

Cohortonderzoek COOL⁵⁻¹⁸

Technisch rapport meting HAVO-5 in 2010
(tweede versie)

Jos Keuning | Jolijn Hendriks | Djurre Zijsling

Colofon

Het Cohortonderzoek Onderwijsloopbanen (COOL⁵⁻¹⁸) wordt in het voortgezet onderwijs uitgevoerd door GION en Cito in opdracht van de Programmaraad voor het Onderwijsonderzoek (PROO).

Bij vermelding van dit rapport in een publicatie dient gerefereerd te worden aan:

Keuning, J., Hendriks, J., & Zijsling, D. (2013). *Cohortonderzoek COOL⁵⁻¹⁸: Technisch rapport meting HAVO-5 in 2010 (tweede versie)*. Groningen: GION.

Projectmanagement

Greetje van der Werf, Hans Kuyper, Jos Keuning

Werving

Service Unit, Cito

Logistieke ondersteuning

Service Unit, Cito

Ontwerp vragenlijst

Hans Kuyper

Psychometrie en methodologie

Jos Keuning, Djurre Zijsling, Jolijn Hendriks

Redactie

Hans Kuyper, Djurre Zijsling, Jos Keuning

Eindredactie databestand

Djurre Zijsling

© December 2013. GION, Gronings Instituut voor Onderzoek van het Onderwijs

Niets van deze uitgave mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op welke andere wijze dan ook zonder voorafgaande schriftelijke toestemming van de Directeur van het Instituut.

Ten geleide bij de tweede versie

Het voorliggende rapport geeft een beschrijving van de gegevens die als onderdeel van COOL⁵⁻¹⁸ in het schooljaar 2009/10 zijn verzameld in de vijfde klas van het HAVO. De eerste versie van dit rapport en het bijbehorende databestand zijn in november 2012 deponéerd bij DANS.

Zoals niet ongebruikelijk bleken in een later stadium zowel het rapport als het databestand enige onvolkomenheden te bevatten. In deze tweede versie zijn deze gecorrigeerd.

De verschillen t.o.v. de eerste versie zijn de volgende.

In het hoofdstuk over de leerlingvragenlijst wordt in §4.3.7 de schaal “examenvaardigheden” beschreven. In de tekst heeft de schaalscore de naam “**LV_EXVA**”, in de variabelenlijst in hoofdstuk 5 en in het databestand heeft deze de naam “**LV11_S**”. In de tweede versie heeft deze overal de naam “**LV_EXVA**”.

In de tweede versie van het rapport zijn enkele – voornamelijk cosmetische – wijzingen aangebracht en is de layout aangepast aan de layout van de overige rapporten die betrekking hebben op COOL⁵⁻¹⁸.

INHOUDSOPGAVE

1	INLEIDING	1
2	DATAVERZAMELING	3
2.1	WERVING	3
2.2	MATERIAAL EN PROCEDURE	5
2.3	REPRESENTATIVITEIT EN RESPONS	6
3	EXAMENGEGEVENS	15
3.1	NEDERLANDS	15
3.1.1	<i>Kenmerken</i>	15
3.1.2	<i>Scoretabel</i>	17
3.2	WISKUNDE	18
3.2.1	<i>Kenmerken</i>	18
3.2.2	<i>Scoretabel</i>	21
3.3	ENGELS	21
3.3.1	<i>Kenmerken</i>	21
3.3.2	<i>Scoretabel</i>	22
4	LEERLINGVRAGENLIJST	23
4.1	INLEIDING	23
4.2	PROCEDURES	23
4.2.1	<i>Invulinstructie</i>	23
4.2.2	<i>Ontbrekende waarden</i>	24
4.2.3	<i>Dubbele coderingen</i>	24
4.2.4	<i>Constructie schaalscores</i>	24
4.2.5	<i>Weigeraars</i>	24
4.3	SYSTEMATISCHE BESCHRIJVING LEERLINGVRAGENLIJST	25
4.3.1	<i>Tijdbesteding</i>	25
4.3.2	<i>Schoolwerk in vakanties</i>	26
4.3.3	<i>Schoolverzuim</i>	27
4.3.4	<i>Bijles</i>	28
4.3.5	<i>Examenvoorbereiding</i>	29
4.3.6	<i>Slaagverwachting</i>	30
4.3.7	<i>Examenvaardigheden</i>	30
4.3.8	<i>Welbevinden met klasgenoten</i>	31
4.3.9	<i>Persoonlijkheid</i>	31
4.3.10	<i>Vrienden/vriendinnen</i>	33
4.3.11	<i>Waargenomen autonomie</i>	34
5	BESCHRIJVING DATABESTAND	37
	LITERATUUR	41
	BIJLAGEN	43
1	SCORETABEL 1 ^{STE} TIJDVAK EXAMEN HAVO NEDERLANDS	43
2A	SCORETABEL 1 ^{STE} TIJDVAK EXAMEN HAVO WISKUNDE A	44

2B	SCORETABEL 1 ^{STE} TIJDVAK EXAMEN HAVO WISKUNDE B	45
3	SCORETABEL 1 ^{STE} TIJDVAK EXAMEN HAVO ENGELS	46

1 INLEIDING

Het CohortOnderzoek OnderwijsLoopbanen (COOL⁵⁻¹⁸) wordt in opdracht van de ProgrammaRaad voor het OnderwijsOnderzoek (PROO) uitgevoerd door een consortium bestaande uit Cito, GION, ITS en Kohnstamm Instituut met het CBS als extra betrokken partij. In COOL⁵⁻¹⁸ worden leerlingen gevolgd in hun schoolloopbaan vanaf groep 2 van de basisschool tot en met het eindexamen HAVO/VWO, dan wel het tweede jaar in het MBO. ITS en Kohnstamm Instituut zijn verantwoordelijk voor het veldwerkonderzoek in het primair onderwijs; Cito en GION voeren het veldwerkonderzoek uit in het voortgezet onderwijs en het middelbaar beroepsonderwijs. In de fase voorafgaand aan het veldwerkonderzoek, levert het CBS informatie over de meest recente plaats in het onderwijsstelsel van de leerlingen die eerder hebben deelgenomen. In de fase na het veldwerkonderzoek is het CBS verantwoordelijk voor het zogenaamde verRINnen van de veldwerkbestanden, waardoor het mogelijk wordt deze veldwerkbestanden binnen de beveiligde omgeving van het CBS te koppelen aan de onderwijsnummerbestanden en andere door het CBS beheerde bestanden.

In dit rapport wordt een technische beschrijving gegeven van de gegevens die als onderdeel van COOL⁵⁻¹⁸ in het schooljaar 2009/10 zijn verzameld in de vijfde klas van het HAVO. Een aantal van de leerlingen van wie in het voorjaar van 2010 gegevens zijn verzameld in HAVO-5 had eerder deelgenomen aan de dataverzameling van COOL⁵⁻¹⁸ in de derde klas van het voortgezet onderwijs die in het voorjaar van 2008 heeft plaatsgevonden, en/of aan de meting van PRIMA die in 2005 in groep 8 van het basisonderwijs heeft plaatsgevonden. Dit waren de 'target' leerlingen voor de deelname in HAVO-5. De doelstelling was een zo groot mogelijk aantal van deze target leerlingen te laten participeren in de nieuwe dataverzameling in HAVO-5. Naast deze target leerlingen heeft een vrij groot aantal van hun klasgenoten deelgenomen aan deze dataverzameling. Er zijn in HAVO-5 twee soorten gegevens verzameld, namelijk examengegevens en vragenlijstgegevens. De examengegevens betreffen de vakken Nederlands, wiskunde en Engels. De vragenlijstgegevens betreffen diverse aspecten, die mogelijk van invloed zijn op de schoolloopbaan dan wel van intrinsiek belang worden geacht.

Hoofdstuk 2 van dit rapport gaat in op de dataverzameling, namelijk de werving van scholen, de gebruikte materialen en de procedures. Tevens wordt ingegaan op de respons en de representativiteit daarvan. Hoofdstuk 3 is gewijd aan de examengegevens en hoofdstuk 4 aan de vragenlijst die bij de leerlingen is afgenomen. Hoofdstuk 5 bevat een beknopte beschrijving van het databestand, dat enerzijds gedeponeerd is bij DANS (Data Archiving and Networked Systems) en anderzijds geplaatst is binnen de 'remote access' omgeving van het CBS. De versie van het bestand die binnen de CBS omgeving is geplaatst, bevat tevens gegevens over de onderwijsposities van de leerlingen. Tevens kunnen binnen deze omgeving koppelingen gemaakt worden met andere databestanden die beheerd worden door het CBS.

2 DATAVERZAMELING

2.1 Werving

In de tweede helft van schooljaar 2009/2010 heeft een vervolgmeting van COOL⁵⁻¹⁸ plaatsgevonden in HAVO-5. Voor deze meting was het uitgangspunt om vooral leerlingen in de onderzoeksgroep op te nemen die eerder aan PRIMA en/of COOL⁵⁻¹⁸ hadden deelgenomen. Concreet werden de volgende leerlingen beschouwd als target leerlingen:

1. Leerlingen die in groep 8 van het primair onderwijs deelgenomen hadden aan PRIMA en in leerjaar 3 van het voortgezet onderwijs aan COOL⁵⁻¹⁸;
2. Leerlingen die in groep 8 van het primair onderwijs deelgenomen hadden aan PRIMA, maar in leerjaar 3 van het voortgezet onderwijs (helaas) niet aan COOL⁵⁻¹⁸;
3. Leerlingen die in groep 8 van het primair onderwijs niet deelgenomen hadden aan PRIMA, maar in leerjaar 3 van het voortgezet onderwijs wel aan COOL⁵⁻¹⁸ (de klasgenoten van subgroep 1).

Om de scholen met target leerlingen te kunnen benaderen, heeft het CBS voorafgaand aan de werving een bestand geleverd met de onderwijsposities van de target leerlingen in het vorige schooljaar, dat wil zeggen schooljaar 2008/2009. Het was niet mogelijk om de onderwijsposities van de leerlingen in schooljaar 2009/2010 te krijgen, omdat het CBS ten tijde van de werving nog niet beschikte over het nieuwste onderwijsnummerbestand. Het bestand van het CBS bracht dus enige onzekerheid met zich mee. Er was bekend welke target leerlingen in schooljaar 2008/2009 op welke school in HAVO-4 zaten, maar het was niet zeker of de leerlingen ook daadwerkelijk waren doorgestroomd naar HAVO-5.

Om de beoogde onderzoeksgroep op tijd beschikbaar te hebben, is in februari 2010 door het Cito gestart met het benaderen van scholen. In totaal is bij 155 scholen schriftelijk de bereidheid gepeild om deel te nemen. Van deze groep zijn 44 scholen benaderd met de vraag om met alle leerlingen in HAVO-5 deel te nemen (= collectieve deelname). Dit waren de scholen die bij de eerste meting van COOL⁵⁻¹⁸ in schooljaar 2007/2008 ook al met HAVO leerlingen mee hadden gedaan. De overige 111 scholen zijn benaderd met de vraag om alleen met de ex-PRIMA leerlingen in HAVO-5 deel te nemen (= individuele deelname). Op deze scholen zaten geen leerlingen die eerder aan COOL⁵⁻¹⁸ deel hadden genomen. Op de 155 aangeschreven scholen zaten volgens de gegevens van het CBS, 2014 ex-COOL⁵⁻¹⁸ leerlingen (subgroepen 1 en 3) en 1155 ex-PRIMA leerlingen (subgroep 2). Naast deze target leerlingen zou op de collectief aangeschreven scholen een substantieel aantal klasgenoten meedoen. Naar verwachting zou het gaan om ongeveer 70 klasgenoten per school. Als alle aangeschreven scholen bereid zouden zijn om deel te nemen aan de meting in HAVO-5 zouden in totaal dus zo'n 7000 leerlingen in de onderzoeksgroep opgenomen worden.

Tegen het einde van februari 2010 hadden 75 van de 155 aangeschreven scholen gereageerd, hetzij met het bericht om deel te willen nemen, hetzij met het bericht om niet deel te willen nemen. Een beperkt aantal scholen is toen nogmaals telefonisch of per mail benaderd (16 in totaal). Het ging om scholen die al eerder aan COOL⁵⁻¹⁸ mee hadden gedaan. Begin maart hadden 58 scholen toegezegd mee te willen werken: 26 scholen collectief en 32 scholen individueel. De redenen waarom scholen niet mee wilden werken waren divers. Er waren echter meerdere scholen die aangaven dat het onderzoek de vorige keer te veel tijd had gekost, zowel van de leerlingen als van de organisatoren op school. Op de scholen die hadden toegezegd

mee te willen werken, zaten volgens de gegevens van het CBS 1456 ex-COOL⁵⁻¹⁸ leerlingen (subgroepen 1 en 3) en 369 ex-PRIMA leerlingen (subgroep 2). Daarnaast zouden er naar verwachting $26 \times 70 = 1820$ klasgenoten in de onderzoeksgroep zitten. Op basis van de gegevens die op dat moment beschikbaar waren, werd dus gerekend op een minimale steekproefomvang van 3500 leerlingen, van wie ongeveer de helft target leerling zou zijn. Een dergelijke samenstelling en omvang van de onderzoeksgroep werd als adequaat gezien en daarom is de actieve werving van scholen toen stopgezet.

In de eerste helft van maart 2010 zijn de gegevens van de leerlingen opgevraagd bij de 58 scholen die hadden toegezegd mee te zullen werken aan COOL⁵⁻¹⁸. Deze leerlinggegevens waren nodig om de vragenlijsten op naam te kunnen maken. Bovendien kon op basis van de leerlinggegevens een definitieve schatting gemaakt worden van het totale aantal deelnemende leerlingen. Aan de scholen die collectief deel zouden nemen, is gevraagd om de volgende gegevens digitaal naar Cito te sturen:

1. Naam van de leerling (voornaam en achternaam)
2. Klas (de aanduiding, zoals die op de school gebruikt wordt)
3. Geboortedatum
4. Geslacht

Bij de scholen die individueel zouden deelnemen, is een andere werkwijze gevolgd. Met behulp van de gegevens van het CBS is voor elke school een lijst gemaakt met daarop de namen, het geslacht en de geboortedatum van de ex-PRIMA leerlingen op die school. Er is aan deze scholen gevraagd om de gegevens te controleren en aan te geven in welke klas de betreffende leerlingen zaten. Om onduidelijke redenen hebben drie scholen die zich hadden aangemeld, ook na herhaaldelijk aandringen onzerzijds, geen gehoor gegeven aan de vraag om de gegevens van de leerlingen toe te sturen of te controleren. De betreffende scholen zijn uiteindelijk dus alsnog afgevalen.

De 55 overgebleven scholen hadden in totaal 3093 leerlingen aangemeld. Op basis van de namen van de leerlingen is een koppeling gemaakt met het databestand van de eerste meting van COOL⁵⁻¹⁸. Op deze manier kon nagegaan worden welke leerlingen al eerder aan COOL⁵⁻¹⁸ deel hadden genomen. Van de 3093 leerlingen kwamen er 843 ook voor in het databestand uit 2007/2008 (subgroepen 1 en 3). Daarnaast bleken er, na controle door de 'individueel' deelnemende scholen, 330 ex-PRIMA leerlingen in de onderzoeksgroep te zitten. Van de totale onderzoeksgroep bleek dus iets meer dan een derde aangemerkt te kunnen worden als target leerling. De omvang en samenstelling van de onderzoeksgroep zoals die bleek te zijn na tussenkomst van de scholen week dus nogal af van de eerste schatting. Het aanzienlijke verschil tussen de eerste schatting en de uiteindelijk gerealiseerde omvang en samenstelling van de onderzoeksgroep is onder andere het gevolg van het enigszins 'verouderd' zijn van de gegevens van het CBS over de onderwijsposities van de target leerlingen. Zo bleek na tussenkomst van de scholen dat bij de 'individueel' deelnemende scholen gemiddeld iets meer dan 20 procent van de leerlingen niet doorgestroomd was naar HAVO-5. Bij de 'collectief' deelnemende scholen lag dat percentage gemiddeld iets hoger, namelijk op ruim 26 procent. Door deze uitval komt de verhouding tussen het aantal target leerlingen en het aantal klasgenoten vanzelfsprekend anders te liggen.

Samenvattend: iets minder dan de helft van de scholen die in schooljaar 2007/2008 hadden deelgenomen aan COOL⁵⁻¹⁸ was bereid om opnieuw deel te nemen. Van de scholen die benaderd zijn voor individuele deelname, en die niet eerder hadden deelgenomen aan COOL⁵⁻¹⁸, was ongeveer een derde bereid mee te werken. Als we naar de percentages op leerlingniveau kijken, zien we dat er volgens de gegevens van het CBS maximaal 2014 COOL⁵⁻¹⁸ leerlingen (subgroepen 1 en 3) en 1985 ex-PRIMA leerlingen (subgroep 2) in schooljaar 2009/2010 in HAVO-5 hadden kunnen zitten. Van de COOL⁵⁻¹⁸ groep hebben er 843

meegedaan (bijna 42 procent) aan de meting in HAVO-5 en van de ex-PRIMA groep 302 (ruim 15 procent). Als er gecorrigeerd wordt voor het enigszins ‘verouderd’ zijn van de CBS gegevens liggen deze percentages respectievelijk op 56.5 procent ($843 \div [2014 \times 0.74]$) en 19.0 procent ($302 \div [1985 \times 0.80]$).

