswetenschappen nu spelen?

Bij tekst 1 en 2 sluit het cv aan bij betekenis 1. Bij tekst 3 bij betekenis 2. Leerlingen kunnen niet weten dat ze het begrip ‘rol’ op verschillende manieren moeten opvatten, dus er kwamen veel antwoorden gebaseerd op betekenis 1. De vraag scoort de hoogste rir-waarde (30), dus je zou er weinig aan willen veranderen, maar toch willen we hier de vraag stellen of de opzet van de vraag in een

toekomstig examen niet beter tot zijn recht zou komen in een korte schrijfpodracht.

Vraag 28: ook hier begripsverwarring, nu over 'consequentie'. Dit soort kwesties maakt dat er veel discussie is over het correctievoorschrift.

Vraag 33: de kandidaat moet twee mogelijke parafrases geven van de titel van tekst 4, 'Wat is slim?'. Een parafrase geef je over het algemeen om een lastig stuk tekst in eigen woorden uit te leggen. Bij de eenvoudige woorden van de titel hebben de leerlingen duidelijk geworsteld met de vraag wat ze aan moesten met parafrases. Op het forum en bij de docentbespreking werd duidelijk dat leerlingen alle kanten opgingen. De vraagstelling had hier helderder mogen zijn, bijvoorbeeld: Geeft twee vragen die de titel van tekst 4 kunnen vervangen.

4.2.3 Meerkeuzevragen

Sommige meerkeuzevragen lijken goed te werken, maar andere vragen lijken niet bij te dragen aan het discriminerend vermogen van het examen. We bespreken hier vraag 3, 16 en 29.

Vraag 3: een erg lage rir. Slechts 54% heeft de vraag goed, maar blijkbaar is er weinig verband met de andere scores. Er zijn ook meerkeuzevragen met een $P > .70$ en een rir $> .10$ (bijvoorbeeld vraag 5).

Vraag 16: Over meerkeuzevragen was er op het forum relatief weinig discussie omdat er weinig invloed op is uit te oefenen; deze vraag kwam waarschijnlijk daarom niet aan de orde. De rir (4) is laag. Het foute antwoord B werd door 44% gekozen, het goede antwoord D door 46%. Voor beide antwoorden zijn er argumenten aan te dragen. De resultaten bij vraag 16 vragen om een nadere analyse. Is B (gebaseerd op alinea 2 van tekst 3) nu echt minder goed dan D (gebaseerd op alinea 4 en/of 5)?

Vraag 29: met rir 3 is deze mc-vraag niet sterk. De antwoorden A (fout, 48%) en B (goed, 48%) liggen nogal dicht bij elkaar. In A ligt de nadruk op samenleving, bij B op docenten. De kandidaat moet naar het slot van de tekst kijken en ziet dat daar leerlingen, docenten en ouders genoemd worden. Het is dan niet vreemd dat men voor 'samenleving' (A) kiest, in plaats van voor het veel beperktere 'docenten'. Het ministerie van Onderwijs communiceert wel vaker met de gehele samenleving (denk aan meisjes en bèta) en niet alleen met docenten.