

# Centraal Schriftelijke Examens Muziek 2023 tijdvak 1

Peter Schoenmakers, toetsdeskundige muziek Stichting Cito

Na drie jaar was 2023 een vrijwel 'normaal' examenjaar. De enige twee maatregelen die nog golden, waren de extra voorbereidingstijd voor de leerlingen tussen tijdvak 1 en tijdvak 2 en ook werd er bij de normering nog rekening gehouden met de docentenoordelen. Vooraf was er enige onrust of de leerlingen die in 2023 examens deden niet benadeeld werden door het schrappen van vrijwel alle 'coronamaatregelen' die voor de examens 2021 en 2022 golden. De resultaten van de examens muziek laten in 2023 echter een beeld zien dat te vergelijken is met de examens vóór de coronatijd.

Na afloop van een van de evaluatievergaderingen werd mij gevraagd hoe er bij de normering omgegaan wordt met eventuele tijdnoed in een examen. In dit artikel ga ik daar uitgebreid op in. Aan het eind van het artikel geef ik een overzicht in cijfers van de verschillende examens.

## Lengte examens

Het is belangrijk dat de examens betrouwbaar zijn en goed discrimineren. Daarmee wordt in deze context bedoeld dat de examens een zo goed mogelijk onderscheid moeten maken tussen vaardige leerlingen en minder vaardige leerlingen. De examens muziek blijken keer op keer hoog te scoren wat betreft de betrouwbaarheid. Een van de factoren die een positief effect hebben op de betrouwbaarheid is de lengte van een toets; hoe meer meetmomenten, hoe betrouwbaarder de toets. In de examens maken we daarom optimaal gebruik van de tijdsduur die het examen mag duren. Voor het examen muziek vmbo is dat twee uur en voor de examens muziek havo en vwo tweeënhalf uur. Tijdens de constructie monitoren we voortdurend de lengte. Daarbij kijken we niet alleen naar de lengte van de fragmenten, maar ook bijvoorbeeld naar de hoeveelheid vragen, de hoeveelheid tekst, het aantal te behalen scorepunten, de verdeling van de verschillende vraagvormen (meerkeuzevragen, open vragen, voorgestructureerde vragen) en het aantal notenvoorbeelden. In de Quikscan, die docenten na afname van het examen kunnen invullen, kan aangegeven worden of het examen te lang (1), precies goed (2) of te kort (3) was. In het algemeen ervaren docenten de examens als (iets) te lang. Het gemiddelde van de laatste vijf examens van tijdvak 1 is op een schaal van 3 voor vmbo 1,9, havo 1,7 en vwo 1,5.

Het is belangrijk dat een leerling het cijfer krijgt dat hij/zij verdient. Ook is het belangrijk dat het cijfer niet ten onrechte lager is, omdat de leerling in tijdnoed is gekomen. Daarom wordt er bij de vaststelling van de N-term goed naar het aspect van eventuele tijdnoed gekeken.

## Tijdnoed

Enige tijdnoed bij de leerlingen die de leerstof in onvoldoende mate beheersen, hoeft toetstechnisch geen probleem te zijn. Pas wanneer leerlingen die de leerstof in voldoende mate beheersen niet meer voldoende tijd krijgen om te laten zien wat ze waard zijn, is er sprake van problematische tijdnoed. Er moet in dat geval onderzocht worden of het effect van de tijdnoed voor leerlingen met voldoende beheersing van de leerstof anders uitpakt. In de analyse van de non-respons (niet beantwoorde vragen) wordt het effect van het missen van scorepunten geschat op basis van de aanname dat elke non-respons wordt veroorzaakt door tijdnoed; ook als het bijvoorbeeld vraag 1 betreft. Zeker bij toetsen die volledig uit open vragen bestaan, levert dat een overschatting van de tijdnoedproblematiek op. Uit de analyses bleek dat bij geen van de examens muziek in tijdvak 1 de Tijd Nood Compensatiefactor groter dan 0,05 was (zie de bijlage voor uitleg over de Tijd Nood Compensatiefactor). Het is van belang dat docenten een 'N' (non-respons) invullen in Wolf

wanneer een leerling een vraag niet beantwoord heeft (en niet de score 0); alleen dan kunnen we goed achterhalen hoe groot de non-respons op een vraag is.