2.2 Materiaal en procedure

De dataverzameling in HAVO-5 bestond uit twee onderdelen. Enerzijds zijn de resultaten van de leerlingen op het centraal schriftelijk eindexamen opgevraagd. Het ging hierbij om de eerste tijdvakexamens Nederlands, Engels en wiskunde A en B. Dit zijn dezelfde vakken waarvoor in de COOL-meting in VO-3 toetsen zijn afgenomen (tekstbegrip Nederlands, wiskunde en Engels). In de tweede plaats is een vragenlijst aan de deelnemende leerlingen voorgelegd. Met behulp van deze vragenlijst is geprobeerd om inzicht te krijgen in de volgende concepten: tijdsbesteding, schoolwerk in vakanties, schoolverzuim, bijles, examenvorbereiding, slaagverwachting, examenvaardigheden, welbevinden met klasgenoten, persoonlijkheid, vrienden/vriendinnen, en waargenomen autonomie. De informatie uit de vragenlijsten werd uiteraard vertrouwelijk behandeld. De antwoorden werden anoniem verwerkt en de resultaten werden niet naar de scholen teruggekoppeld. Bijna alle vragen hadden gesloten antwoordmogelijkheden.

Om de meting op de scholen goed te laten verlopen is een handleiding naar scholen gestuurd. De handleiding was bedoeld voor contactpersonen, examencoördinatoren en afnamebegeleiders. De contactpersoon was degene die het onderzoeksmateriaal ontving, verspreidde binnen de school, en terugstuurde naar Cito. De examencoördinator was degene die ervoor zorgde dat de examenresultaten van alle deelnemende leerlingen via het computerprogramma WOLF naar Cito gestuurd werden. De afnamebegeleider was verantwoordelijk voor de afname van de vragenlijst. Hij zorgde er bij de afname voor dat de vragenlijsten bij de juiste leerlingen terechtkwamen, dat de vragenlijsten weer ingenomen werden en dat er tijdens het beantwoorden van de vragenlijsten individueel te werk werd gegaan. Vanzelfsprekend was het mogelijk dat personen een dubbelfunctie vervulden. In de handleiding was precies aangegeven hoe de vragenlijstafname georganiseerd moest worden en hoe de digitale verzending van examenresultaten via WOLF moest plaatsvinden. Naast deze specifieke instructies bevatte de handleiding informatie over de achtergronden van COOL⁵⁻¹⁸ en de fasering van de meting in HAVO-5.

De materialen zijn in dozen naar de scholen gestuurd. Het materiaal was zoveel mogelijk per klas geordend. Elke verzenddoos bevatte een geleidebon, een leerlinglijst, een handleiding en een pakket met vragenlijsten. In de vragenlijsten kwamen persoonlijke en dus vertrouwelijke vragen voor. Het was nadrukkelijk niet de bedoeling dat de betrokkenen op school de ingevulde vragenlijsten zouden inzien. Daarom bevatten de verzenddozen ook enveloppen waarmee de leerlingen hun vragenlijst konden ‘verzegelen’. De materialen waren uniek gecodeerd en op naam gesteld. Als een afnamebegeleider materiaal tekort kwam, kon er dus niet zomaar gekopieerd worden. Dan zouden de antwoorden van twee verschillende leerlingen immers bij één en dezelfde leerling terechtkomen. Elke school kreeg daarom een set met reservematerialen. Op het reservemateriaal was alleen een unieke barcode aangebracht; de gegevens van de leerling moesten handmatig ingevuld worden. De verwerking van de materialen heeft plaatsgevonden bij Cito.

De meting in HAVO-5 liep van 22 maart 2010 tot en met 11 juni 2010. Kort voor de start van het onderzoek zijn alle materialen naar de deelnemende scholen gestuurd. De afname van de vragenlijst kon flexibel plaatsvinden vanaf 22 maart. De afname van de vragenlijst nam maximaal 1 lesuur in beslag. Er is aan de scholen gevraagd om de vragenlijsten vóór 26 april te retourneren. De leerlingen konden de vragenlijsten dan ruim voor de examenperiode maken en de eerste verwerking van de vragenlijsten kon

tijdens de examenperiode plaatsvinden. Verreweg de meeste scholen hebben aan deze oproep gehoor gegeven. In het databestand geeft de variabele **DAT_LV** aan op welke datum de leerlingen de vragenlijsten hebben ingevuld. De examens werden vanzelfsprekend gemaakt op de voorgeschreven datums voor het eerste tijdvak. Het examen Nederlands werd gemaakt op 17 mei, het examen Engels op 27 mei en de examens wiskunde A en B op 18 mei. Voor de volledigheid zijn deze datums ook in het databestand opgenomen. De betreffende variabelen hebben respectievelijk de naam: **DAT_NED**, **DAT_ENG**, **DAT_WIA** en **DAT_WIB**. Voor het inzenden van de examenresultaten van de leerlingen die deelnamen aan COOL⁵⁻¹⁸ moesten scholen de inzenddatums aanhouden zoals die elk jaar vastgesteld worden. Nadat de scholen de examenresultaten ingestuurd hadden, heeft er een terugrapportage naar de scholen plaatsgevonden.

2.3 Representativiteit en respons

In totaal hebben 3024 leerlingen verspreid over 55 scholen deelgenomen aan de dataverzameling. Van 69 leerlingen in de oorspronkelijke groep van 3093 leerlingen (zie paragraaf 2.1) zijn dus van geen enkel onderdeel gegevens beschikbaar. Een leerling is tot de onderzoeksgroep gerekend als er gegevens beschikbaar waren van ten minste één onderdeel. In het databestand geeft de variabele **NTOETS** het aantal onderdelen aan waaraan een leerling heeft deelgenomen. Iets meer dan 40% van de leerlingen (1230 in totaal) heeft deelgenomen aan alle vier onderdelen. Het gaat dan om de vragenlijst, het examen Nederlands, het examen Engels en het examen wiskunde A of B. Naast de leerlingen die aan alle onderdelen hebben deelgenomen, zijn er 275 (9.1%) leerlingen die deelgenomen hebben aan één onderdeel, 487 (16.1%) die deelgenomen hebben aan twee onderdelen en 1032 (34.1%) die deelgenomen hebben aan drie onderdelen. Als een bepaald onderdeel voor een bepaalde leerling niet aanwezig is, betekent dat niet automatisch dat de leerling het betreffende onderdeel ook niet gemaakt heeft. Dit komt doordat de examenresultaten van de COOL⁵⁻¹⁸-leerlingen soms zeer moeilijk terug te vinden waren in de totale WOLF-database met daarin ook de examenresultaten van alle leerlingen die niet aan COOL⁵⁻¹⁸ deelnamen. Aan scholen is weliswaar nadrukkelijk gevraagd om bij het invoeren van de examenresultaten in WOLF exact dezelfde identificatie aan te houden als op de vragenlijsten, maar dit is niet altijd gebeurd. Als de door de school gekozen identificatie te veel afweek van de identificatie die in COOL⁵⁻¹⁸ gebruikt werd (en die door scholen zelf was opgegeven) kon de leerling niet teruggevonden worden in de database. Ook kunnen er door de inconsistenties in de identificatie van de leerlingen (op zeer beperkte schaal) bepaalde examenresultaten bij de verkeerde leerlingen terechtgekomen zijn.

Als een leerling deelgenomen heeft aan een onderdeel betekent dat niet automatisch dat deze leerling ook alle items van dat onderdeel beantwoord heeft. In het databestand geven de variabelen **MIS_NED**, **MIS_ENG**, **MIS_WIA** en **MIS_WIB** per leerling aan op hoeveel procent van de items een antwoord ontbreekt. Bij de berekening van het percentage ontbrekende antwoorden is uitgegaan van het aantal items in een examen en *niet* van het aantal scorepunten dat een leerling kon behalen. Dit betekent bijvoorbeeld dat bij het examen Nederlands is uitgegaan van 19 items (d.w.z. 18 open of gesloten antwoordvragen en 1 samenvattingsopdracht) en niet van de schaallengte van 52 scorepunten. Tabel 2.1 laat zien dat verreweg de meeste leerlingen bij de vakken Nederlands en Engels alle items beantwoord hebben. Bij meer dan 95 procent van de leerlingen is het percentage ontbrekende antwoorden gelijk aan 0. Het percentage leerlingen dat meer dan 5 procent van de items niet beantwoord heeft is bij deze vakken bovendien erg klein, namelijk minder dan 1 procent. Bij wiskunde A en wiskunde B is het percentage leerlingen met een ontbrekend antwoord aanzienlijk groter. Bij beide vakken heeft bijna een kwart van de leerlingen ten minste 1 item niet beantwoord. In sommige situaties kan het zinvol zijn om de leerlingen met relatief veel ontbrekende antwoorden buiten beschouwing te laten. Bij veel ontbrekende antwoorden kan er immers aan getwijfeld

worden of een leerling wel gemotiveerd heeft deelgenomen. Gezien het *high-stakes* karakter van de afname ligt dat bij de examens echter niet voor de hand. Bij de berekening van de cijfers en de transformatie daarvan in Z-scores zijn de items met een ontbrekend antwoord fout gerekend. Tabel 2.1 bevat geen informatie over het aantal ontbrekende antwoorden bij de leerlingvragenlijst. De reden hiervoor is dat dit aantal per onderdeel varieert en er bij het maken van de schaalvariabelen schaalspecifieke regels met betrekking tot ontbrekende waarden zijn gehanteerd. Zie hiervoor de informatie in Hoofdstuk 4.

Tabel 2.1

Percentage missing per onderdeel

Onderdeel	Percentage ontbrekende antwoorden							Totaal
	0	≤ 5	≤ 10	≤ 20	≤ 50	≤ 99	100	
Nederlands	2142	40	0	5	1	3	0	2191
Wiskunde A	1261	146	78	78	26	2	1	1592
Wiskunde B	555	0	70	79	31	1	0	736
Engels	2178	102	6	2	0	1	0	2289
Leerlingvragenlijst	----	----	----	----	----	----	----	2457

Zoals eerder aangegeven is geprobeerd om zoveel mogelijk leerlingen in de onderzoeksgroep op te nemen die eerder aan PRIMA en/of COOL⁵⁻¹⁸ hebben deelgenomen. Analyses lieten zien dat in totaal 770 leerlingen in de onderzoeksgroep deelgenomen hebben aan de meting van COOL-VO3 in schooljaar 2007/2008 (subgroep 3, zie paragraaf 2.1). Daarnaast bevat de onderzoeksgroep 73 leerlingen die niet alleen aan de meting van COOL-VO3 hebben deelgenomen, maar daarnaast ook drie jaar eerder aan het PRIMA cohortonderzoek hadden deelgenomen (subgroep 1). Ten slotte zijn er 330 leerlingen die wel aan het PRIMA cohortonderzoek hadden deelgenomen, maar in schooljaar 2007/2008 niet aan COOL⁵⁻¹⁸ (subgroep 2) hebben deelgenomen. Het totale aantal target leerlingen komt daarmee op 1173; een kleine 40 procent van de totale onderzoeksgroep. In de groep met target leerlingen zijn er 428 (36.5%) leerlingen die deelgenomen hebben aan alle vier onderdelen. Daarnaast zijn er 118 (10.1%) leerlingen die deelgenomen hebben aan één onderdeel, 195 (16.6%) leerlingen die deelgenomen hebben aan twee onderdelen en 406 (34.6%) leerlingen die aan drie onderdelen hebben deelgenomen. Van 26 target leerlingen zijn in het geheel geen gegevens beschikbaar. Deze leerlingen staan dan ook niet in het databestand. De target leerlingen kunnen via de variabele **LL_ID** gekoppeld worden aan het databestand van COOL-VO3 in schooljaar 2007/2008

Om de representativiteit te kunnen beoordelen is de onderzoeksgroep geanalyseerd in het perspectief van een aantal relevante achtergrondvariabelen. De representativiteit is geëvalueerd in relatie tot de volgende variabelen: (1) regionale spreiding, (2) mate van verstedelijking, (3) sekse, (4) leeftijd en (5) profielkeuze. Voor het bepalen van de populatiegegevens is gebruikgemaakt van gegevens van DUO. Concreet is uitgegaan van het bestand met het aantal examenkandidaten en geslaagden in schooljaar 2009/2010 zoals dat gedownload kan worden via www.data.duo.nl. Omdat niet alle leerlingen in de onderzoeksgroep hebben deelgenomen aan alle onderdelen zijn de analyses niet alleen voor de totale onderzoeksgroep verricht, maar ook voor elk van de onderdelen afzonderlijk. De beoordeling van de representativiteit per onderdeel in hierbij wel enigszins problematisch, omdat het niet zeker is dat de populatie voor elk onderdeel 'identiek' is. Met name bij de examens wiskunde verschilt de deelname sterk van de algehele populatie met examenkandidaten. Zo wordt wiskunde B veel meer door jongens dan door meisjes gekozen.

Het databestand met populatiegegevens van DUO bevat over dergelijke patronen geen informatie. De leerlingen die geen toestemming hebben gegeven voor het koppelen van hun gegevens aan andere databestanden (zie de variabele **TOESTEM**) zijn niet meegenomen in de representativiteitsanalyses (zie ook paragraaf 4.2.5). Hierdoor wijken de leerlingaantallen in de representativiteitsanalyse af van de eerder genoemde 3024 aan het begin van deze paragraaf.

Representativiteit naar regionale spreiding

Eerst is gekeken naar de spreiding van scholen en leerlingen over Nederland. Er is onderscheid gemaakt in vier regio's. Regio *Noord* omvat de provincies Groningen, Friesland en Drenthe; regio *Oost* de provincies Overijssel, Gelderland en Flevoland; regio *West* de provincies Utrecht, Noord-Holland, Zuid-Holland en Zeeland en regio *Zuid* de provincies Noord-Brabant en Limburg. De verdeling van scholen en leerlingen in de populatie en onze onderzoeksgroep naar regio staat in Tabel 2.2a en Tabel 2.2b.

Tabel 2.2a

Aantal en percentage scholen in de populatie en de onderzoeksgroep naar regio

Regio	Onderdeel						Populatie
	Totaal	ENG	NED	WIA	WIB	LV	
Noord	11 (20.0)	11 (20.0)	11 (20.4)	11 (21.2)	11 (24.4)	11 (23.9)	47 (9.7)
Oost	10 (18.2)	10 (18.2)	10 (18.5)	10 (19.2)	7 (15.6)	8 (17.4)	103 (21.2)
West	19 (34.5)	19 (34.5)	19 (35.2)	17 (32.7)	16 (35.6)	15 (32.6)	240 (49.5)
Zuid	15 (27.3)	15 (27.3)	14 (25.9)	14 (26.9)	11 (24.4)	12 (26.1)	95 (19.6)
Totaal	55	55	54	52	45	46	485

Tabel 2.2b

Aantal en percentage leerlingen in de populatie en de onderzoeksgroep naar regio

Regio	Onderdeel						Populatie
	Totaal	ENG	NED	WIA	WIB	LV	
Noord	1068 (35.4)	768 (33.6)	782 (35.7)	589 (37.0)	215 (29.2)	938 (38.2)	5015 (10.2)
Oost	695 (23.0)	521 (22.8)	477 (21.8)	315 (19.8)	178 (24.2)	533 (21.7)	10958 (22.3)
West	584 (19.3)	457 (20.0)	434 (19.8)	311 (19.5)	152 (20.7)	486 (19.8)	22479 (45.7)
Zuid	674 (22.3)	542 (23.7)	495 (22.6)	376 (23.6)	191 (26.0)	497 (20.3)	10774 (21.9)
Totaal	3021	2288	2188	1591	736	2454	49226

We zien dat de verdeling van scholen en leerlingen over de vier regio's in de onderzoeksgroep tamelijk sterk afwijkt van de verdeling in de populatie. Er hebben ten opzichte van de populatie tamelijk veel scholen en leerlingen uit het noorden meegedaan aan COOL⁵⁻¹⁸ in HAVO-5 en tamelijk weinig scholen en leerlingen uit het westen. Het aantal deelnemende scholen en leerlingen in het oosten en zuiden loopt wel redelijk in de pas met de populatie. Als we de totale onderzoeksgroep afzetten tegen de populatie dan zien we dat er zowel op schoolniveau als leerlingniveau sprake is van een significant verschil tussen onze onderzoeksgroep en de aantallen scholen en leerlingen die we op grond van de gegevens van DUO in elke regio mogen verwachten: (schoolniveau) $\chi^2 = 10.413$, $df = 3$, $p = 0.015$; (leerlingniveau) $\chi^2 = 2337.618$, $df = 3$, $p = 0.000$. In bepaalde analyses kan het dus wenselijk zijn om terug te wegen voor de variabele *regio*. Een alternatief kan zijn om de variabele *regio* expliciet in een model op te nemen. In het databestand geeft de variabele **SCH_REG** aan in welke provincie de school stond. De hierboven gehanteerde indeling in regio's is op basis van deze variabele eenvoudig te maken.

Representativiteit naar mate van verstedelijking

Vervolgens is de representativiteit van de onderzoeksgroep geanalyseerd in het licht van de achtergrondvariabele *mate van verstedelijking*. Er zijn vijf groepen onderscheiden: scholen of leerlingen in (1) niet verstedelijkt gebied, (2) weinig verstedelijkt gebied, (3) matig verstedelijkt gebied, (4) sterk verstedelijkt gebied, en (5) zeer sterk verstedelijkt gebied. De verdeling van scholen en leerlingen in de populatie en onze onderzoeksgroep naar mate van verstedelijking staat in Tabel 2.3a en Tabel 2.3b.

Tabel 2.3a

Aantal en percentage scholen in de populatie en de onderzoeksgroep naar mate van verstedelijking

Verstedelijking	Onderdeel						Populatie
	Totaal	ENG	NED	WIA	WIB	LV	
Niet	2 (3.6)	2 (3.6)	2 (3.7)	2 (3.8)	2 (4.4)	2 (4.3)	21 (4.3)
Weinig	11 (20.0)	11 (20.0)	11 (20.4)	11 (21.2)	9 (20.0)	8 (17.4)	79 (16.3)
Matig	16 (29.1)	16 (29.1)	15 (27.8)	16 (30.8)	14 (31.1)	15 (32.6)	113 (23.3)
Sterk	15 (27.3)	15 (27.3)	15 (27.8)	14 (26.9)	11 (24.4)	12 (26.1)	170 (35.1)
Zeer sterk	11 (20.0)	11 (20.0)	11 (20.4)	9 (17.3)	9 (20.0)	9 (19.6)	102 (21.0)
Totaal	55	55	54	52	45	46	485

Tabel 2.3b*Aantal en percentage leerlingen in de populatie en de onderzoeksgroep naar mate van verstedelijking*

Verstedelijking	Totaal	Onderdeel					Populatie
		ENG	NED	WIA	WIB	LV	
Niet	255 (8.4)	154 (6.7)	248 (11.3)	135 (8.5)	72 (9.8)	239 (9.7)	1783 (3.6)
Weinig	960 (31.8)	791 (34.6)	636 (29.1)	526 (33.1)	286 (38.9)	695 (28.3)	8276 (16.8)
Matig	735 (24.3)	627 (27.4)	557 (25.4)	472 (29.6)	160 (21.7)	633 (25.8)	12932 (26.3)
Sterk	598 (19.8)	392 (17.1)	425 (19.4)	245 (15.4)	136 (18.5)	512 (20.8)	18310 (37.2)
Zeer sterk	473 (15.7)	324 (14.2)	322 (14.8)	213 (13.4)	82 (11.1)	375 (15.3)	7925 (16.1)
Totaal	3021	2288	2188	1591	736	2454	49226

We zien dat er op schoolniveau sprake is van een lichte ondervertegenwoordiging van scholen in sterk verstedelijkt gebied en een lichte oververtegenwoordiging van scholen in weinig en matig verstedelijkt gebied. Op leerlingniveau is het beeld min of meer vergelijkbaar, maar enigszins extremer. De afwijking tussen de aantallen scholen in onze onderzoeksgroep en de aantallen scholen die verwacht mogen worden op grond van de gegevens van DUO zijn niet significant: $\chi^2 = 2.295$, $df = 4$, $p = 0.682$. Op leerlingniveau is er wel sprake van een significant verschil tussen de aantallen die we op grond van de gegevens van DUO mogen verwachten en de aantallen in onze onderzoeksgroep: $\chi^2 = 846.739$, $df = 4$, $p = 0.000$. Dit betekent dat het in bepaalde analyses wenselijk kan zijn om terug te wegen voor de variabele *mate van verstedelijking*. In het databestand geeft de variabele **SCH_STED** aan in hoeverre de school in verstedelijkt gebied stond. Merk op dat er in de analyses vanuit is gegaan dat de stedelijke graad voor de scholen en de leerlingen identiek is. Dit hoeft vanzelfsprekend niet per sé zo te zijn. Als een leerling ver van de school woont, kan de stedelijke graad van de woonomgeving van de leerling afwijken van de stedelijke graad van de schoolomgeving.

Representativiteit naar sekse

Na de beoordeling van de representativiteit van de onderzoeksgroep in het licht van de variabelen *regio* en *mate van verstedelijking* is gekeken naar de verdeling van jongens en meisjes. De verdeling van leerlingen in de populatie en onze onderzoeksgroep naar sekse staat in Tabel 2.4. Volgens de gegevens van DUO hebben er in schooljaar 2009/2010 iets meer meisjes dan jongens deelgenomen aan het HAVO-examen, namelijk 51.9 procent meisjes tegenover 48.1% jongens. In onze onderzoeksgroep zien we datzelfde beeld terug, met uitzondering van wiskunde B. Bij dit examen hebben we duidelijk meer gegevens van jongens dan van meisjes verkregen. De tegengestelde richting bij wiskunde B hoeft niet te wijzen op een gebrek aan representativiteit. Het is immers algemeen bekend dat in Nederland veel meer jongens dan meisjes wiskunde B kiezen. Een dergelijk patroon is echter niet af te leiden uit de gegevens van DUO, omdat het bestand alleen informatie geeft over het totale aantal examenkandidaten. Er wordt geen uitsplitsing gemaakt naar de afzonderlijke vakken. Als we de totale COOL⁵⁻¹⁸-onderzoeksgroep afzetten tegen de populatie dan zien we dat de aantallen jongens en meisjes in onze onderzoeksgroep niet significant afwijken van de aantallen die op grond van de gegevens van DUO verwacht mogen worden: $\chi^2 = 1.929$, df

= 1, $p = 0.150$. Er is met andere woorden geen reden om in analyses terug te wegen voor de variabele *seks*. In het databestand geeft de variabele **SEKSE** weer of de leerling een jongen of een meisje was.

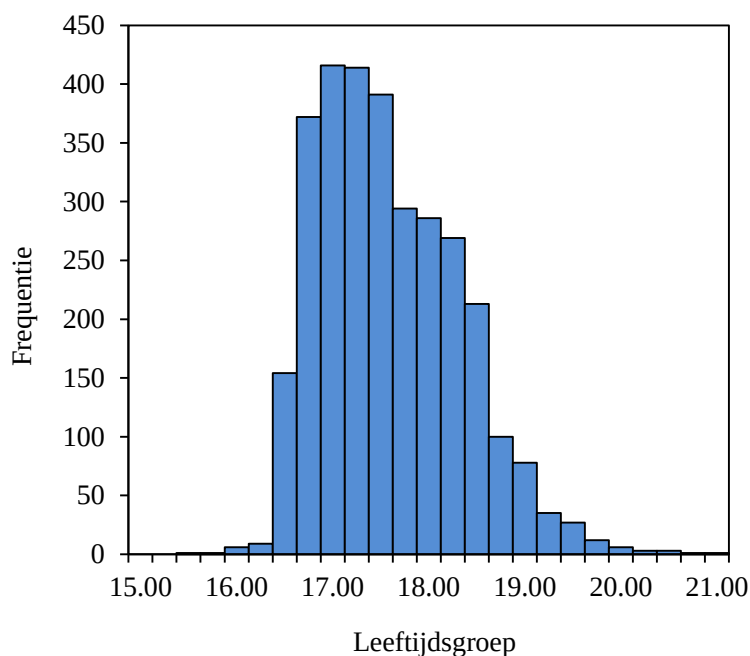
Tabel 2.4

Aantal en percentage leerlingen in de populatie en de onderzoeksgroep naar seks

Seks	Onderdeel						Populatie
	Totaal	ENG	NED	WIA	WIB	LV	
Jongen	1414 (46.8)	1114 (48.7)	993 (45.4)	684 (43.0)	492 (66.8)	1154 (47.0)	23662 (48.1)
Meisje	1607 (53.2)	1174 (51.3)	1195 (54.6)	907 (57.0)	244 (33.2)	1300 (53.0)	25564 (51.9)
Totaal	3021	2288	2188	1591	736	2454	49226

Representativiteit naar leeftijd

In de vierde stap is de representativiteit van de totale onderzoeksgroep geanalyseerd vanuit de achtergrondvariabele leeftijd. In het databestand bevat de variabele **LEEFTIJD** de leeftijd van de leerlingen op 1 maart 2010. Er is gekozen om bij de constructie van de variabele *leeftijd* uit te gaan van een vaste referentiedatum om zo het aantal variabelen in het databestand enigszins beperkt te houden. Bovendien lijkt het er voor de analyses niet veel toe te doen hoe oud de leerling exact was op het tijdstip van afname. Figuur 2.1 laat voor verschillende leeftijdsgroepen het aantal leerlingen in de onderzoeksgroep zien.



Figuur 2.1. *Leeftijdsverdeling onderzoeksgroep HAVO-5*

We zien dat de verdeling van de leeftijden enigszins rechtsscheef verdeeld is. Zoals verwacht mag worden heeft de meerderheid van de leerlingen een leeftijd die 'past' bij een onvertraagde schoolcarrière. De

overige leerlingen hebben ergens gedurende de schoolloopbaan een klas overgeslagen (1.0%) of vertraging opgelopen (43.0%). Onderzoek heeft laten zien dat de verhouding tussen de reguliere en vertraagde leerlingen in jaargroep 8 van het basisonderwijs rond de 80/20 ligt (zie bijvoorbeeld Heesters, van Berkel, van der Schoot, & Hemker, 2007). De COOL⁵⁻¹⁸-onderzoeksgroep lijkt qua leeftijdsopbouw dus min of meer representatief te zijn voor de populatie. Door zittenblijven zal het percentage vertraagde leerlingen in HAVO-5 immers groter zijn dan in groep 8 van het basisonderwijs. Als we kijken naar de leeftijdsopbouw bij de afzonderlijke onderdelen, dan zien we een vergelijkbaar beeld. Alleen bij het onderdeel wiskunde B ligt het percentage vertraagde leerlingen duidelijk iets lager (36.3%). Dit is niet zo verwonderlijk aangezien het vooral de goede leerlingen zijn die wiskunde B kiezen (en het vooral de minder goede leerlingen zijn die vertraging oplopen).

Representativiteit naar profiel

Ten slotte is de representativiteit van de onderzoeksgroep geanalyseerd in relatie tot de achtergrondvariabele *profiel*. Er zijn vijf groepen onderscheiden: (1) N&T – Natuur en Techniek, (2) N&G – Natuur en Gezondheid, (3) E&M – Economie en Maatschappij, (4) C&M – Cultuur en Maatschappij, en (5) Combinatieprofiel. Een leerling met een combinatieprofiel volgt meer dan één profiel, bijvoorbeeld zowel het profiel N&T als het profiel N&G. In het databestand geeft de variabele **OW_PROFIEL** aan voor welk profiel een leerling heeft gekozen. De verdeling van leerlingen in de populatie en onze onderzoeksgroep naar profiel staat in Tabel 2.5. De kolomtotalen wijken af van de kolomtotalen in de voorgaande tabellen, omdat niet van alle leerlingen de profielkeuze bekend is. We zien dat de verdeling van leerlingen over de verschillende profielen in onze onderzoeksgroep grote gelijkenis vertoont met de verdeling in de populatie. Statistisch gezien wijkt de verdeling in de onderzoeksgroep dan ook niet significant af van de verdeling die we op grond van de gegevens van DUO mogen verwachten: $\chi^2 = 7.956$, $df = 4$, $p = 0.093$.

Tabel 2.5

Aantal en percentage leerlingen in de populatie en de onderzoeksgroep naar profiel

Profiel	Totaal	Onderdeel					Populatie
		ENG	NED	WIA	WIB	LV	
C&M	480 (17.1)	333 (14.8)	405 (18.7)	180 (11.5)	1 (0.1)	369 (16.4)	8368 (17.0)
E&M	1158 (41.3)	924 (41.1)	870 (40.3)	966 (61.8)	79 (10.9)	911 (40.5)	21167 (43.0)
N&G	488 (17.4)	406 (18.1)	374 (17.3)	329 (21.0)	99 (13.7)	401 (17.8)	8861 (18.0)
N&T	308 (11.0)	271 (12.1)	234 (10.8)	0 (0.0)	289 (39.9)	253 (11.3)	4923 (10.0)
Combinatie	367 (13.1)	313 (13.9)	278 (12.9)	89 (5.7)	257 (35.4)	314 (14.0)	5907 (12.0)
Totaal	2801	2247	2161	1564	725	2248	49226

Als we kijken naar de profielkeuze van leerlingen binnen de afzonderlijke onderdelen, dan zien we grote verschillen ten opzichte van de profielkeuze in de populatie. Niet zozeer als het gaat om de (verplichte) onderdelen Nederland en Engels, maar wel als het gaat om de examens wiskunde. Dit is echter niet vreemd. De examens wiskunde zijn immers rechtstreeks gekoppeld aan een bepaald profiel. In het HAVO doen de leerlingen die gekozen hebben voor het profiel N&T verplicht eindexamen in wiskunde B. De leerlingen die voor een ander profiel hebben gekozen doen in principe eindexamen in wiskunde A, maar de leerlingen mogen ook kiezen voor wiskunde B. Een uitzondering hierop vormen de leerlingen die het profiel C&M hebben gekozen; de C&M-leerlingen zijn niet verplicht om eindexamen in wiskunde te doen. Het is zeer opmerkelijk dat er 1 leerling in het databestand zit die naar eigen opgave gekozen heeft voor het profiel C&M en toch eindexamen heeft gedaan in wiskunde B. Mogelijk is hier sprake van een invulfout. C&M-leerlingen hebben immers alleen wiskunde A als keuzemogelijkheid.

Samenvattend laten de representativiteitsanalyses zien dat de totale COOL⁵⁻¹⁸-onderzoeksgroep ($N = 3024$) voldoende representatief is voor de populatie als het gaat om de variabelen *seks*, *leeftijd* en *profiel*. Dit betekent dat het niet nodig is om in analyses te controleren of terug te wegen voor deze variabelen. Dat geldt niet voor de variabelen *regio* en *mate van verstedelijking*. Bij deze variabelen zien we namelijk dat de aantallen in de onderzoeksgroep wel significant afwijken van de aantallen die we mogen verwachten op grond van gegevens van DUO. Afhankelijk van de onderzoeksvraag en de analyse kan het dus wenselijk zijn om deze variabelen expliciet mee te nemen.

3 EXAMENGEGEVENS

3.1 Nederlands

3.1.1 Kenmerken

Het eerste tijdvak examen HAVO Nederlands bevatte 18 items en een samenvattingsopdracht. De samenvattingsopdracht werd op 10 aspecten beoordeeld. In totaal gaat het dus om 28 items. Bij 7 items moesten leerlingen het juiste antwoord kiezen uit vier alternatieven. De overige items hadden de vorm van een open-antwoordvraag. In het examen kwamen verschillende deeltaarigheden aan de orde, namelijk (a) analyseren en interpreteren, (b) beargumenteren en (c) samenvatten. Hieronder staan twee voorbeelden uit het examen. Het eerste voorbeeld (item 2) heeft betrekking op de deeltaarigheid analyseren en interpreteren. In het tweede voorbeeld (item 3) gaat het om de deeltaarigheid beargumenteren. Het volledige HAVO examen Nederlands is te vinden op www.cito.nl. Kies op de website achtereenvolgens: voortgezet onderwijs – centrale examens – schriftelijke examens havo/vwo – examens 2010 – havo 1e tijdvak – Nederlands. Op de website is ook het tekstboekje te vinden dat bij het examen Nederlands hoort (en waar in de voorbeelditems naar wordt verwezen).

Tekst 1 Hoera! Een nieuw klimaat

tekstfragment 1

Naar verwachting zullen klimaatveranderingen in de tweede helft van deze eeuw en daarna doorgaan. De thans optredende klimaatverandering werkt nog eeuwen door, vooral in de zeespiegelstijging (uitzetting door temperatuurverhoging). Op die zeer lange termijn worden zeespiegelstijgingen van enkele tot vele meters verwacht. Laag-Nederland krijgt hier steeds meer last van. Het is de vraag of met conventionele technieken de veiligheid op het huidige niveau kan worden gehouden. De opgave wordt versterkt door de combinatie met mogelijk verder verhoogde piekafvoeren van de Rijn en de Maas en een voortgaande bodemdaling.

bron: www.pbl.nl

“Intussen worden rondom hem de dijken verhoogd en zal een beperkt aantal mensen moeten verhuizen om het hoofd boven het water te houden. Slim waterbeheer zorgt voor een doeltreffende afvoer van rivierwater en er komen meer gebieden waar overtollig rivierwater naartoe wordt geloodst.”
(regels 75-83)

- 2 **Welke kritische vraag kan een lezer bij de maatregelen in het citaat uit de hoofdttekst stellen gelet op tekstfragment 1?**
Gebruik voor je antwoord maximaal 25 woorden.

De gedachtegang van alinea 4 berust op het volgende argumentatieschema:
bewering – tegenwerping op de bewering – weerlegging van de tegenwerping.
De bewering is: “Er zijn geen eenduidige aanwijzingen dat we aan gezondheid gaan inleveren (...).” (regels 95-97)

- 3 **Formuleer de tegenwerping op de genoemde bewering.**
Gebruik maximaal 10 woorden.

Er is geen voorbeeld opgenomen van de deelvaardigheid *samenvatten*. Dit onderdeel nam ten opzichte van de andere onderdelen in het examen dan ook een aparte positie in. Leerlingen moesten namelijk een goedlopende samenvatting in correct Nederlands schrijven van een bepaalde tekst. Bij het beoordelen van de samenvattingsopdracht werd speciaal gelet op het taalgebruik en het aantal woorden. Er werden punten in mindering gebracht als de samenvatting te lang was of incorrecte formuleringen bevatte. Bij de overige onderdelen moesten leerlingen vragen beantwoorden over een tekst. Soms konden leerlingen het juiste antwoord kiezen uit een viertal antwoordalternatieven. In andere gevallen moesten leerlingen zelf een antwoord formuleren. Leerlingen konden maximaal 52 punten behalen op het examen. Tabel 3.1 laat zien hoe het HAVO examen Nederlands precies was opgebouwd.

Tabel 3.1

Aantal items in het HAVO examen uitgesplitst naar vraagtype en vraagvorm

Vraagtype	Open	Gesloten	Totaal
Analyseren en interpreteren	8	5	13
Argumentatieve vaardigheden	3	2	5
Samenvatten	10	0	10
Totaal	21	7	28

In het examen nam de samenvattingsopdracht een belangrijke plaats in. Ongeveer één derde van de eindscore werd namelijk bepaald door de prestatie op deze opdracht. Hoewel de vaardigheid *samenvatten* thuishoort in het examen, maakt het type opdracht de vergelijking met andere toetsresultaten van het vak Nederlands wel moeilijker. De toetsen Nederlands Begrijpend Lezen die bij de meting in leerjaar 3 gebruikt zijn, bevatten bijvoorbeeld uitsluitend korte teksten met gesloten vragen (zie Zijlsing et al., 2009). Het is goed mogelijk dat leerlingen anders presteren op een dergelijke toets tekstbegrip dan op een examen dat bestaat uit twee lange teksten en een samenvattingsopdracht. De lengte van de teksten kan bijvoorbeeld van invloed zijn op het eindresultaat. Daarnaast wordt bij het examen niet uitsluitend een beroep gedaan op leesvaardigheid, maar ook op schrijfvaardigheid. Om deze reden is in het databestand ook een variabele opgenomen die de score op het examen Nederlands weergeeft zonder de samenvattingsopdracht. Bij het bepalen van deze score zijn dus alleen de eerste 18 items meegenomen. De variabele heeft de naam **NED_RSCO1**. De variabele waarin de samenvattingsopdracht wel is meegeteld heeft de naam **NED_RSCO2**.