De drie vragen van respectievelijk het examen vmbo, havo en vwo waarbij de non-respons het groots was, staan hieronder. In het algemeen kan men stellen dat non-respons zich vooral voordoet bij vragen naar een technische term en solfègevragen. Deze vragen worden overgeslagen omdat leerlingen het antwoord niet weten of de vraag (te) moeilijk vinden en niet vanwege tijdnoed.

### Non-respons VMBO 2023 tijdvak 1

totaal aantal leerlingen in de toets- en itemanalyse: 948

% non-respons = percentage leerlingen (uit de analyse) dat de vraag niet heeft beantwoord

# = aantal leerlingen (uit de analyse) dat de vraag niet heeft beantwoord

zij bijlage voor de uitleg over de p-waarde en de rir-waarde:

vraag	% non respons	#	p-waarde	rir-waarde
6	4,96	47	.27	.35

In het fragment spelen violen en een contrabas.

**Geef de Italiaanse technische term voor de speeltechniek op de violen en de contrabas in het fragment.**

correctievoorschrift:

**maximumscore 1**  
pizzicato

vraag	% non respons	#	p-waarde	rir-waarde
39	4,75	45	.45	.33

In het videofragment speelt een bassist mee met het einde van de song. Behalve tokkelen gebruikt de bassist nog meer speeltechnieken.

**Geef de technische term voor nog een andere speeltechniek op de basgitaar in het fragment.**

correctievoorschrift:

**maximumscore 1**

een van de volgende:

- slappen
- glissando / slide
- hammer-on
- pulling / pull-off
- popping

vraag	% non respons	#	p-waarde	rir-waarde
43	4,11	39	.18	.39

Het fragment is het einde van de treurmars voor de Franse koning. Er klinkt een eenvoudige melodie. Deze is in de uitwerkbijlage onvolledig afgedrukt.

**Noteer de ontbrekende aanwijzingen. Gebruik:**

*rit.* (= ritenuto)

***f***

***p***

**Let op: sommige aanwijzingen moet je vaker gebruiken.**

correctievoorschrift:

**maximumscore 2**



indien zes aanwijzingen juist	2
indien vijf of vier aanwijzingen juist	1
indien minder dan vier aanwijzingen juist	0

*Opmerking*

*De aanwijzing rit. mag op de derde of vierde tel van maat 5 genoteerd staan.*

## Non-respons HAVO 2023 tijdvak 1

totaal aantal leerlingen in de toets- en itemanalyse: 960

vraag	% non respons	#	p-waarde	rir-waarde
20	2,81	27	.15	.43

Van het fragment is de zangmelodie in de uitwerkbijlage onvolledig afgedrukt.

**Noteer de ontbrekende noten (toonhoogte en ritme) inclusief één toevallig voorteken.**

correctievoorschrift:


**maximumscore 2**

een van de volgende:

- ritme geheel juist 1
- toonhoogte exclusief voorteken geheel juist 1
- voorteken juist 1

*Opmerking*

*Bij het beoordelen van de toonhoogte exclusief voorteken, wordt de 'bes' in maat 3, en indien van toepassing de noot met het foutief geplaatst toevallig voorteken, genegeerd.*

vraag	% non respons	#	p-waarde	rir-waarde
24	2,81	27	.13	.36
<p>Het fragment is couplet 1 en duurt acht maten. In de begeleiding klinkt voortdurend hetzelfde ritme.</p> <p><b>Noteer één maat van het ritme van de begeleiding.</b></p> <p>correctievoorschrift:</p> <p><b>maximumscore 1</b></p> 				

vraag	% non respons	#	p-waarde	rir-waarde
40	3,96	38	.44	.25
<p>In het fragment wordt Belle door een blaasinstrument verklankt.</p> <p><b>Geef de technische term voor de articulatiwijze van de melodie.</b></p> <p>correctievoorschrift:</p> <p><b>maximumscore 1</b></p> <p>legato</p>				