Op basis van een steekproef van examenleerlingen zijn de psychometrische eigenschappen van de examens Nederlands in kaart gebracht. De afnamegegevens van de leerlingen in de steekproef zijn via het computerprogramma WOLF naar Cito gestuurd. Elke school is verplicht om de gegevens van de alfabetisch eerste vijf (of soms tien) leerlingen in te sturen. Veel scholen kiezen er echter voor om de afnamegegevens van al hun examenleerlingen in te sturen. De steekproef is gebaseerd op alle voor een bepaalde datum aangeleverde afnamegegevens. In Tabel 3.2 worden de belangrijke eigenschappen op een rij gezet. De gemiddelde *p*-waarde in Tabel 3.2 is een getal tussen 0 en 1 dat de gemiddelde moeilijkheidsgraad van het examen weergeeft. Het is de gemiddelde toetsscore gedeeld door het maximaal haalbare aantal punten. De betrouwbaarheid is uitgedrukt in termen van de asymptotische *greatest lower bound* (GLB; Ten Berge & Sočan, 2004). Hoewel ook deze maat voor betrouwbaarheid een onderschatting is van de werkelijke betrouwbaarheid, ligt hij wel dicht bij de werkelijke betrouwbaarheid dan Cronbach's alpha. De standaardmeetfout, ten slotte, is een indicatie voor de onnauwkeurigheid van een meting. De grootte van de standaardmeetfout van een toets hangt af van de betrouwbaarheid en de standaarddeviatie van de

toetsscores en kan gebruikt worden om een betrouwbaarheidsinterval voor de ware score van een leerling te construeren. De resultaten uit Tabel 3.2 zijn ook terug te vinden op www.cito.nl. Kies op de website achtereenvolgens: voortgezet onderwijs – centrale examens – examenverslagen – oude verslagen – 2010 – resultaten per examen – toelichting / applicatie.

Tabel 3.2

Eigenschappen 1^{ste} tijdvak examen HAVO Nederlands

Eigenschap	HAVO
Totaal aantal leerlingen	50169
Steekproefgrootte	32443
Aantal items	28
Gemiddelde <i>p</i> -waarde	0.47
Minimumscore	0
Maximumscore	52
Gemiddelde score	24.7
Gemiddeld cijfer	6.0
Percentage onvoldoendes (<5.5)	28.0
Standaarddeviatie	5.8
Betrouwbaarheid (asyp. GLB)	0.43
Standaardmeetfout	4.3

3.1.2 Scoretabel

In Nederland worden de examenresultaten van leerlingen uitgedrukt in een cijfer op een schaal van 0 tot 10. De cijfers komen tot stand na de normering van het betreffende examen door het CvE (College voor Examens). Die normering is mede gebaseerd op de door de scholen aan Cito gerapporteerde scores van duizenden leerlingen. De ruwe scores van de leerlingen en de daarmee corresponderende cijfers kunnen ook vanuit andere benaderingen beschreven en geïnterpreteerd worden. Het gaat dan om de relatieve positie van een score of cijfer in de populatie. In de scoretabel in Bijlage 1 wordt de relatieve positie van een cijfer op het HAVO-examen Nederlands op verschillende manieren tot uitdrukking gebracht, namelijk in een percentiele rang (perc.), een *Z*-score, en een *T*-score. De percentiele rang bij een bepaald cijfer (of eigenlijk bij een bepaalde ruwe score) op een examen is gelijk aan het verwachte percentage leerlingen in de gehele populatie van leerlingen dat hetzelfde of een lager cijfer heeft behaald. Een percentiele rang van 29 betekent dus dat 29 procent van de examenleerlingen een lager cijfer heeft behaald. De percentielen voor de totale populatie zijn geschat op basis van alle leerlingen in de steekproef. Bij het HAVO examen Nederlands gaat het in totaal dus om 32443 leerlingen. Hoewel een percentiele rang eenvoudig te interpreteren is, kleven er ook nadelen aan. Bij de cijfers of scores rond het gemiddelde worden kleine verschillen in prestatie in de percentielscores namelijk uit elkaar getrokken. *Z*-scores en *T*-scores hebben daar geen last van.

Een *Z*-score geeft aan in welke mate een cijfer van het gemiddelde afwijkt. Het gemiddelde van de *Z*-scores is 0. Een cijfer dat als *Z*-score 0 heeft komt dus overeen met het gemiddelde cijfer. Cijfers met een negatieve *Z*-score liggen onder het gemiddelde en cijfers met een positieve *Z*-score liggen boven het gemiddelde. *Z*-scores zijn in de regel geen gehele getallen en bovendien is ongeveer de helft van de voorkomende scores negatief. Om dat ongemak te verhelpen worden *Z*-scores ook wel naar *T*-scores getransformeerd. *T*-scores zijn net als *Z*-scores gestandaardiseerde scores. Ze hebben alleen een ander

gemiddelde en een andere standaardafwijking. Het gemiddelde van de *T*-scores bedraagt altijd 50 en de standaardafwijking is 10. In het databestand zijn geen percentielen of *T*-scores opgenomen. Deze kunnen met behulp van Bijlage 1 eenvoudig bepaald worden. De ruwe scores, de cijfers en de bijbehorende *Z*-scores staan wel in het databestand. De variabelen hebben respectievelijk de naam **NED_RSCO1**, **NED_RSCO2**, **NED_CFR** en **NED_ZSCO**. Bijlage 1 is ook terug te vinden op www.cito.nl. Kies op de website achtereenvolgens: voortgezet onderwijs – centrale examens – examenverslagen – oude verslagen – 2010 – betekenis van cijfers – toelichting / applicatie.

3.2 Wiskunde

3.2.1 Kenmerken

Bij de overgang van leerjaar 3 naar leerjaar 4 kiezen leerlingen voor een bepaald profiel. In 2010 konden leerlingen uit een viertal profielen kiezen, namelijk uit Cultuur en Maatschappij (C&M), Economie en Maatschappij (E&M), Natuur en Gezondheid (N&G), en Natuur en Techniek (N&T). De leerlingen die gekozen hebben voor het profiel E&M of N&G doen in principe eindexamen in wiskunde A. De leerlingen die gekozen hebben voor het profiel N&T doen eindexamen in wiskunde B. De C&M-leerlingen maken geen examen wiskunde. De examens wiskunde bevatten uitsluitend open-antwoordvragen. Er zijn verschillende deelvaardigheden onderscheiden. Tabel 3.3 laat zien hoe de verschillende examens wiskunde waren opgebouwd. We zien dat het onderdeel *kansrekening en statistiek* in de wiskunde A examens een belangrijke plaats innam. In de wiskunde B examens kwam dit onderdeel niet aan bod. Daar werd aandacht besteed aan algebra en meetkunde.

Tabel 3.3

Aantal items in de HAVO examens wiskunde uitgesplitst naar vraagtype

Vraagtype	Wiskunde A	Wiskunde B
Analyse	13	9
Algebra	0	3
Kansrekening en statistiek	10	0
Meetkunde	0	5
Totaal	23	17

Hieronder staat van elke deelvaardigheid een voorbeelditem. De voorbeelden hebben achtereenvolgens betrekking op algebra, *kansrekening en statistiek*, *meetkunde* en *analyse*. De volledige HAVO examens wiskunde zijn te vinden op www.cito.nl. Kies op de website achtereenvolgens: voortgezet onderwijs – centrale examens – schriftelijke examens havo/vwo – examens 2010 – havo 1e tijdvak – wiskunde A / wiskunde B.

Snijpunt (wiskunde B)

De functies f en g zijn gegeven door $f(x) = 2^{4x+1}$ en $g(x) = 4 \cdot 4^x$.

- 13 Bereken op algebraïsche wijze de coördinaten van het snijpunt van de grafiek van f en g .**

Een tenniswedstrijd (wiskunde A)

De finale in het herenenkelspel van het tennistoernooi Australian Open van 2007 ging tussen de tennissers Roger Federer en Fernando Gonzalez. Roger Federer won. Hij speelde al negen keer eerder tegen Fernando Gonzalez en al die wedstrijden won hij.

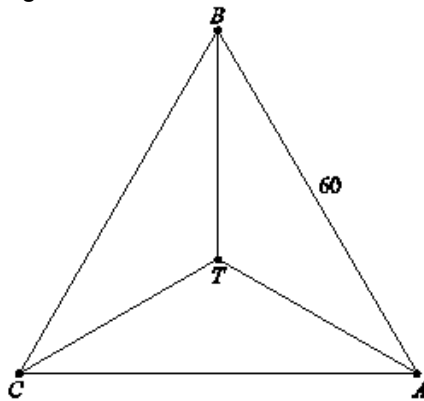
Wanneer twee spelers met hetzelfde krachtsverschil als Federer en Gonzalez tegen elkaar spelen, is de kans dat de sterkste de wedstrijd wint, gelijk aan 0,94. We bekijken 10 van dergelijke wedstrijden.

- 1 Bereken de kans dat de sterkste speler in die 10 wedstrijden 10 keer van de andere speler wint.

Tetraëder van Bottrop (wiskunde B)

Een regelmatige tetraëder is een viervlak waarvan de vier grensvlakken de vorm van een gelijkzijdige driehoek hebben. In figuur 1 is een bovenaanzicht van de regelmatige tetraëder $ABC.T$ te zien. Hierin is ABC het grondvlak en T de top. Er geldt $AB = 60$.

figuur 1



De lengte van CT in het bovenaanzicht van figuur 1 is ongeveer 35.

- 5 Bereken exact de lengte van CT in het bovenaanzicht van figuur 1.

Diersoorten (wiskunde B)

Uit onderzoek is gebleken dat er een verband bestaat tussen de lengte van diersoorten en het aantal diersoorten met die lengte. Met de lengte van een diersoort wordt bedoeld de gemiddelde lengte van volwassen dieren van die soort. Het blijkt dat er weinig lange diersoorten zijn en veel korte diersoorten. Uit gegevens die de onderzoeker Dobson verzamelde, blijkt dat bij benadering de volgende formule geldt:

$$S = \frac{700}{L^2}$$

Hierin is L de lengte in meter en S het aantal diersoorten met die lengte. Deze formule geldt voor $0,01 \leq L \leq 10$. Het aantal diersoorten van 10 cm lang is veel groter dan het aantal diersoorten van 50 cm lang.

foto



1 Bereken met behulp van de bovenstaande formule hoeveel maal zo groot.

De voorbeelditems laten zien dat het wiskunde A examen inhoudelijk tamelijk fundamenteel verschilt van het wiskunde B examen. Ook de toetsen wiskunde die bij een eerdere meting in leerjaar 3 gebruikt zijn, wijken inhoudelijk behoorlijk af van de eindexamens (zie Zijssling et al., 2009). Zo concentreren de toetsen voor leerjaar 3 zich op de rekenvaardigheid van leerlingen, terwijl het in de eindexamens ook om andere vaardigheden gaat. Ook de vraagvorm kan van invloed zijn op het eindresultaat. In de toetsen voor leerjaar 3 moeten leerlingen altijd het juiste antwoord kiezen uit 3 of 4 alternatieven, terwijl zowel het wiskunde A als het wiskunde B examen alleen open-antwoordvragen bevatten. Door de inhoudelijke verschillen kan het problematisch zijn om de resultaten van leerlingen op de verschillende wiskundetoetsen met elkaar te vergelijken. Het is immers goed mogelijk dat de prestaties van leerlingen mede bepaald worden door het type items en/of de vraagvorm.

De items in de examens werden polytoom gescoord. Bij het examen wiskunde A konden leerlingen maximaal 82 punten behalen en bij het examen wiskunde B maximaal 77 punten. Net zoals bij het examen Nederlands zijn op basis van een steekproef van examenleerlingen de psychometrische eigenschappen in kaart gebracht. In Tabel 3.4 worden de belangrijke eigenschappen van de examens wiskunde op een rij gezet. De resultaten uit Tabel 3.4 zijn ook terug te vinden op www.cito.nl. Kies op de website achtereenvolgens: voortgezet onderwijs – centrale examens – examenverslagen – oude verslagen – 2010 – resultaten per examen – toelichting / applicatie.

Tabel 3.4*Eigenschappen 1^{ste} tijdvak examens HAVO wiskunde A en B*

Eigenschap	Wiskunde A	Wiskunde B
Totaal aantal leerlingen	32030	12770
Steekproefgrootte	26698	10949
Aantal items	23	17
Gemiddelde <i>p</i> -waarde	0.64	0.49
Minimumscore	0	0
Maximumscore	82	77
Gemiddelde score	52.5	37.4
Gemiddeld cijfer	6.3	5.9
Percentage onvoldoendes (<5.5)	27.0	38.0
Standaarddeviatie	12.8	11.4
Betrouwbaarheid (asympt. GLB)	0.85	0.77
Standaardmeetfout	4.9	5.5

3.2.2 Scoretabel

In het databestand zijn voor elke wiskundevariant drie variabelen opgenomen. De ruwe score heeft de naam **WL._RSCO**, het cijfer heeft de naam **WL._CFR** en de bijbehorende Z-score heeft de naam **WL._ZSCO**. Op de puntjes staat de wiskundevariant. In geval van wiskunde A worden de ruwe scores dus gegeven door de naam WIA_RSCO. Net als bij het examen Nederlands kunnen de examenresultaten van leerlingen ook op andere manieren uitgedrukt worden. In Bijlage 2a en 2b wordt de relatieve positie van leerlingen op de HAVO examens wiskunde A en B tot uitdrukking gebracht in een percentiele rang en een *T*-score. In paragraaf 3.1.2 is uitgelegd wat de verschillende scores precies betekenen. Bijlage 2a en 2b zijn ook terug te vinden op www.cito.nl. Kies op de website achtereenvolgens: voortgezet onderwijs – centrale examens – examenverslagen – oude verslagen – 2010 – betekenis van cijfers – toelichting / applicatie.

3.3 Engels

3.3.1 Kenmerken

Het eerste tijdvak examen Havo Engels bevatte 11 verschillende teksten en 44 items. Bij de meerderheid van de items moesten leerlingen het juiste antwoord kiezen uit drie, vier of vijf antwoordalternatieven. De overige items hadden een iets afwijkende vorm. Soms moesten leerlingen zelf een antwoord formuleren, terwijl leerlingen in andere gevallen een antwoord moesten geven volgens een bepaalde structuur. Hieronder staat een voorbeeld van een dergelijk item. In het examen is geen onderscheid gemaakt in deelvraagdigheden. Leerlingen konden maximaal 49 punten behalen op het examen. Bij enkele items konden leerlingen dus meer dan 1 punt behalen. Het volledige HAVO examen Engels is te vinden op www.cito.nl. Kies op de website achtereenvolgens: voortgezet onderwijs – centrale examens – schriftelijke examens havo/vwo – examens 2010 – havo 1e tijdvak – Engels.

Tekst 7 Roll up, roll up, and watch the Mona Lisa weep

- 31 Geef van elk van de volgende beweringen aan of deze wel of niet overeenkomt met de inhoud van de alinea's 4 tot en met 6.
- 1 De bestseller *The Da Vinci Code* heeft veel mensen gestimuleerd zich grondig te verdiepen in de schilderkunst.
 - 2 De kwaliteit van beroemde kunstwerken is meer dan eens beneden de maat.
 - 3 Grote groepen mensen bezoeken bekende kunstwerken meer vanwege hun populariteit dan vanwege hun artistieke belang.
 - 4 Om schade te voorkomen zouden alleen nog kunstkeners absolute topstukken mogen bezichtigen.

Noteer het nummer van elke bewering, gevolgd door "wel" of "niet".

Net zoals bij de examens Nederlands en wiskunde zijn de psychometrische eigenschappen in kaart gebracht op basis van een steekproef van examenleerlingen. In Tabel 3.5 worden de belangrijke eigenschappen van het examens Engels op een rij gezet. De resultaten uit Tabel 3.5 zijn ook terug te vinden op www.cito.nl. Kies op de website achtereenvolgens: voortgezet onderwijs – centrale examens – examenverslagen – oude verslagen – 2010 – resultaten per examen – toelichting / applicatie.

Tabel 3.5

Eigenschappen 1^{ste} tijdvak examen HAVO Engels

Eigenschap	Havo
Totaal aantal leerlingen	50168
Steekproefgrootte	32123
Aantal items	44
Gemiddelde <i>p</i> -waarde	0.63
Minimumscore	0
Maximumscore	49
Gemiddelde score	31.0
Gemiddeld cijfer	6.1
Percentage onvoldoendes (<5.5)	33.0
Standaarddeviatie	7.5
Betrouwbaarheid (asympt. GLB)	0.85
Standaardmeetfout	2.9

3.3.2 Scoretabel

In het databestand zijn voor Engels drie variabelen opgenomen. De ruwe score heeft de naam **ENG_RSCO**, het cijfer heeft de naam **ENG_CFR** en de bijbehorende Z-score heeft de naam **ENG_ZSCO**. Net als bij de examens Nederlands en wiskunde kunnen de examenresultaten van leerlingen ook op andere manieren uitgedrukt worden. In Bijlage 3 wordt de relatieve positie van leerlingen op het HAVO examen Engels tot uitdrukking gebracht in een percentiele rang en een *T*-score. In paragraaf 3.1.2 is uitgelegd wat de verschillende scores precies betekenen. Bijlage 3 is ook terug te vinden op www.cito.nl. Kies op de website achtereenvolgens: voortgezet onderwijs – centrale examens – examenverslagen – oude verslagen – 2010 – betekenis van cijfers – toelichting / applicatie.

4 LEERLINGVRAGENLIJST

4.1 Inleiding

De in de leerlingvragenlijst opgenomen variabelen dienen ter operationalisatie van de in het onderzoeksvoorstel opgenomen concepten. In volgorde van verschijnen in de vragenlijst betreft het de volgende concepten: tijdbesteding, schoolwerk in vakanties, schoolverzuim, bijles, examenvorbereiding, slaagverwachting, examenvaardigheden, welbevinden met klasgenoten, persoonlijkheid, vrienden/vriendinnen, en waargenomen autonomie (opvoedstijl ouders). De uitwerking van deze concepten komt respectievelijk in paragrafen 4.3.1 tot en met 4.3.11 ter sprake. In deze paragrafen worden ook de bronnen vermeld, waarnaar bij gebruik van de betreffende concepten verwezen dient te worden. Daarvoor worden in paragraaf 4.2 enkele procedures toegelicht. Naast de in het onderzoeksvoorstel opgenomen concepten heeft het GION ook zelf een aantal concepten in de vragenlijst opgenomen. De betreffende variabelen zijn niet in het COOL⁵⁻¹⁸ bestand opgenomen en komen derhalve ook niet in dit rapport aan de orde.

4.2 Procedures

De door de scholen naar het Cito geretourneerde vragenlijsten zijn daar optisch ingelezen, volgens de wensen van het GION. Deze wensen betroffen het omgaan met dubbele antwoorden. Hoewel in de invulinstructie expliciet stond vermeld dat er per vraag slechts één antwoord mocht worden gegeven (tenzij anders aangegeven), is het de ervaring van het GION dat ook dan sommige respondenten bij sommige vragen meer dan één antwoord aankruisen (of omcirkelen). In verreweg de meeste van deze gevallen gaat het daarbij om twee aanliggende antwoorden op Likertschaal-achtige items. De enige zinvolle interpretatie van dergelijke antwoorden is dat de respondent heeft willen aangeven dat zijn/haar antwoord ‘er tussen in’ ligt. Aangezien dergelijke antwoordpatronen zinvol verwerkt kunnen worden, zijn ze ook met unieke codes ingelezen. Deze door het GION geprefereerde aanpak wijkt af van de standaardbenadering bij optisch inlezen – en veelal ook bij handmatige invoer – waarbij dubbele antwoorden hetzij direct gelijk worden gesteld aan ‘geen antwoord’ of ‘fout antwoord’, hetzij één algemene code voor ‘dubbel antwoord’ krijgen toegekend. Na het inlezen zijn de data als een EXCEL-bestand aan het GION overgedragen. De verdere werkzaamheden zijn door het GION verricht. Deze werkzaamheden betreffen het omzetten naar SPSS, het toekennen van variabele namen en het ‘labelen’ van de antwoordcategorieën, het hercoderen van met name de coderingen die een dubbel antwoord representeren, het definiëren van ontbrekende datacoderingen, en het aanmaken van schaalscores.

4.2.1 Invulinstructie

De vragenlijst begon met de volgende invulinstructie: *“Je hebt vast wel eens eerder een vragenlijst ingevuld, bijvoorbeeld twee jaar geleden voor ditzelfde onderzoek. Dit is dan een vervolgmeting. Het kan ook zijn dat je twee jaar geleden niet hebt meegedaan. Het is belangrijk dat je de vragenlijst zo goed mogelijk invult. Neem rustig de tijd voor het lezen en beantwoorden van de vragen. Bij de meeste vragen moet je één hokje aankruisen, liefst met een zwart potlood. Bij sommige vragen mag je meer dan één hokje aankruisen. Dat staat er dan bij. Er zijn ook vragen waarbij je een getal moet invullen. Dan staan er twee of drie vakjes. Gebruik voor een antwoord met één getal alleen het rechter vakje. Is je antwoord*

bijvoorbeeld '7' vul dit dan op deze wijze in: (gevolgd door voorbeeld van een 7 ingevuld in een tweevaks ruimte). Als je je hebt vergist, gum je het foute antwoord uit en kruis je een ander hokje aan, of schrijf je een ander antwoord op. Vul voor je met de vragen begint eerst de datum op de voorkant in. Alle vragen in deze vragenlijst zijn 'vertrouwelijk'. Dat betekent dat de onderzoekers er zeer zorgvuldig mee omgaan en dat niemand anders de antwoorden te zien krijgt. Stop daarom wanneer je klaar bent de ingevulde vragenlijst in bijgaande envelop en plak deze daarna dicht."

4.2.2 Ontbrekende waarden

Welke codering voor ontbrekende waarden is gebruikt, hangt in de eerste plaats af van het aantal antwoordmogelijkheden op het betreffende item. Als dat aantal hoogstens zeven bedraagt, is in principe code 9 gebruikt; anders code 99 of 999, afhankelijk van de hoogste antwoordmogelijkheid. In een aantal gevallen, bijvoorbeeld als het om een 'schaal' gaat, horen de items zodanig bij elkaar dat het zinvol is om onderscheid te maken tussen leerlingen die één of enkele items niet hebben beantwoord en leerlingen die geen enkel item hebben beantwoord, dat wil zeggen de hele vraag niet hebben beantwoord. In deze gevallen geldt: 9, 99 of 999 = hele vraag niet beantwoord; 8, 98 of 998 = specifiek item niet beantwoord.

4.2.3 Dubbele coderingen

Zoals al aangegeven zijn dubbele antwoorden specifiek, ofwel 'als zodanig', ingelezen, en kan er onderscheid worden gemaakt tussen 'zinvol dubbel' en 'anders dubbel'. De 'zinvol dubbele' antwoorden betreffen aankruisingen van twee naast elkaar liggende antwoordcategorieën in geval van Likertschaal-achtige items. De coderingen die dergelijke antwoordpatronen aangeven zijn gehercodeerd naar de waarde die de respondent redelijkerwijs bedoeld heeft. Indien bijvoorbeeld de 4 en 5 beide zijn aangekruist, is dat ingelezen als 4,5, en gehercodeerd naar 4,5. In de overige gevallen, indien twee niet aanliggende antwoordcategorieën zijn aangekruist, of wanneer er meer dan twee categorieën zijn aangekruist, is de betreffende dubbelcodering gehercodeerd naar 7, of eventueel naar 97 of 997, welke codes bijvoorbeeld zijn gelabeld als 'overig' en als ontbrekend zijn gedeclareerd.

4.2.4 Constructie schaalscores

Voor de diverse in de vragenlijst opgenomen itemsets die elk één of meerdere onderliggende concepten beogen te meten, is nagegaan hoeveel items zijn beantwoord. Tevens is per schaal een betrouwbaarheidsanalyse uitgevoerd. Een beslispunt bij het aanmaken van de schaalscores is het maximale aantal items dat eventueel niet beantwoord mag zijn. In de meest strikte opvatting moeten alle items van een bepaalde schaal zijn beantwoord, voordat als schaalscore het gemiddelde over de items mag worden berekend. Dit leidt ertoe dat relatief veel leerlingen een ontbrekende schaalscore moeten krijgen toegekend. In de praktijk is het daarom gebruikelijk om soepeler te zijn, door bijvoorbeeld één of enkele ontbrekende itemscores toe te staan. Wij hebben getracht een redelijk evenwicht te vinden tussen de volgende twee wenselijkheden: enerzijds optimale nauwkeurigheid, anderzijds een schaalscore toekennen aan zo veel mogelijk leerlingen. Het spreekt voor zich dat bepaalde negatief geformuleerde items eerst gehercodeerd (gespiegeld) zijn, voordat het itemgemiddelde is berekend.

4.2.5 Weigeraars

Aan het eind van de leerlingvragenlijst werd allereerst opgemerkt "Wij danken je hartelijk voor het invullen van deze vragenlijst!" en daaronder: "Het is de bedoeling dat de antwoorden die je hebt gegeven met behulp van jouw 'onderwijsnummer' gekoppeld gaan worden aan andere gegevens, vooral die van jouw eerdere deelname aan het onderzoek (indien van toepassing) en gegevens over jouw schoolcarrière. Als je

niet wilt dat jouw antwoorden gekoppeld gaan worden, kan je een e-mail sturen naar cool@cito.nl met als onderwerp: koppeling HAVO-5.”

Er zijn op het Cito drie van dergelijke meldingen binnengekomen. De vragenlijstgegevens van deze leerlingen komen wel voor in de resultaten die in de volgende paragrafen worden gepresenteerd en ook in het vragenlijstbestand, maar het is onmogelijk gemaakt de gegevens aan andere databronnen te koppelen. Het gevolg hiervan is dat ook de examengegevens van deze leerlingen niet aan andere databronnen kunnen worden gekoppeld. In het databestand geeft de variabele **TOESTEM** aan of de leerling (passief) toestemming heeft gegeven om zijn of haar gegevens te koppelen aan andere databestanden.

4.3 Systematische beschrijving leerlingvragenlijst

De vragenlijst was ontworpen om optisch te worden ingelezen. De antwoordmogelijkheden waren vierkante hokjes die moesten worden aangekruist of waar soms een getal (bijvoorbeeld een aantal uren per week) moest worden ingevuld. Er werd dus geen gebruik gemaakt van cijfers bij de hokjes van de antwoordmogelijkheden. In de volgende sub-paragrafen maken we echter wel gebruik van cijfers om de antwoordmogelijkheden mee aan te duiden. Dit zijn tevens de coderingen die in het bestand zijn gebruikt. De antwoordmogelijkheden zelf zijn doorgaans gebruikt als ‘value labels’. Het vragenlijstdeel van het SPSS-bestand bevat de gegevens van 2457 leerlingen, die minstens één vraag van de vragenlijst hebben beantwoord. In het totale SPSS-bestand geeft de variabele **DUM_LV** aan of de leerling al dan niet in het vragenlijstdeel van het bestand voorkomt. De in de rest van dit hoofdstuk vermelde percentages hebben betrekking op deze 2457 cases.

Het eerste deel van de vragenlijst begon met de volgende korte inleiding: *“Dit schooljaar stond in het teken van het eindexamen. Je hebt een aantal schoolonderzoeken gemaakt en binnenkort is het Centraal Schriftelijk Examen. De volgende vragen gaan over het afgelopen schooljaar.”*

4.3.1 Tijdbesteding

Vraag 1 luidde: *Hoeveel tijd per week heb je ongeveer aan de volgende activiteiten besteed? Het gaat om een schatting van het aantal uren per week, in de weken dat je naar school ging.*

De antwoordmogelijkheden bij deze vraag waren open. Bij elk van negen activiteiten (zie Tabel 4.1) werd gevraagd een aantal uren in twee hokjes op te schrijven. Dit zijn variabelen **LV01_1** tot en met **LV01_9**. Het gebruik van open antwoorden, in tegenstelling tot gesloten antwoorden, waarbij de invuller kan kiezen uit een aantal categorieën (bv. 0-5 uur, 6-10 uur, 11-15 uur, etc.) heeft als voordeel dat een nauwkeurig beeld van het gevraagde wordt verkregen, maar tegelijkertijd ook het nadeel dat het de respondent de ruimte geeft om ‘rare’ antwoorden te geven, bijvoorbeeld een onmogelijk groot aantal uren.

Bij het laatste item van deze vraag, de activiteit ‘anders, namelijk ...’, kon op de puntjes een specificatie worden opgeschreven. De gegeven antwoorden zijn ondergebracht in een aparte variabele **LV01_9TN**. Deze variabele is numeriek; de value labels zijn de oorspronkelijk opgeschreven antwoorden. Wij hebben ervan afgezien de antwoorden in te dikken tot een beperkt aantal inhoudelijk zinvolle categorieën. Van de 143 leerlingen die hier een specificatie hebben opgeschreven, hadden 6 echter geen aantal uren opgeschreven. Zij hebben op variabele **LV01_9** voor het aantal uren de waarde 997 toegekend gekregen, die gelabeld is als ‘# uren niet gepecificeerd’ en als ontbrekend is gedeclareerd. Tevens zijn er 92 leerlingen die op dit laatste item wel een aantal uren, maar geen activiteit hebben opgeschreven. Zij hebben

op de variabele **LV01_9TN** codering 97 toegekend gekregen, die gelabeld is als ‘activiteit niet gespecificeerd’ en als ontbrekende waarde is gedeclareerd.

Nadat bij het laatste item het wel/niet hebben opgeschreven van een activiteit en het wel/niet hebben opgeschreven van het aantal uren op elkaar waren afgestemd, is over de negen items geteld bij hoeveel items een aantal uren – eventueel ook 0 (nul) – is ingevuld. Het blijkt dat 3 leerlingen op geen enkel item een aantal uren hadden opgeschreven. Deze leerlingen hebben vraag 1 dus in het geheel niet beantwoord. Zij hebben op elk van de negen variabelen **LV01_1** tot en met **LV01_9** de waarde 999 toegekend gekregen, die gelabeld is als ‘vraag 1 niet beantwoord’ en als ontbrekend is gedeclareerd. Aan de overige ontbrekende waarden is 998 toegekend, die gelabeld is als ‘item niet beantwoord’ en ook als ontbrekend is gedeclareerd. De antwoorden op de negen items van vraag 1 staan samengevat in Tabel 4.1. Zoals hierboven reeds is opgemerkt, laat het gebruik van open antwoorden ruimte voor het geven van ‘onwaarschijnlijke’ antwoorden. Uit de in kolom V vermelde ranges van de urenaantallen blijkt dat deze inderdaad voorkomen. Welke waarden zo laag of zo hoog zijn dat ze als ‘onmogelijk’ (of als ‘niet serieus ingevuld’) moeten worden gekwalificeerd, is tot op zekere hoogte een arbitraire keuze die aan de toekomstige gebruiker wordt overgelaten.

Tabel 4.1
Tijdbesteding

I ¹	II ¹	III ¹	IV ¹	V ¹	VI ¹
1	naar school gaan (reistijd + tijd op school)	18	2436 (0)	1-90	30.5
2	televisie kijken	25	2429 (33)	0-99	9.1
3	huiswerk maken, leren, schoolonderzoeken voorbereiden	19	2435 (28)	0-90	7.4
4	baantje	98	2356 (481)	0-97	7.4
5	omgaan met vriend(inn)en	40	2414 (11)	0-93	13.3
6	sporten	78	2376 (385)	0-75	4.8
7	computeren (recreatief, dus niet voor school)	27	2427 (12)	0-99	11.2
8	uitgaan	139	2315 (335)	0-62	5.4
9	anders, namelijk ...	2219	235 (6 ²)	1-99	14.7

Noot 1: I = itemnummer, II = activiteit, III = aantal itemspecifieke ontbrekende waarden (hierin zijn dus niet opgenomen de 3 leerlingen die de gehele vraag niet hebben beantwoord), IV = het aantal valide antwoorden, met tussen haakjes het aantal keren dat 0 uren is geantwoord, V = de range van de valide antwoorden, VI = het gemiddelde.

Noot 2: Deze 6 leerlingen hebben wel een activiteit maar geen aantal uren gespecificeerd, dus het gemiddelde is berekend over de overige 229.

4.3.2 Schoolwerk in vakanties

Vraag 2 luidde: *Hoeveel tijd heb je in de schoolvakanties aan schoolwerk besteed? Het gaat om een schatting van het totaal aantal dagen (8 uur per dag) in de herfstvakantie, kerstvakantie en voorjaarsvakantie.*

De vraag is niet beantwoord door 48 leerlingen (code 99). De variabele is **LV02**. Het gemiddelde van de antwoorden van de andere 2409 leerlingen is 4.8 met een range van 0 tot 84. Ook op deze vraag komen dus onwaarschijnlijk hoge waarden voor. Wij achten het waarschijnlijk dat een aantal leerlingen over het hoofd heeft gezien dat hier naar het aantal dagen werd gevraagd en een schatting van het aantal uren heeft opgeschreven. Er valt niet met zekerheid uit te maken of dat zo is en welke leerlingen dat zijn.

4.3.3 Schoolverzuim

Vraag 3a luidde: *Hoeveel schooldagen ben je niet naar school gegaan en waarom niet? Het gaat om een schatting van het totaal aantal dagen (8 uur per dag) vanaf het begin van het schooljaar tot nu.*

Het antwoordregime bij deze vraag was vergelijkbaar met dat bij vraag 1. Bij vier mogelijke redenen van schoolverzuim moest een getal (het aantal dagen) worden ingevuld. Dit zijn de variabelen **LV03_A1** tot en met **LV03_A4**. Het vierde item was een zelf in te vullen reden (zie Tabel 4.2). De bewerkingen zijn dan ook analoog aan die bij vraag 1. Ook hier zijn bij het laatste item inconsistenties tussen het wel/niet hebben opgeschreven van een reden en het wel/niet hebben ingevuld van het aantal dagen. Van de 384 leerlingen die bij ‘andere reden, namelijk ...’ een reden hebben opgeschreven, hebben 7 niet het aantal dagen gespecificeerd en hebben 2 als aantal ‘0 dagen’ ingevuld. Bij die 7 leerlingen is voor het aantal dagen de waarde 997 toegekend, die is gelabeld als ‘# dagen niet gespecificeerd’ en als ontbrekende waarde is gedeclareerd. Tevens zijn er bij dit item 142 leerlingen die wel een aantal dagen, maar geen verzuimreden hebben ingevuld. Bij hen is op **LV03_A4N** – de variabele die de ingevulde reden bevat – de waarde 997 toegekend, die is gelabeld als ‘reden niet gespecificeerd’ en als ontbrekend is gedeclareerd.