## Non-respons VWO 2023 tijdvak 1

totaal aantal leerlingen in de toets- en itemanalyse: 1053

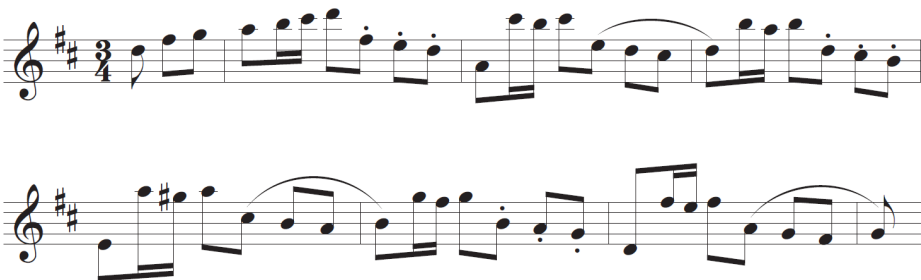
vraag	% non respons	#	p-waarde	rir-waarde
5	2,94	31	.64	.45

Van het fragment is de melodie van de hobo in de uitwerkbijlage afgedrukt.

**Noteer onder de haken alle ontbrekende staccato- en legato-tekens.**

correctievoorschrift:

**maximumscore 2**



- staccato geheel juist 1
- legato geheel juist 1

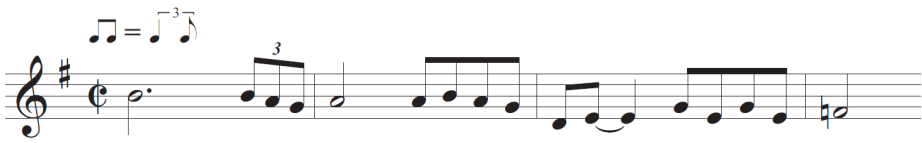
vraag	% non respons	#	p-waarde	rir-waarde
35	3,23	34	.37	.32

Het fragment is het einde van couplet 1.

**Geef de technische term voor de ritmische figuur in de elektrische gitaarpartij en de baspartij aan het eind van het fragment.**

correctievoorschrift:

**maximimscore 1**  
triool

vraag	% non respons	#	p-waarde	rir-waarde
37	9,31	98	.15	.51
<p>Het fragment is het begin van de gitaarsolo en is in de uitwerkbijlage onvolledig afgedrukt.</p> <p><b>Noteer in maat 2 en maat 3 de ontbrekende noten (toonhoogte en ritme). Let op de triolenfeel.</b></p> <p>&gt;&gt; De klank is een halve toon hoger dan de notatie.</p> <p>correctievoorschrift:</p> <p><b>maximumscore 2</b></p>  <ul style="list-style-type: none"> <li>• ritme geheel juist <span style="float: right;">1</span></li> <li>• toonhoogte geheel juist <span style="float: right;">1</span></li> </ul>				

## Overige instrumenten

De Tijd Nood Compensatiefactor is niet het enige waar we naar kijken of er gecompenseerd moet worden voor tijdnoed. Daarvoor hebben we, behalve het oordeel van de docenten, nog een aantal instrumenten.

- In de regel maken de leerlingen het examen in de volgorde zoals het wordt aangeboden, waardoor eventuele tijdnoed voorkomt bij de laatste vragen. Het kan best zo zijn dat leerlingen alle vragen beantwoord hebben waardoor er geen sprake is van non-respons, maar dat er wel sprake was van tijdnoed. Daarom wordt er extra gekeken of de laatste vragen opvallend lage p-waarden en/of rir-waarden hebben die afwijken van soortgelijke vragen in het examen. Mocht dat zo zijn, dan kan er dus sprake zijn van tijdnoed omdat de leerling snel iets heeft ingevuld, bijvoorbeeld een meerkeuzevraag heeft gekokt. Specifiek bij meerkeuzevragen kunnen we in de toets- en itemanalyse zien of het leerlinggedrag afwijkt van wat verwacht zou mogen worden. Dat is bijvoorbeeld het geval als vaardige leerlingen minder vaak het juiste antwoord geven dan minder vaardige leerlingen.

- In de analyse kunnen we ook zien hoeveel leerlingen alle vragen beantwoord hebben. Als dat minder dan 80% is, is dat een extra reden om te onderzoeken of er sprake was van tijdnoed. In de examens muziek 2023 tijdvak 1 bleef het percentage dit jaar steeds boven de 80%.

- Tot slot wordt elke vraag in een examen waarbij het percentage non-respons hoger dan 2% is apart besproken tijdens de normeringsvergadering.

Uiteindelijk hebben de analyses uitgewezen dat er niet gecompenseerd hoefde te worden voor eventuele tijdnoed in een van de examens muziek.