Nadat bij het laatste item het wel/niet ingevuld zijn van een reden en het wel/niet ingevuld zijn van een aantal dagen op elkaar waren afgestemd, is over de vier items geteld bij hoeveel items een aantal dagen – eventueel ook 0 (nul) – is ingevuld. Het blijkt dat 26 leerlingen bij geen enkel item iets hadden ingevuld. Deze hebben vraag 3a dus in het geheel niet beantwoord en bij hen is aan de vier variabelen de waarde 999 toegekend, die is gelabeld als ‘vraag 3a niet beantwoord’ en die als ontbrekend is gedeclareerd. Voor de overige ontbrekende waarden is 998 toegekend, die is gelabeld als ‘item niet beantwoord’ en uiteraard ook als ontbrekend is gedeclareerd. De antwoorden op de vier items van vraag 3a worden samengevat in Tabel 4.2. Ook hier geldt dat ‘onmogelijke’ (hoge) antwoorden voorkomen, die vermoedelijk voor een groot deel veroorzaakt zijn doordat de leerlingen het aantal uren hebben opgeschreven in plaats van het aantal dagen. Wij laten het aan de toekomstige gebruiker over welke waarden hij/zij eventueel buiten beschouwing wil laten.

Tabel 4.2

Verzuimredenen

I ¹	II ¹	III ¹	IV ¹	V ¹	VI ¹
1	ziekte	46	2385 (283)	0-99	6.66
2	spijbelen	372	2059 (1133)	0-70	1.73
3	lesuitval	205	2226 (230)	0-99	4.42
4	andere reden, namelijk	1905	526 (7 ¹)	0-80	4.28

Noot 1: I = itemnummer, II = verzuimreden, III = aantal itemspecifieke ontbrekende waarden (hierin zijn dus niet de 26 leerlingen opgenomen die de gehele vraag niet hebben beantwoord), IV = het aantal valide antwoorden, met tussen haakjes het aantal keren dat 0 dagen is geantwoord, V = de range van de valide antwoorden, VI = gemiddelde.

Noot 2: Deze 7 leerlingen hebben wel een reden maar geen aantal dagen ingevuld, dus het gemiddelde is berekend over de overige 519.

De reden die is ingevuld bij ‘andere reden, namelijk ...’ is opgeslagen in de variabele **LV03_A4N**. Zoals uit het bovenstaande volgt, hebben 26 leerlingen op deze variabele code 999 die staat voor ‘vraag 3a niet beantwoord’; 1905 leerlingen hebben code 998, die staat voor ‘item niet beantwoord’ en die aangeeft dat ze noch een aantal dagen, noch een reden hebben ingevuld; 142 leerlingen hebben code 997, die betekent ‘reden niet gespecificeerd’. Codes 997, 998 en 999 zijn als ontbrekend gedeclareerd. De overige 384 leerlingen hebben op **LV03_A4N** numerieke waarden (lopend van 2 tot en met 211), die zijn gelabeld met

de specificatie van de ‘andere reden’. Verreweg de meeste specificaties zijn unieke formuleringen, maar enkele komen vaker voor. Wij hebben er van afgezien de antwoorden in te dikken tot een beperkt aantal inhoudelijk zinvolle categorieën.

Vraag 3b luidde: *Als je spijbelde, wat deed je dan meestal? Kruis maximaal drie vakjes aan.*

Er waren zeven antwoordmogelijkheden, waarvan er dus maximaal drie mochten worden aangekruist. Er zijn 1379 leerlingen die geen enkel item hebben aangekruist, en daarmee de vraag niet hebben beantwoord. Dit zijn natuurlijk vooral leerlingen die nooit spijbelde. De overige 1078 leerlingen hebben één of meer items aangekruist. Een beperkt aantal leerlingen heeft meer dan drie redenen aangekruist (één leerling zelfs alle zeven). Omdat het mogelijk was meer dan één antwoord aan te kruisen, zijn de antwoorden opgeslagen als zeven (quasi) dichotome variabelen (items) die aangeven of het item (1) niet of (2) wel is aangekruist. Dit zijn variabelen **LV03_B1** tot en met **LV03_B7**. Voor de 1379 leerlingen die geen enkele reden hebben aangekruist, is telkens code 9 toegekend, die is gelabeld als ‘vraag 3b niet beantwoord’ en die als ontbrekend is gedeclareerd. Het aantal aankruisingen per item staat in Tabel 4.3.

Tabel 4.3
Spijbelredenen

Nr	Item	Aantal aankruisingen
1	uitslapen	619
2	rondhangen met vrienden / vriendinnen	409
3	tv kijken / internetten / gamen	250
4	huiswerk maken	272
5	werken aan een profielwerkstuk	233
6	leren voor het examen	184
7	anders	294

In dit geval hebben we niet gevraagd naar een specificatie van ‘anders’.

4.3.4 Bijles

Vraag 4a: *Heb je dit jaar bijles gehad om je beter op het examen voor te bereiden?*

Vraag 4b: *Hoeveel uur bijles heb je (ongeveer) gehad?*

De antwoordmogelijkheden bij vraag 4a (variabele **LV04_A**) waren: ‘nee’ en ‘ja’. Werd het antwoord ‘ja’ aangekruist, dan werd de invuller doorverwezen naar vraag 4b (variabele **LV04_B**) waar het aantal uren bijles per week moest worden ingevuld. Gevraagd werd een schatting te geven van het gemiddelde aantal uren bijles per week.

De hele vraag – dus zowel 4a als 4b – is niet beantwoord door 9 leerlingen. Van de overige 2448 geven 1711 leerlingen aan dat ze geen bijles hebben gehad; bij hen is op vraag 4b voor het aantal uren code 98 toegekend die is gelabeld als ‘n.v.t. - geen bijles’ en die als ontbrekend is gedeclareerd. Van de 737 leerlingen die wel bijles hebben gehad, hebben 6 niet het aantal uren opgegeven. Zij hebben de waarde 97 toegekend gekregen die is gelabeld als ‘# uren niet gespecificeerd’ en die ook als ontbrekend is gedeclareerd. Enigszins merkwaardig zijn de 5 leerlingen die zeggen wel bijles te hebben gehad, maar bij het aantal uren 0 (nul) invulden. Het is mogelijk dat zij in het hele jaar maar een paar uur bijles hebben gehad en dat zij, omdat naar het aantal uren per week werd gevraagd, dit hebben afgerond tot 0.

4.3.5 Examenvoorbereiding

Onder de noemer 'examenvoorbereiding' vallen de vraagnummers 5 tot en met 9.

Vraag 5: *Vind je dat je je goed hebt voorbereid op de schoolonderzoeken?*

De antwoordmogelijkheden op deze vraag (variabele **LV05**) waren: 1) heel slecht, 2) slecht, 3) beetje slecht, 4) gaat wel, 5) beetje goed, 6) goed, 7) heel goed. De vraag is niet beantwoord door 11 leerlingen; bij hen is code 9 toegekend, die is gelabeld als 'geen antwoord' en die als ontbrekend is gedeclareerd. Daarnaast hebben 13 leerlingen een onbruikbaar antwoord ingevuld; bij hen is code 8 toegekend, die is gelabeld als 'overig' en die eveneens als ontbrekend is gedeclareerd. Het gemiddelde van de 2433 leerlingen die een valide antwoord hebben gegeven, is 4.95 ($SD = 1.19$).

Vraag 6: *Denk je dat je je slechter of beter hebt voorbereid op de schoolonderzoeken dan de meeste van je klasgenoten?*

De antwoordmogelijkheden op deze vraag (variabele **LV06**) waren: 1) veel slechter, 2) slechter, 3) iets slechter, 4) even goed, 5) iets beter, 6) beter, 7) veel beter. De vraag is niet beantwoord door 20 leerlingen; bij hen is code 9 toegekend, die is gelabeld als 'geen antwoord' en die als ontbrekend is gedeclareerd. Daarnaast hebben 5 leerlingen een onbruikbaar antwoord ingevuld; bij hen is code 8 toegekend, die is gelabeld als 'overig' en die eveneens als ontbrekend is gedeclareerd. Het gemiddelde van de 2432 leerlingen die een valide antwoord hebben gegeven, is 4.05 ($SD = 1.23$).

Vraag 7: *Wat is je gemiddelde cijfer op de schoolonderzoeken?*

Deze vraag (variabele **LV07**) had een open antwoordmogelijkheid: gevraagd werd het cijfer met één cijfer achter de komma in te vullen ('graag één cijfer achter de komma, dus bijvoorbeeld 6,7'). Het getal is ingelezen als een geheel, tweecijferig getal, dus het opgeschreven gemiddelde SO-cijfer vermenigvuldigd met 10. Bij enkele leerlingen die een ééncijferig getal hadden opgeschreven (bijvoorbeeld '7'), is dit omgezet naar een tweecijferig getal ('70'). De vraag is niet beantwoord door 107 leerlingen; bij hen is code 999 toegekend, die is gelabeld als 'geen antwoord' en die als ontbrekend is gedeclareerd. Het gemiddelde van de 2350 leerlingen die een valide antwoord hebben gegeven, is 65.3 ($SD = 5.7$).

Vraag 8: *Denk je dat je goed voorbereid zult zijn op het centraal schriftelijk examen?*

De antwoordmogelijkheden op deze vraag (variabele **LV08**) waren: 1) heel slecht, 2) slecht, 3) beetje slecht, 4) gaat wel, 5) beetje goed, 6) goed, 7) heel goed. De vraag is niet beantwoord door 5 leerlingen; bij hen is code 9 toegekend, die is gelabeld als 'geen antwoord' en die als ontbrekend is gedeclareerd. Daarnaast hebben 2 leerlingen een onbruikbaar antwoord ingevuld; bij hen is code 8 toegekend, die is gelabeld als 'overig' en die eveneens als ontbrekend is gedeclareerd. Het gemiddelde van de 2450 leerlingen die een valide antwoord hebben gegeven, is 5.45 ($SD = 0.98$).

Vraag 9: *Denk je dat je slechter of beter voorbereid zult zijn op het centraal schriftelijk examen dan de meeste van je klasgenoten?*

De antwoordmogelijkheden op deze vraag (variabele **LV09**) waren: 1) veel slechter, 2) slechter, 3) iets slechter, 4) even goed, 5) iets beter, 6) beter, 7) veel beter. De vraag is niet beantwoord door 10 leerlingen; bij hen is code 9 toegekend, die is gelabeld als 'geen antwoord' en die als ontbrekend is gedeclareerd. Daarnaast hebben 2 leerlingen een onbruikbaar antwoord gegeven; bij hen is code 8 toegekend, die is

gelabeld als ‘overig’ en die eveneens als ontbrekend is gedeclareerd. Het gemiddelde van de 2445 leerlingen die een valide antwoord hebben gegeven is 4.30 ($SD = 0.96$).

4.3.6 Slaagverwachting

Vraag 10: *Hoe groot is volgens jou de kans dat je in één keer zult slagen voor het eindexamen?*

De antwoordmogelijkheid op deze vraag (variabele **LV10**) was open: gevraagd werd de kans als een percentage in te vullen. De vraag is niet beantwoord door 11 leerlingen; bij hen is code 999 toegekend, die is gelabeld als ‘geen antwoord’ en die als ontbrekend is gedeclareerd. Het gemiddelde van de 2446 leerlingen met een valide antwoord is 74.8 ($SD = 18.9$) procent kans, met een minimum van 0 (nul) en een maximum van 100 procent kans.

4.3.7 Examenvaardigheden

Vraag 11: *Hieronder staat een aantal vaardigheden die van belang kunnen zijn tijdens de onderdelen van het Centraal Schriftelijk Examen.*

Hoe goed denk jij dat jij in staat bent om tijdens het CSE ...

De negen vaardigheden waarnaar werd gevraagd, zijn variabelen **LV11_A** tot en met **LV11_I** en staan vermeld in de tweede kolom van Tabel 4.4. De antwoordmogelijkheden waren: 1) heel slecht 2) vrij slecht 3) gaat wel 4) vrij goed 5) heel goed. Er zijn 4 leerlingen die geen van de items hebben beantwoord. Zij hebben op elk item code 9 (‘vraag 11 niet beantwoord’) toegekend gekregen. Daarnaast komen op de meeste items nog enkele ontbrekende antwoorden voor (code 8 ‘item niet beantwoord’). Aan onbruikbare antwoorden is code 7 (‘overig’) toegekend. Codes 7, 8 en 9 zijn als ontbrekend gedeclareerd. Uit een telling van het aantal valide antwoorden binnen leerlingen over de negen items blijkt dat 2431 leerlingen (98.9%) alle items valide hebben beantwoord, dat 18 leerlingen (0.7%) één niet valide item hebben en dat 4 leerlingen (0.2%) geen enkel item valide hebben beantwoord. De overige 4 leerlingen (0.2%) hebben meerdere items onbeantwoord gelaten.

Tabel 4.4

Examenvaardigheden tijdens het CSE

Nr	Item	<i>M</i>
a	je te herinneren wat je geleerd hebt	3.72
b	kalm te blijven, niet in paniek te raken	3.70
c	de tijd goed te verdelen	3.29
d	goed geconcentreerd te blijven	3.25
e	van tevoren te bedenken hoe je te werk zult gaan	3.08
f	systematisch te werk gaan	3.37
g	je creativiteit te gebruiken	3.65
h	een verstandige gokstrategie toe te passen bij meerkeuzevragen	3.37
i	je antwoorden te controleren	3.58

De betrouwbaarheid (Cronbach’s α) van de negen items samen bedraagt .60, hetgeen erop wijst dat de items nauwelijks een homogene (intern consistente) schaal vormen. Er zijn echter ook geen “slechte” items, dat wil zeggen items die door weglating uit de schaal de betrouwbaarheid ervan zouden verhogen. Aan het bestand is een nieuwe variabele **LV_EXVA**, gelabeld als ‘Schaal examenvaardigheden’, toegevoegd. De

score is het gemiddelde van de negen items, waarbij één itemantwoord mocht ontbreken. Op deze schaal variabele hebben 2449 (99.7%) leerlingen een score. Het gemiddelde bedraagt 3.45 ($SD = 0.44$). De acht leerlingen die geen score hebben, zijn ‘system missing’ op deze schaal.

4.3.8 Welbevinden met klasgenoten

De items waarmee dit concept is gemeten zijn afkomstig van Peetsma, Wagenaar, & Kat (2001). De betreffende vraag (vraag 13) luidde: *Wat vond je het afgelopen schooljaar van je klasgenoten?*

De vraag bestond uit zes items in de vorm van beweringen, met als antwoordmogelijkheden: 1) klopt helemaal niet, 2) klopt niet, 3) klopt soms wel, soms niet, 4) klopt, 5) klopt precies. Dit zijn items **LV13_A** tot en met **LV13_F** – zie Tabel 4.5. Uit een telling van het aantal antwoorden binnen leerlingen over de items, blijkt dat 2423 leerlingen (98.6%) alle zes items hebben beantwoord, 25 leerlingen (1.0%) één item hebben overgeslagen, 1 leerling (< 0.1%) twee items heeft overgeslagen en 8 leerlingen (0.3%) geen enkel item hebben beantwoord. Deze 8 leerlingen hebben op alle items code 9 (‘vraag 13 niet beantwoord’) toegekend gekregen. Daarnaast komen item-specifieke ontbrekende waarden voor (code 8: ‘item niet beantwoord’). Voor enkele onbruikbare dubbele antwoorden is code 7, gelabeld ‘overig’ toegekend. Codes 7, 8 en 9 zijn als ontbrekend gedeclareerd. De betrouwbaarheid (Cronbach’s α - na spiegeling van de items b en e) van de schaal bedraagt .83. Er zijn geen items die door weglating de betrouwbaarheid van de schaal zouden verhogen. Aan het bestand is een nieuwe variabele **LV13_KLAS** (‘welbevinden met klasgenoten’) toegevoegd waarvan de score is berekend door het gemiddelde te nemen op de zes items (na spiegeling van de antwoorden op items b en e), waarbij één itemscore mocht ontbreken. Op deze nieuwe variabele hebben 2448 (99.6%) leerlingen een score. Het gemiddelde daarvan bedraagt 3.89 ($SD = 0.61$). De negen leerlingen die geen score hebben, zijn ‘system missing’ op deze schaal.

Tabel 4.5

Welbevinden met klasgenoten

Nr	Item	<i>M</i>
a	Ik had veel contact met mijn klasgenoten	3.69
<u>b</u>	Ik had liever in een andere klas gezeten	2.16
c	Wij hadden een leuke klas	3.78
d	Ik kon goed met mijn klasgenoten overweg	3.89
<u>e</u>	Ik voelde mij soms alleen in de klas	1.76
f	Ik vond het leuk om met mijn klasgenoten om te gaan	3.87

4.3.9 Persoonlijkheid

Het persoonlijkheidsprofiel van de leerlingen is vastgesteld met de ‘Five-Factor Personality Inventory’ (FFPI; Hendriks, Hofstee, & De Raad, 1999, 2011), waarmee scores worden verkregen op de vijf factoren Extraversie, Mildheid, Ordelijkheid, Emotionele Stabiliteit en Intellectuele Autonomie. Extraversie beschrijft iemands mate van sociale deelname, energieniveau en expressiviteit. Mildheid beschrijft hoe iemand met anderen omgaat. Ordelijkheid beschrijft hoe iemand met taken, spullen en zijn of haar omgeving omgaat. Emotionele Stabiliteit beschrijft hoe bestendig iemand is tegen emotionele en fysieke prikkels, zoals indrukken, drukte en geluiden. Intellectuele Autonomie beschrijft iemands mate van onafhankelijkheid in denken en beslissen.