## Onderzoek lengte examens

In de tijd voordat de examens muziek met de computer werden afgenomen, vond de afname met een cd plaats (en daarvoor een muziekcassette en in de eerste jaren een muziekband). Deze cd werd klassikaal afgespeeld. Op de cd was alle tekst door een professionele nieuwslezeres ingesproken en hadden de examenmakers bepaald hoe vaak een fragment te horen was en hoelang de beantwoordingstijd per vraag was. Alle leerlingen doorliepen dus in hetzelfde tempo het examen. Een extra bijzonderheid was dat de cd pas vijf of tien minuten (afhankelijk van het examen) na aanvang van de examenzitting mocht worden gestart, zodat leerlingen eerst het examen globaal konden bekijken. Halverwege het examen was er een verplichte pauze van vijf of tien minuten. De examenzitting was, net als tegenwoordig, twee uur voor vmbo en tweeënhalf uur voor havo en vwo. Deze afnamemodus met een klassikale cd heeft voor het vmbo tot en met 2005 bestaan en voor havo en vwo tot en met 2013. Leerlingen kunnen tegenwoordig achter de computer het examen in hun eigen tempo maken. Het voordeel is dat fragmenten zo vaak beluisterd kunnen worden als de leerling zelf wil en dat de leerling aan het einde van het examen nogmaals fragmenten kan terugluisteren; iets dat bij de klassikale afname niet mogelijk was. Wel schuilt er het gevaar dat leerlingen te lang bij een vraag stil blijven staan door fragmenten eindeloos te herhalen. Om een beeld te krijgen over de lengte van de huidige examens muziek ten opzichte van de examens die klassikaal met een cd werden afgenomen is een analyse uitgevoerd. Hierbij is onderzocht hoe vaak een fragment gemiddeld te horen was in de examens met een klassikale afname en hoelang de totale afspeeltijd gemiddeld van de muziek was bij één keer luisteren. Voor vmbo is gekeken naar de examens uit tijdvak 1 uit 2003, 2004 en 2005. Dit waren de laatste drie examens met een klassikale afname, maar met het nieuwe vmbo-programma zoals we dat nu ook nog in grote lijnen kennen. De examens muziek (mavo) voor 2003 hadden een ander programma. Voor havo en vwo is gekeken naar de laatste vijf examens met een klassikale afname (2009 t/m 2013). De gemiddelde lengte van de fragmenten is daarna vergeleken met de gemiddelde lengte van de fragmenten uit de laatste vijf examens muziek uit tijdvak 1 (2018, 2019, 2021, 2022 en 2023). Ook is gekeken naar het gemiddeld aantal vragen, gemiddeld aantal scorepunten, en het gemiddeld aantal meerkeuzevragen in de examens in de twee tijdperiodes. Voor vwo is ook nog het gemiddeld aantal woorden in de examens berekend, omdat docenten name de vwo-examens als (iets) te lang worden beoordelen.

VMBO		
	2003 t/m 2005	2018, 2019, 2021, 2022, 2023
gemiddelde lengte fragmenten	21:21	18:44
gemiddeld aantal keer te beluisteren	2,1	vrij
gemiddeld aantal vragen	49	50
gemiddeld aantal scorepunten	57	67
gemiddeld aantal meerkeuzevragen	31	19

<b>HAVO</b>		
	2009 t/m 2013	2018, 2019, 2021, 2022, 2023
gemiddelde lengte fragmenten	17:13	14:46
gemiddeld aantal keer te beluisteren	2,2	vrij
gemiddeld aantal vragen	52	51
gemiddeld aantal scorepunten	74	75
gemiddeld aantal meerkeuzevragen	9	11

<b>VWO</b>		
	2009 t/m 2013	2018, 2019, 2021, 2022, 2023
gemiddelde lengte fragmenten	20:21	19:41
gemiddeld aantal keer te beluisteren	2,2	vrij
gemiddeld aantal vragen	53	49
gemiddeld aantal scorepunten	86	81
gemiddeld aantal meerkeuzevragen	3	7
gemiddeld aantal woorden	2120	2400

### **Uitkomst onderzoek lengte examen**

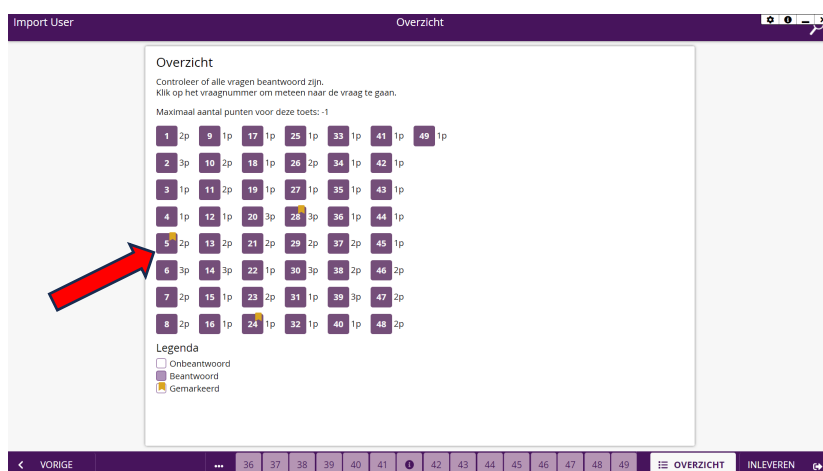
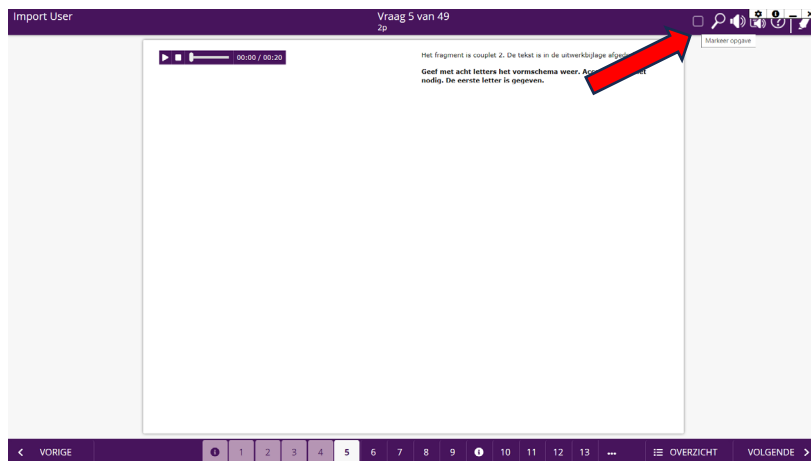
De gemiddelde lengte van de fragmenten is op alle niveaus de laatste jaren afgenomen. Voor de meeste lezers zal het opvallend zijn dat de fragmenten bij de klassikale afname gemiddeld maar iets meer dan twee keer afgespeeld werden. Dit vergde een grote mate van concentratie van de leerlingen. Daarom was er ook een verplichte pauze halverwege het examen.

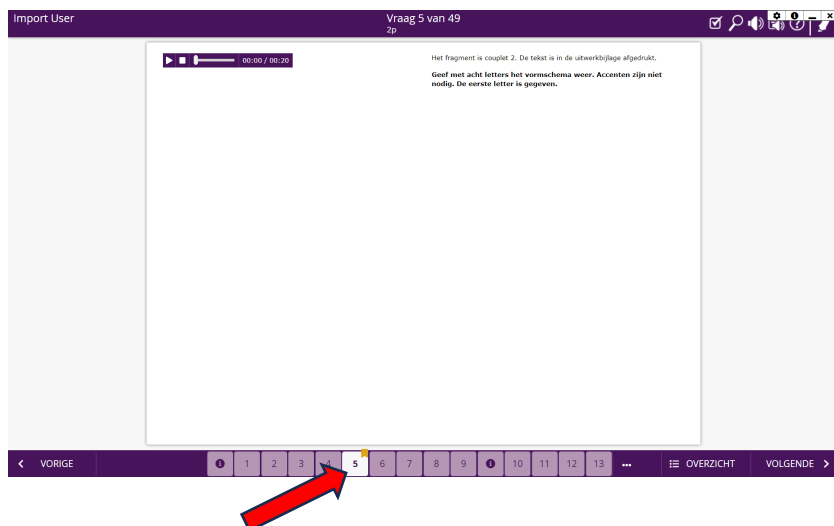
Het gemiddeld aantal vragen in de verschillende examens is over de jaren heen ongeveer gelijk gebleven. De laatste jaren bestaat het vwo-examen uit iets minder vragen. Dit komt door de analyse-opdracht. Deze opdracht bestaat uit meerdere deelvragen waarvoor scorepunten te behalen zijn. Ook de verdeling van de opgavenblokken is in de loop van de tijd veranderd. Tot en met 2015 bestond het vmbo-examen uit negen blokken. Om meer ruimte te geven aan populaire muziek, is één blok met klassieke muziek geschrapt. Ook in het havo- en vwo-examen is meer ruimte gekomen voor populaire muziek. In 2011 is in het havo-examen één blok met klassieke muziek ingeruild voor een extra blok populaire muziek. Sinds 2019 heeft het vwo-examen een themablok met populaire muziek. Dit themablok is groter dan de overige drie blokken. Het blok omvat ongeveer 40% van het examen.