De FFPI is voor volwassenen gemaakt, maar bleek met enkele kleine aanpassingen tevens geschikt voor leerlingen vanaf ongeveer 11 jaar (Hendriks et al., 2011; Hendriks, Kuyper, Offringa & Van der Werf, 2008). Bij gebruik van de FFPI-factorscores dient naar deze laatste twee publicaties te worden verwezen. De FFPI bestaat uit 100 korte gedragsbeschrijvende zinnen. Per factor zijn er 20 items, waarvan er 10 de positieve ‘pool’ betreffen, en 10 de negatieve. Enkele voorbeelditems zijn:

- *Ik neem tijd voor een praatje* (Extraversie)
- *Ik probeer ruzie te voorkomen* (Mildheid)
- *Ik werk graag volgens schema* (Ordelijkheid)
- *Ik zet tegenslag snel opzij* (Emotionele Stabiliteit)
- *Ik weet wat ik wil* (Intellectuele Autonomie)

In de vragenlijst waren de items verspreid over twee plaatsen, met elk 50 items. De reden hiervoor was het aanbrengen van enige afwisseling in de vragenlijst wat betreft het soort vragen dat moest worden beantwoord. De eerste keer werden de FFPI-items als volgt ingeleid:

Vraag 14: *De volgende uitspraken gaan over hoe je zelf vindt dat je bent.*

Het vraagteken betekent ‘geen nee en geen ja’. Denk niet te lang na: je eerste gedachte is vaak het best. Wanneer je een uitspraak (echt) niet snapt, mag je het hokje ‘snap ik niet’ aankruisen.

De antwoordmogelijkheden stonden boven de kolommen met hokjes en luiden: NEE! nee ? ja JA! en ‘snap ik niet’. De tweede keer (vraag 20) was de inleiding hetzelfde, met als enige verschil dat in de eerste zin het woordje ‘weer’ was toegevoegd: *‘De volgende uitspraken gaan weer ...’*.

De antwoorden zijn ingelezen met codes 1 tot en met 5 voor de vijf inhoudelijke antwoorden en code 6 voor ‘snap ik niet’. Bij twee aanliggende antwoorden werd in dit geval de meest extreme van de twee codes genomen, dus bij 12 werd dit 1, bij 23 werd dit 2, bij 34 werd dit 4 en bij 45 werd dit 5. Indien zowel een itemantwoord was gegeven maar ook ‘snap ik niet’ was aangekruist, werd (toch) gekozen voor dat itemantwoord. De andere mogelijkheden zijn als ontbrekend behandeld. De antwoorden zijn verwerkt met het standaardscoringsprogramma van de FFPI, dat een maximum van vijf ontbrekende antwoorden per factorpool toestaat (maximaal 50% ontbrekende oordelen). Bij maximaal vijf ontbrekende waarden op een factorpool vervangt het scoringsprogramma de ontbrekende antwoorden door het gemiddelde van de persoon op de wel door hem of haar ingevulde items van de betreffende factorpool. Bij meer dan vijf ontbrekende antwoorden op een factorpool wordt geen score voor die factor berekend. Factorscores worden berekend via een gewogen somming van de itemantwoorden op alle 100 items, waarbij (per factor verschillende) itemgewichten worden gebruikt die zijn bepaald in de algemene normgroep volwassenen voor de FFPI (Hendriks et al., 1999, 2011), waarmee de structuur van de FFPI bij (jong)adolescenten vrijwel identiek is (Hendriks et al., 2008, 2011). Behalve vanwege teveel ontbrekende itemantwoorden kunnen leerlingen om nog twee andere redenen geen scores op de FFPI hebben. De eerste is als een leerling te vaak hetzelfde antwoord heeft gegeven (meer dan 20 op een bladzijde, of meer dan 70 in totaal). De tweede is als het scoringsprogramma een slompatroon (12345432123.. etc.) in de antwoorden ontdekte, wat naar alle waarschijnlijkheid duidt op het niet serieus hebben ingevuld van de FFPI. Van de 2457 leerlingen hebben er 2381 (96.9%) scores op alle vijf factoren toegekend gekregen en hebben 51 leerlingen (2.1%) op geen enkele factor een score toegekend gekregen. De overige 25 leerlingen (1.0%) hebben op minimaal één en maximaal vier factoren een score.

Bij de 2381 leerlingen met valide factorscores op alle vijf factoren is de betrouwbaarheid van de scores bepaald. De interne consistentie (gestratificeerde α) was .85 voor Extraversie, .81 voor Mildheid, .84 voor Ordelijkheid, .85 voor Emotionele Stabiliteit, en .73 voor Autonomie. Om auteursrechtelijke redenen zijn

alleen de FFPI-factorscores (en geen itemscores) in het bestand opgenomen. Tabel 4.6 toont de basisinformatie over de vijf factoren.

Tabel 4.6

Basisinformatie FFPI: variabele naam, factor naam, aantal, gemiddelde en standaard deviatie

Variabele	Naam	<i>N</i>	<i>M</i>	<i>SD</i>
LV_FFPI1	Extraversie	2401	0.93	0.89
LV_FFPI2	Mildheid	2382	1.91	0.90
LV_FFPI3	Ordelijkheid	2402	0.00	1.07
LV_FFPI4	Emotionele Stabiliteit	2405	0.92	1.02
LV_FFPI5	Intellectuele Autonomie	2402	0.87	0.84

4.3.10 Vrienden/vriendinnen

Er werd gevraagd (vraag 15a): *Hoeveel goede vrienden en vriendinnen heb je?*

De antwoordmogelijkheden bij deze vraag waren open; gevraagd werd zowel het aantal goede vrienden (**LV15_A1**) als het aantal goede vriendinnen (**LV15_A2**) in te vullen. Aangezien er bij beide vragen twee antwoordvakjes waren, is het maximale aantal 99 (dat komt in beide gevallen voor). Er zijn 41 leerlingen die noch een aantal vrienden, noch een aantal vriendinnen hadden opgeschreven. Zij hebben vraag 15a niet beantwoord (code 999). De overige 2416 leerlingen (98.3%) hadden bij minstens één van de twee vragen een aantal opgeschreven, waarbij ook 0 (nul) enkele malen als aantal voorkomt (9 leerlingen hadden tweemaal 0 opgeschreven). In de gevallen dat er slechts bij één van de twee vragen een aantal was opgeschreven is het niet opgeschreven aantal op de andere vraag door ons op 0 (nul) gesteld. Het gemiddelde aantal goede vrienden bedraagt 7.8 (*SD* = 10.2), het gemiddelde aantal goede vriendinnen 6.8 (*SD* = 8.4). Enkele zeer hoge aantallen komen ons onwaarschijnlijk voor. Ter illustratie: de correlatie tussen het aantal goede vrienden en vriendinnen bedraagt .62. Indien op beide variabelen geselecteerd wordt op een bovengrens van 20, wordt de correlatie .17. Aangezien op die leeftijd jongens vooral jongens als vriend hebben en meisjes vooral meisjes, komt deze veel lagere correlatie ons geloofwaardiger voor.

Vraag 15b: *Waarvan ken je je meeste vriend(inn)en? Kruis maximaal drie vakjes aan.*

Deze vraag had zeven antwoordmogelijkheden (items) die al dan niet konden worden aangekruist. Deze zijn, net als bij vraag 3b, omgezet naar (quasi) dichotome variabelen (1 = niet aangekruist, 2 = aangekruist). De variabelenamen zijn **LV15_B1** tot en met **LV15_B7**. Vraag 15b is in het geheel niet beantwoord door 16 leerlingen; bij hen hebben alle variabelen de waarde 9 gekregen, die is gelabeld als ‘vraag 15b niet beantwoord’ en als ontbrekende waarde is gedeclareerd. Omdat het aankruisen van het eerste item – ‘ik heb geen vrienden’ – niet samengaat met het tevens aankruisen van een van de volgende items, is bij de leerlingen die dat toch hadden gedaan, de 2 op het eerste item vervangen door een 1. Bij de 5 leerlingen die zeiden geen vrienden te hebben en op het eerste item dus de waarde 2 hebben, is bij de zes overige items de waarde 8 ingevuld, die is gelabeld als ‘niet van toepassing - geen vrienden’ en die is gedeclareerd als ontbrekende waarde. Het aantal aankruisingen per item staat in Tabel 4.7.

Tabel 4.7*Waarvan ken je je vrienden?*

Nr	Item	Aantal aankruisingen
1	ik heb geen vrienden	5
2	van school	1566
3	van de buurt waar ik woon	864
4	van een (sport)vereniging of club	942
5	zowel van school als daarbuiten	1296
6	van internet	122
7	ergens anders van	501

4.3.11 Waargenomen autonomie

Er werd gevraagd (vraag 22): *Wie neemt bij jou thuis de beslissingen over de volgende onderwerpen?*

De onderwerpen (items) zijn te vinden in Tabel 4.8. Dit zijn variabelen **LV22_A** tot en met **LV22_Q**. De items zijn gebaseerd op de literatuur over ‘opvoedstijlen’ en betreffen het concept ‘authoritatieve’ opvoedstijl. Wikipedia geeft hiervan de volgende omschrijving: ‘Authoritative parenting is characterized by a child-centered approach that holds high expectations of maturity, compliance to parental rules and directions, while allowing for an open dialogue about those rules and behaviors between parent and child’. Geschikte verwijzingen naar dit concept zijn Baumrind (1978) en Steinberg, Elmen & Mounts (1989). Ongeveer de helft van de betreffende items is ingebracht door Veenstra (Veenstra & Kuyper, 2004) in VOCL’93. De overige items zijn door het GION toegevoegd om een bredere dekking te verkrijgen.

Tabel 4.8*Waargenomen autonomie: wie neemt thuis de beslissingen?*

I ¹	II ¹	III ¹	IV ¹	V ¹
a	hoe laat je thuis moet zijn	2289	82	3.30
b	of je een baantje mag hebben	2268	103	3.82
c	of je thuis alcohol mag drinken	2182	180	3.47
d	of je op feestjes alcohol mag drinken	2212	152	4.26
e	naar welke tv-programma’s je kijkt	2229	129	4.60
f	welke vervolgopleiding je gaat volgen	2322	40	3.94
g	wanneer je je huiswerk maakt	2278	95	4.65
h	hoe lang je internet of computerspellen speelt	2263	100	4.54
i	welke internetsites je bezoekt	2250	120	4.80
j	of je uit mag gaan	2278	86	3.83
k	hoe je op vakantie gaat (of je zelf op vakantie mag gaan)	2298	72	3.49
l	van welke verenigingen je lid bent	2170	200	3.85
m	of je softdrugs mag gebruiken	1343	1010	3.19
n	of je uit huis gaat wonen	1952	415	3.48
o	of je de ‘pil’ mag gebruiken	1289	1055	3.58
p	of je naar de kerk gaat	1311	1045	3.95
q	waar je je zak- en kleedgeld aan besteedt	2178	193	4.52

Noot 1: I = itemnummer, II = item, III = aantal leerlingen met valide antwoord, IV = aantal keer ‘niet van toepassing’ aangekruist, V = gemiddelde.

De antwoordmogelijkheden bij deze vraag waren: 1) mijn ouders verzorgers, **zonder** overleg met mij; 2) mijn ouders/verzorgers, **na** overleg met mij; 3) ik en mijn ouders **samen**; 4) ik zelf, **na** overleg met mijn ouders/verzorgers; 5) ik zelf, **zonder** overleg met mijn ouders/verzorgers; en 6) niet van toepassing. Aanliggende dubbelcoderingen werden behouden; bijvoorbeeld, indien zowel 'ik en mijn ouders **samen**' als 'ik zelf, **na** overleg met mijn ouders/verzorgers' als antwoord op een vraag waren aangekruist, dan werd het antwoord als 3.5 gecodeerd. Overige dubbelcoderingen werden gehercodeerd naar code 7 en gelabeld als 'overig'. Er zijn 56 leerlingen die geen enkel item hebben beantwoord. Zij hebben op alle items code 9, gelabeld als 'vraag 22 niet beantwoord' toegekend gekregen. Overgeslagen items kregen de code 8, gelabeld als 'item niet beantwoord'. De coderingen 6, 7, 8 en 9 zijn als ontbrekend gedeclareerd.

Op deze items hebben we volgens twee verschillende benaderingen een schaal bepaald. De eerste benadering is als het ware de normale, namelijk het berekenen van het gemiddelde over de items. Een hoge score op deze schaal (**LV_AUTON**) geeft aan dat de ouders van de leerling in diens perceptie een 'laissez faire' opvoedingsstijl hanteren. Hun kind ervaart relatief veel autonomie. De tweede benadering bestaat uit het bepalen van de proportie antwoorden 'ik, na overleg met mijn ouders' en meet de 'authoritatieve' opvoedingsstijl in strikte zin. Deze schaal hebben we **LV_AUTH** genoemd. Wat betreft het omgaan met ontbrekende waarden zijn we als volgt te werk gegaan. Er mocht op maximaal zes items een ontbrekende waarde voorkomen, waarvan maximaal vier maal 'niet van toepassing' (code 6) en maximaal twee maal anderszins 'ontbrekend' (codes 7, 8, 9). Er zijn 2126 leerlingen (86.5%) die aan dit criterium voldoen. De betrouwbaarheid van de twee schalen is bepaald op de volledige cases, dat wil zeggen de leerlingen die alle items valide hadden beantwoord. Dat zijn er slechts 549. Een belangrijke reden van dit lage aantal is dat alle jongens het item over het gebruik van de anticonceptie pil met 'niet van toepassing' zullen hebben beantwoord. Ook op de items over gebruik van soft drugs en kerkbezoek komen zeer veel 'niet van toepassing' antwoorden voor. De betrouwbaarheid van **LV_AUTON** is .92, van **LV_AUTH** .80. Er zijn geen slechte items die de betrouwbaarheid verlagen, in die zin dat het weglaten ervan de betrouwbaarheid zou verhogen. Het gemiddelde op **LV_AUTON** is 4.00 ($SD = 0.58$). Het gemiddelde op **LV_AUTH** is 0.23 ($SD = 0.17$).

5 BESCHRIJVING DATABESTAND

Het databestand bevat de examen- en vragenlijstgegevens van de 3024 leerlingen die aan ten minste één van de onderdelen van de gegevensverzameling in 2010 in HAVO-5 hebben deelgenomen. In de voorgaande twee hoofdstukken is al een groot deel van de variabelen van het databestand geïntroduceerd. Enkele administratieve variabelen zijn nog niet besproken, maar de precieze betekenis van deze variabelen kan eenvoudig worden afgeleid uit de ‘variable labels’ en ‘value labels’ in het databestand. Onderstaand overzicht geeft een korte toelichting bij elke variabele uit het databestand. Op basis van de informatie in de voorgaande twee hoofdstukken, de samenvattende informatie in dit hoofdstuk en beide soorten labels moet het naar onze inschatting mogelijk zijn om probleemloos met het databestand te werken. Bij gebruik van de variabelen uit het databestand dient gerefereerd te worden naar de bron die is genoemd in het hoofdstuk waar de variabele beschreven wordt.

Administratieve variabelen

LL_ID	Leerlingidentificatie
SCH_ID	Schoolidentificatie
KLAS_ID	Klasidentificatie
DUM_KLID	Klasidentificatie bruikbaar? (Nee = klasindeling binnen de school onbekend)
LEERJAAR	Leerjaar
TARGET	Heeft de leerling eerder aan PRIMA en/of COOL ⁵⁻¹⁸ deelgenomen?
DEELNAME	Heeft de school ‘collectief’ of ‘individueel’ deelgenomen?
TOESTEM	Is er toestemming voor de koppeling van de gegevens aan andere databestanden?
DUM_NED	Zijn er gegevens beschikbaar van het 1 ^{ste} tijdvakexamen Nederlands?
DUM_WIA	Zijn er gegevens beschikbaar van het 1 ^{ste} tijdvakexamen Wiskunde A?
DUM_WIB	Zijn er gegevens beschikbaar van het 1 ^{ste} tijdvakexamen Wiskunde B?
DUM_ENG	Zijn er gegevens beschikbaar van het 1 ^{ste} tijdvakexamen Engels?
DUM_LV	Is de leerlingvragenlijst HAVO-5 beantwoord?

Variabelen beschreven in hoofdstuk 2

DAT_NED	Op welke datum is het 1 ^{ste} tijdvakexamen Nederlands gemaakt?
DAT_WIA	Op welke datum is het 1 ^{ste} tijdvakexamen Wiskunde A gemaakt?
DAT_WIB	Op welke datum is het 1 ^{ste} tijdvakexamen Wiskunde B gemaakt?
DAT_ENG	Op welke datum is het 1 ^{ste} tijdvakexamen Engels gemaakt?
DAT_LV	Op welke datum is de vragenlijst gemaakt?
MIS_NED	Percentage ontbrekende antwoorden bij het 1 ^{ste} tijdvakexamen Nederlands
MIS_WIA	Percentage ontbrekende antwoorden bij het 1 ^{ste} tijdvakexamen Wiskunde A
MIS_WIB	Percentage ontbrekende antwoorden bij het 1 ^{ste} tijdvakexamen Wiskunde B
MIS_ENG	Percentage ontbrekende antwoorden bij het 1 ^{ste} tijdvakexamen Engels
NTOETS	Van hoeveel onderdelen zijn gegevens beschikbaar?
SCH_REG	In welke provincie staat de school?
SCH_STED	Urbanisatiegraad van de school
SEKSE	Geslacht leerling
LEEFTIJD	Leeftijd van de leerling op 01-03-2010
OW_TYPE	Welk type onderwijs volgt de leerling?
OW_PROFIEL	Voor welk profiel heeft de leerling gekozen?