Wat betreft het aantal te behalen scorepunten valt het volgende op. Vooral in het vmbo-examen is het aantal te behalen scorepunten aanzienlijk toegenomen. Deze toename valt te verklaren uit het feit dat de huidige vmbo-examens uit minder meerkeuzevragen en meer open vragen bestaan. In het vwo-examen lijkt een afname van het gemiddeld aantal te behalen scorepunten het geval te zijn. Dit cijfer is echter enigszins vertroebeld, omdat er 102 scorepunten behaald konden worden in het examen 2009. Over de jaren 2010 t/m 2013 is het gemiddeld te behalen scorepunten 82. Wel zien we een toename van het aantal woorden in het vwo-examen.

## Timemanagement

Alle scholen zijn verplicht om minimaal één keer met hun leerlingen te oefenen in Facet. Hiervoor moet de school een zogenaamde 'Proef op de Som' afname inplannen. Het doel van de Proef op de Som is de echte afnames goed voor te bereiden voor alle betrokkenen: leerlingen, docenten, systeembeheerders en examensecretarissen. Bij de Proef op de Som afname moeten alle leerlingen tegelijk oefenen met de afname in Facet. Voor de leerlingen is dit een moment om ook inhoudelijk te oefenen. Daarnaast ervaren leerlingen hoe het is om de aandacht de gehele examenduur vast te houden. Daarbij zou ik adviseren dat de leerlingen de opdracht krijgen om fragmenten gemiddeld maximaal drie keer te beluisteren. In dat geval houdt een vmbo-leerling één minuut beantwoordingstijd per vraag over en een havo- en vwo-leerling bijna twee minuten beantwoordingstijd per vraag zonder in tijdnood te komen. Wijs leerlingen erop dat ze niet eindeloos fragmenten moeten blijven beluisteren in de hoop het goede antwoorden te horen. Ze kunnen in dat dat geval beter verder gaan naar de volgende vraag. Ik citeer hier een collega-muziekdocent: "Als je het na vier keer luisteren niet hoort, hoor je het na zeven keer ook niet." In Facet bestaat de mogelijkheid om een vraag te markeren. Hierdoor kun je aan het einde van het examen makkelijk zien welke vraag nog enige aandacht nodig heeft.





## Systemcheck

Voor de volledigheid; de Proef op de Som staat los van de verplichte systeemcheck die een school voor 15 december moet uitvoeren. CvTE ziet (strenger) toe dat elke school die een examenafname in Facet heeft, de systeemcheck heeft doorlopen. Wanneer de systeemcheck en de Proef op Som goed doorlopen worden, is de kans op problemen met de instellingen minimaal. Overigens kunnen leerlingen altijd oefenen met examens in Facet via de oefenomgeving: [Facet oefenomgeving](#).

## De examens in cijfers

### VMBO 2023-1

aantal scholen	94
aantal leerlingen	1011
aantal vragen	51
maximale score	68
gemiddelde score	35,4
gemiddelde p-waarde	.52
N-term	1,5
gemiddeld cijfer	6,2
betrouwbaarheid (Cronbach's $\alpha$ )	.77

Quickscan:

moeilijkheidsgraad	2,29 (van 5)
lengte examen	1,88 (van 3)
inhoudelijke aansluiting op het onderwijs	3,05 (van 5)
waardering	6,4 (van 10)

De vragen waar de vmbo-leerlingen de meeste moeite mee hadden waren vraag 29, 43 en 45.

<b>vraag</b>	<b>p-waarde</b>	<b>rir-waarde</b>
29	.17	.16
<p>Het fragment is het refrein.</p> <p><b>Welk instrument klinkt?</b></p> <p><b>A</b> fagot <b>B</b> saxofoon <b>C</b> trombone <b>D</b> tuba</p> <p>correctievoorschrift: <b>A</b></p>		

<b>vraag</b>	<b>p-waarde</b>	<b>rir-waarde</b>
43	.18	.39
zie boven pagina 4		

vraag	p-waarde	rir-waarde
45	.20	.21

Welk instrument verklankt de kanonschoten? Omcirkel in de uitwerkbijlage het nummer van de juiste afbeelding en geef de (Nederlandse) naam van het instrument.



correctievoorschrift:

**maximumscore 1**

afbeelding: 2

naam instrument: grote trom

*Opmerking*

*Alleen als beide antwoorden juist zijn, 1 scorepunt toekennen.*

Vraag 43 heeft een rir-waarde van .39 en is daarmee de vraag die in dit examen het best onderscheid maakt tussen vaardige en minder vaardige leerlingen.

Het examen vmbo 2023-2 werd door negen leerlingen gemaakt.

## Havo 2023-1

aantal scholen	122
aantal leerlingen	1132
aantal vragen	49
maximale score	79
gemiddelde score	44,59
gemiddelde p-waarde	.56
N-term	1,2
gemiddeld cijfer	6,3
betrouwbaarheid (Cronbach's $\alpha$ )	.77

Quickscan:

moeilijkheidsgraad	2,81 (van 5)
lengte examen	2,08 (van 3)
inhoudelijke aansluiting op het onderwijs	3,45 (van 5)
waardering	7,11 (van 10)

De vragen waar de havo-leerlingen de meeste moeite mee hadden waren vraag 8, 20 en 24.

vraag	p-waarde	rir-waarde
8	.08	.33
<p>Het fragment is een gedeelte van couplet 3. De melodie is in de uitwerkbijlage onvolledig afgedrukt.</p> <p><b>Noteer de ontbrekende noten. Het ritme is gegeven.</b></p> <p>correctievoorschrift:</p> <p><b>maximumscore 2</b></p> <p>Wo - on nin man man fo wo kè - nou (A-loug - bi - rin)</p> <p>per haak geheel juist</p>		

vraag	p-waarde	rir-waarde
20	.15	.43
zie boven pagina 5		

vraag	p-waarde	rir-waarde
24	.13	.36
zie boven pagina 6		

Vraag 20 en vraag 28 (ook een solfègevraag) hebben beide een rir-waarde van .48 en zijn daarmee de vragen die in dit examen het best onderscheid maken tussen vaardige en minder vaardige leerlingen. Vraag 28 heeft echter een p-waarde van .40. Aangezien de p-waarde bij vraag 28 dichterbij .50 ligt, heeft dit toetstechnisch de voorkeur.

Het examen havo 2023-2 werd door negentien leerlingen gemaakt.

### VWO 2023-1



aantal scholen	118
aantal leerlingen	1192
aantal vragen	48
maximale score	81
gemiddelde score	44,19
gemiddelde p-waarde	.55
N-term	1,6
gemiddeld cijfer	6,5
betrouwbaarheid (Cronbach's $\alpha$ )	.84


Quickscan:

moeilijkheidsgraad	2,21 (van 5)
lengte examen	1,44 (van 3)
inhoudelijke aansluiting op het onderwijs	3,17 (van 5)
waardering	6,23 (van 10)



De vragen waar de vwo-leerlingen de meeste moeite mee hadden waren vraag 11, 20 en 37.

vraag	p-waarde	rir-waarde
11	.06	.14
<p>Het fragment is de afsluiting van een tutti-gedeelte. De partij van viool 1 is in de uitwerkbijlage (enigszins vereenvoudigd) afgedrukt.</p> <p><b>Geef met een haak de hemiool aan.</b></p> <p>correctievoorschrift: <b>maximumscore 1</b></p>  <p>aanvulling op het correctievoorschrift: ook goed:</p> 		

vraag	p-waarde	rir-waarde
20	.18	.17
<p>Fragment 1 is het begin van thema 1. De partij van viool 1 is weergegeven.</p>  <p>In fragment 2 wordt het motief onder de haak ritmisch gevarieerd.</p> <p><b>Op welke manier wordt het motief onder de haak ritmisch gevarieerd in fragment 2? Geef de technische term.</b></p> <p>correctievoorschrift: <b>maximumscore 1</b> vergroting</p>		

<b>vraag</b>	<b>p-waarde</b>	<b>rir-waarde</b>
37	.15	.51
zie boven pagina 8		

Vraag 37 en vraag 27 (de grote analyse-opdracht) hebben beide een rir-waarde van .51 en zijn daarmee in dit examen de vragen die het best onderscheid maken tussen vaardige en minder vaardige leerlingen. Vraag 27 heeft echter een p-waarde van .43. Aangezien de p-waarde van vraag 27 dichterbij .50 ligt, heeft dit toetstechnisch de voorkeur.