Variabelen beschreven in hoofdstuk 3

NED_RSCO1	Ruwe score op het 1 ^{ste} tijdvakexamen Nederlands zonder de samenvattingsopdracht
NED_RSCO2	Ruwe score op het 1 ^{ste} tijdvakexamen Nederlands
NED_CFR	Cijfer 1 ^{ste} tijdvakexamen Nederlands
NED_ZSCO	Relatieve niveau op het 1 ^{ste} tijdvakexamen Nederlands uitgedrukt in een Z-score
WIA_RSCO	Ruwe score op het 1 ^{ste} tijdvakexamen Wiskunde A
WIA_CFR	Cijfer 1 ^{ste} tijdvakexamen Wiskunde A
WIA_ZSCO	Relatieve niveau op het 1 ^{ste} tijdvakexamen Wiskunde A uitgedrukt in een Z-score
WIB_RSCO	Ruwe score op het 1 ^{ste} tijdvakexamen Wiskunde B
WIB_CFR	Cijfer 1 ^{ste} tijdvakexamen Wiskunde B
WIB_ZSCO	Relatieve niveau op het 1 ^{ste} tijdvakexamen Wiskunde B uitgedrukt in een Z-score
ENG_RSCO	Ruwe score op het 1 ^{ste} tijdvakexamen Engels
ENG_CFR	Cijfer 1 ^{ste} tijdvakexamen Engels
ENG_ZSCO	Relatieve niveau op het 1 ^{ste} tijdvakexamen Engels uitgedrukt in een Z-score

Variabelen beschreven in hoofdstuk 4

LV01_1	# uren naar school
LV01_2	# uren televisie
LV01_3	# uren huiswerk
LV01_4	# uren baantje
LV01_5	# uren vriend(inn)en
LV01_6	# uren sporten
LV01_7	# uren computeren
LV01_8	# uren uitgaan
LV01_9	# uren anders
LV01_9TN	anders, nl - welke activiteit
LV02	# dagen schoolwerk in vakanties
LV03_A1	# dagen ziek
LV03_A2	# dagen spijbelen
LV03_A3	# dagen lesuitval
LV03_A4	# dagen andere uitval
LV03_A4N	anders nl. - reden uitval
LV03_B1	uitslapen
LV03_B2	rondhangen met vriend(inn)en
LV03_B3	tv kijken
LV03_B4	huiswerk maken
LV03_B5	werken aan werkstuk
LV03_B6	leren voor examen
LV03_B7	anders
LV04_A	Heb je bijles gehad?
LV04_B	# uren bijles per week
LV05	Goed voorbereid op SO?
LV06	Beter/slechter voorbereid dan klasgenoten?
LV07	Gemiddeld cijfer SO*10
LV08	Goed voorbereid op CSE?
LV09	Beter/slechter voorbereid dan klasgenoten?
LV10	Kans in 1* te slagen
LV11_A	herinneren wat je geleerd hebt

LV11_B	kalm blijven, geen paniek
LV11_C	tijd goed verdelen
LV11_D	goed geconcentreerd blijven
LV11_E	bedenken hoe te werk te gaan
LV11_F	systematisch te werk gaan
LV11_G	creativiteit gebruiken
LV11_H	verstandige gokstrategie
LV11_I	antwoorden controleren
LV_EXVA	Schaal examenvaardigheden
LV13_A	veel contact met klasgenoten
LV13_B	liever in andere klas
LV13_C	hadden leuke klas
LV13_D	kon goed met klasgenoten overweg
LV13_E	voelde me soms alleen in klas
LV13_F	vond leuk met klasgenoten om te gaan
LV_KLAS	Welbevinden met klasgenoten
LV15_A1	#goede vrienden
LV15_A2	#goede vriendinnen
LV15_B1	heb geen vrienden
LV15_B2	school
LV15_B3	buurt waar ik woon
LV15_B4	(sport-)club/vereniging
LV15_B5	school en daarbuiten
LV15_B6	internet
LV15_B7	anders
LV22_A	hoe laat je thuis moet zijn
LV22_B	of je een bijbaantje mag hebben
LV22_C	of je thuis alcohol mag drinken
LV22_D	of je op feestjes alcohol mag drinken
LV22_E	naar welke tv-programma's je kijkt
LV22_F	welke vervolgopleiding je gaat volgen
LV22_G	wanneer je je huiswerk maakt
LV22_H	hoe lang internet en/of computerspellen speelt
LV22_I	welke internetsites je bezoekt
LV22_J	of je uit mag gaan
LV22_K	hoe je op vakantie gaat / zelf op vak. mogen
LV22_L	van welke verenigingen je lid bent
LV22_M	of je softdrugs mag gebruiken
LV22_N	of je uit huis gaat wonen
LV22_O	of je de 'pil' mag gebruiken
LV22_P	of je naar de kerk gaat
LV22_Q	waar je je zak- kleedgeld aan besteedt
LV_AUTON	Autonoom beslissen
LV_AUTH	'Authoritative' opv.stijl
LV_FFPI1	Extraversie
LV_FFPI2	Mildheid
LV_FFPI3	Ordelijkheid
LV_FFPI4	Emotionele Stabiliteit

LV_FFPI5	Intellectuele Autonomie
LV_NFACT	Aantal valide factorscores

LITERATUUR

- Baumrind, D. (1978). Parental disciplinary patterns and social competence in children. *Youth and Society*, 9, 238-276.
- Heesters, K., Berkel, S. van, Schoot, F. van der, & Hemker, B. (2007). *Balans van het leesonderwijs aan het einde van de basisschool: Uitkomsten van de vierde peiling in 2005*. Arnhem: Cito.
- Hendriks, A. A. J., Hofstee, W. K. B., & De Raad, B. (1999). *Handleiding bij de Five Factor Personality Inventory (FFPI)*. Lisse: Swets Test Publishers.
- Hendriks, A. A. J., Hofstee, W. K. B., & De Raad, B. (2011). *Handleiding bij de Five Factor Personality Inventory II (FFPI-II)*. Houten: Bohn Stafleu van Loghum.
- Hendriks, A. A. J., Kuyper, H., Offringa, G. J., & Werf, M. P. C. van der (2008). Assessing Young Adolescents' Personality With the Five-Factor Personality Inventory. *Assessment*, 15, 304-316.
- Peetsma, T. T. D., Wagenaar, E., & Kat, E. de (2001). School motivation, future time perspective and well-being of high school students in segregated and integrated schools in the Netherlands and the role of ethnic self-description. In J. Koppen, I. Lunt, & C. Wulf (Eds.). *Education in Europe; Culture, Values, Institutions in transition* (pp. 54-74). Münster/New York: Waxmann.
- Steinberg, L., Elmen, J. D., & Mounts, N. S. (1989). Authoritative parenting, psychosocial maturity, and academic success among adolescents. *Child Development*, 60, 1424-1436.
- Ten Berge, J.M.F & Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, 69, 613-625.
- Veenstra, D., & Kuyper, H. (2004). Effective Students and Families: The Importance of Individual Characteristics for Achievement in High School. *Educational Research and Evaluation*, 10, 41-70.
- Zijsling, D., Keuning, J., Kuyper, H., Batenburg, T. van, & Hemker, B. (2009). *Technisch rapport eerste meting van COOL⁵⁻¹⁸ in het derde leerjaar van het voortgezet onderwijs*. Groningen: GION.

BIJLAGEN

1 Scoretabel 1^{ste} tijdvak examen HAVO Nederlands

Score	Cijfer	Perc.	Z-score	T-score	Score	Cijfer	Perc.	Z-score	T-score
0	1.0	0	-4.3	7	27	6.4	65	0.4	54
1	1.3	0	-4.1	9	28	6.5	71	0.6	56
2	1.7	0	-3.9	11	29	6.7	77	0.7	57
3	2.0	0	-3.8	12	30	6.9	82	0.9	59
4	2.4	0	-3.6	14	31	7.1	86	1.1	61
5	2.6	0	-3.4	16	32	7.2	90	1.3	63
6	2.7	0	-3.2	18	33	7.4	93	1.4	64
7	2.9	0	-3.1	19	34	7.6	95	1.6	66
8	3.1	1	-2.9	21	35	7.8	97	1.8	68
9	3.3	1	-2.7	23	36	7.9	98	2.0	70
10	3.4	1	-2.5	25	37	8.1	99	2.1	71
11	3.6	1	-2.4	26	38	8.3	99	2.3	73
12	3.8	2	-2.2	28	39	8.5	100	2.5	75
13	4.0	3	-2.0	30	40	8.6	100	2.7	77
14	4.1	4	-1.9	31	41	8.8	100	2.8	78
15	4.3	5	-1.7	33	42	9.0	100	3.0	80
16	4.5	7	-1.5	35	43	9.1	100	3.2	82
17	4.6	9	-1.3	37	44	9.3	100	3.3	83
18	4.8	12	-1.2	38	45	9.4	100	3.5	85
19	5.0	16	-1.0	40	46	9.5	100	3.7	87
20	5.2	20	-0.8	42	47	9.6	100	3.9	89
21	5.3	25	-0.6	44	48	9.7	100	4.0	90
22	5.5	31	-0.5	45	49	9.7	100	4.2	92
23	5.7	37	-0.3	47	50	9.8	100	4.4	94
24	5.9	44	-0.1	49	51	9.9	100	4.6	96
25	6.0	51	0.1	51	52	10.0	100	4.7	97
26	6.2	58	0.2	52					

2a Scoretabel 1^{ste} tijdvak examen HAVO wiskunde A

Score	Cijfer	Perc.	Z-score	T-score	Score	Cijfer	Perc.	Z-score	T-score
0	1.0	0.0	-4.1	9	42	5.1	20.0	-0.8	42
1	1.1	0.0	-4.0	10	43	5.2	22.0	-0.7	43
2	1.1	0.0	-3.9	11	44	5.3	24.0	-0.7	43
3	1.2	0.0	-3.9	11	45	5.4	26.0	-0.6	44
4	1.2	0.0	-3.8	12	46	5.5	28.0	-0.5	45
5	1.3	0.0	-3.7	13	47	5.7	31.0	-0.4	46
6	1.3	0.0	-3.6	14	48	5.8	33.0	-0.4	46
7	1.4	0.0	-3.6	14	49	5.9	36.0	-0.3	47
8	1.4	0.0	-3.5	15	50	6.0	39.0	-0.2	48
9	1.5	0.0	-3.4	16	51	6.1	42.0	-0.1	49
10	1.6	0.0	-3.3	17	52	6.2	45.0	0.0	50
11	1.7	0.0	-3.2	18	53	6.3	48.0	0.0	50
12	1.8	0.0	-3.2	18	54	6.4	51.0	0.1	51
13	1.9	0.0	-3.1	19	55	6.5	54.0	0.2	52
14	2.0	0.0	-3.0	20	56	6.6	57.0	0.3	53
15	2.1	1.0	-2.9	21	57	6.8	61.0	0.4	54
16	2.3	1.0	-2.9	21	58	6.9	64.0	0.4	54
17	2.4	1.0	-2.8	22	59	7.0	67.0	0.5	55
18	2.5	1.0	-2.7	23	60	7.1	70.0	0.6	56
19	2.6	1.0	-2.6	24	61	7.2	73.0	0.7	57
20	2.7	1.0	-2.5	25	62	7.3	76.0	0.7	57
21	2.8	2.0	-2.5	25	63	7.4	78.0	0.8	58
22	2.9	2.0	-2.4	26	64	7.5	81.0	0.9	59
23	3.0	2.0	-2.3	27	65	7.6	83.0	1.0	60
24	3.1	2.0	-2.2	28	66	7.7	86.0	1.1	61
25	3.2	3.0	-2.2	28	67	7.9	88.0	1.1	61
26	3.4	3.0	-2.1	29	68	8.0	90.0	1.2	62
27	3.5	4.0	-2.0	30	69	8.1	91.0	1.3	63
28	3.6	4.0	-1.9	31	70	8.2	93.0	1.4	64
29	3.7	5.0	-1.8	32	71	8.3	94.0	1.4	64
30	3.8	5.0	-1.8	32	72	8.4	96.0	1.5	65
31	3.9	6.0	-1.7	33	73	8.5	97.0	1.6	66
32	4.0	7.0	-1.6	34	74	8.6	98.0	1.7	67
33	4.1	8.0	-1.5	35	75	8.7	98.0	1.8	68
34	4.2	9.0	-1.4	36	76	8.8	99.0	1.8	68
35	4.3	10.0	-1.4	36	77	9.0	99.0	1.9	69
36	4.5	11.0	-1.3	37	78	9.1	100.0	2.0	70
37	4.6	12.0	-1.2	38	79	9.3	100.0	2.1	71
38	4.7	13.0	-1.1	39	80	9.6	100.0	2.2	72
39	4.8	15.0	-1.1	39	81	9.8	100.0	2.2	72
40	4.9	16.0	-1.0	40	82	10.0	100.0	2.3	73
41	5.0	18.0	-0.9	41					

2b Scoretabel 1^{ste} tijdvak examen HAVO wiskunde B

Score	Cijfer	Perc.	Z-score	T-score	Score	Cijfer	Perc.	Z-score	T-score
0	1.0	0.0	-3.3	17	39	6.1	57.0	0.1	51
1	1.2	0.0	-3.2	18	40	6.2	60.0	0.2	52
2	1.5	0.0	-3.1	19	41	6.3	64.0	0.3	53
3	1.7	0.0	-3.0	20	42	6.4	67.0	0.4	54
4	1.9	0.0	-2.9	21	43	6.5	70.0	0.5	55
5	2.1	0.0	-2.8	22	44	6.6	73.0	0.6	56
6	2.2	0.0	-2.7	23	45	6.8	76.0	0.7	57
7	2.3	0.0	-2.7	23	46	6.9	78.0	0.8	58
8	2.4	0.0	-2.6	24	47	7.0	80.0	0.8	58
9	2.6	0.0	-2.5	25	48	7.1	83.0	0.9	59
10	2.7	1.0	-2.4	26	49	7.2	85.0	1.0	60
11	2.8	1.0	-2.3	27	50	7.3	86.0	1.1	61
12	2.9	1.0	-2.2	28	51	7.5	88.0	1.2	62
13	3.0	1.0	-2.1	29	52	7.6	89.0	1.3	63
14	3.1	2.0	-2.0	30	53	7.7	91.0	1.4	64
15	3.3	2.0	-2.0	30	54	7.8	92.0	1.5	65
16	3.4	3.0	-1.9	31	55	7.9	93.0	1.5	65
17	3.5	3.0	-1.8	32	56	8.0	94.0	1.6	66
18	3.6	4.0	-1.7	33	57	8.2	94.0	1.7	67
19	3.7	5.0	-1.6	34	58	8.3	95.0	1.8	68
20	3.8	6.0	-1.5	35	59	8.4	96.0	1.9	69
21	4.0	7.0	-1.4	36	60	8.5	97.0	2.0	70
22	4.1	8.0	-1.3	37	61	8.6	97.0	2.1	71
23	4.2	10.0	-1.3	37	62	8.7	98.0	2.2	72
24	4.3	12.0	-1.2	38	63	8.9	98.0	2.2	72
25	4.4	14.0	-1.1	39	64	9.0	99.0	2.3	73
26	4.5	16.0	-1.0	40	65	9.1	99.0	2.4	74
27	4.7	18.0	-0.9	41	66	9.2	99.0	2.5	75
28	4.8	20.0	-0.8	42	67	9.3	99.0	2.6	76
29	4.9	23.0	-0.7	43	68	9.4	100.0	2.7	77
30	5.0	26.0	-0.6	44	69	9.5	100.0	2.8	78
31	5.1	29.0	-0.6	44	70	9.6	100.0	2.8	78
32	5.2	33.0	-0.5	45	71	9.6	100.0	2.9	79
33	5.4	36.0	-0.4	46	72	9.7	100.0	3.0	80
34	5.5	40.0	-0.3	47	73	9.8	100.0	3.1	81
35	5.6	43.0	-0.2	48	74	9.8	100.0	3.2	82
36	5.7	47.0	-0.1	49	75	9.9	100.0	3.3	83
37	5.8	50.0	0.0	50	76	9.9	100.0	3.4	84
38	5.9	54.0	0.1	51	77	10.0	100.0	3.5	85

3 Scoretabel 1^{ste} tijdvak examen HAVO Engels

Score	Cijfer	Perc.	Z-score	T-score	Score	Cijfer	Perc.	Z-score	T-score
0	1.0	0.0	-4.1	9	25	5.0	23.0	-0.8	42
1	1.1	0.0	-4.0	10	26	5.2	27.0	-0.7	43
2	1.2	0.0	-3.9	11	27	5.4	31.0	-0.5	45
3	1.3	0.0	-3.7	13	28	5.5	35.0	-0.4	46
4	1.4	0.0	-3.6	14	29	5.7	39.0	-0.3	47
5	1.5	0.0	-3.5	15	30	5.9	44.0	-0.1	49
6	1.6	0.0	-3.3	17	31	6.1	48.0	0.0	50
7	1.7	0.0	-3.2	18	32	6.3	53.0	0.1	51
8	1.9	0.0	-3.1	19	33	6.5	58.0	0.3	53
9	2.1	0.0	-2.9	21	34	6.6	63.0	0.4	54
10	2.2	0.0	-2.8	22	35	6.8	67.0	0.5	55
11	2.4	0.0	-2.7	23	36	7.0	72.0	0.7	57
12	2.6	0.0	-2.5	25	37	7.2	76.0	0.8	58
13	2.8	1.0	-2.4	26	38	7.4	80.0	0.9	59
14	3.0	1.0	-2.3	27	39	7.6	84.0	1.1	61
15	3.2	2.0	-2.1	29	40	7.7	88.0	1.2	62
16	3.3	3.0	-2.0	30	41	7.9	91.0	1.3	63
17	3.5	4.0	-1.9	31	42	8.1	94.0	1.5	65
18	3.7	5.0	-1.7	33	43	8.3	96.0	1.6	66
19	3.9	6.0	-1.6	34	44	8.5	97.0	1.7	67
20	4.1	8.0	-1.5	35	45	8.7	99.0	1.9	69
21	4.3	11.0	-1.3	37	46	8.9	99.0	2.0	70
22	4.4	13.0	-1.2	38	47	9.3	100.0	2.1	71
23	4.6	16.0	-1.1	39	48	9.6	100.0	2.3	73
24	4.8	19.0	-0.9	41	49	10.0	100.0	2.4	74