Het examen vwo 2023-2 werd door eenentwintig leerlingen gemaakt.

## Bijlage

### Tijd Nood Compensatiefactor

We gaan voor de berekening van een eventuele compensatie voor tijdnoed op zoek naar de mate waarin de bedachtzame in wat trager tempo werkende leerling, die de leerstof wel in voldoende mate beheerst, een meer dan gemiddelde aantal non-respons laat zien. Voor het bepalen van een eventuele compensatie vanwege tijdnoed, verdelen we de scores van de leerlingen in vier kwartielen. Het minst vaardige kwartiel beschouwen we als het kwartiel waarvan we niet weten of de leerlingen de leerstof wel in voldoende mate beheersen en laten we daarom verder buiten beschouwing. Vervolgens wordt het potentiële effect van de non-respons op het cijfer berekend voor de drie hoogst scorende kwartielen. Overal waar een vraag is overgeslagen wordt een score ingevuld die overeenkomt met de score die een leerling in dat kwartiel gemiddeld haalt. Op basis van het verschil wordt het maximale effect van tijdnoed op het cijfer bepaald. Van de drie hoogst scorende kwartielen nemen we aan dat ze de leerstof wel in voldoende mate beheersen. Voor de bepaling van de Tijd Nood Compensatiefactor (TNC) gaan we na of de leerlingen in het één na laagst scorende kwartiel ernstiger door non-respons gedupeerd zijn dan het gemiddelde over alle leerlingen met voldoende beheersing van de leerstof. Wanneer de TNC groter dan 0,05 cijferpunt is, wordt toepassing van extra compensatie overwogen.

### p-waarde

De gemiddelde p-waarde of het 'gemiddelde percentage goed' is een maat voor de moeilijkheidsgraad van de toets of vraag. De p-waarde ligt tussen 1.00 en .00. De p-waarde wordt berekend door de gemiddelde score te delen door de maximaal te behalen score. Een p-waarde boven .85 wijst op een (extreem) makkelijke vraag of toets, een p-waarde beneden .30 wijst op een (extreem) moeilijke vraag of toets.

### voorbeelden berekenen p-waarde

maximumscore 1

leerling 1: 1p

leerling 2: 1p

leerling 3: 0p

gemiddelde score = 0,67

(gemiddelde score) 0,67 / (maximumscore) 1 = p-waarde .67

maximumscore 2

leerling 1: 2p

leerling 2: 1p

leerling 3: 1p

gemiddelde score = 1,3

(gemiddelde score) 1,3 / (maximumscore) 2 = p-waarde .65

We streven bij de constructie naar een gemiddeld p-waarde van ca. .50.

**rir-waarde**

De rir-waarde geeft informatie over de discriminerende werking van een item (vraag). De rir-waarde geeft de correlatie aan tussen de itemscore en de score van de overige items in de toets. Het zegt iets over in hoeverre leerlingen die de betreffende vraag juist hebben beantwoord ook de overige vragen in de toets juist hebben beantwoord. We streven bij de constructie naar een rir-waarde  $> .12$ .

**Cronbach's  $\alpha$** 

Een belangrijke kwaliteit van de toets als geheel is de betrouwbaarheid. Binnen de toets- en itemanalyse wordt de betrouwbaarheid geschat door te onderzoeken in hoeverre de verschillende items in de toets allemaal hetzelfde meten oftewel in welke mate ze onderling correleren. De betrouwbaarheid wordt geschat door de berekening van een betrouwbaarheidscoëfficiënt. Dat is altijd een getal dat ligt tussen 0 en 1. Hoe hoger dat getal, hoe betrouwbaarder de toets. Er wordt gestreefd naar een betrouwbaarheid van ten minste  $.70$ .