



# De koppeling van de Staatsexamens Nederlands als Tweede Taal Lezen en Luisteren aan het Europees Referentiekader voor Talen

## Fase 2 Standaardisatie en bepaling van grensscores

Henk Kuijper, Matthieu Brinkhuis, Laura van Hofwegen  
Cito, Arnhem, maart 2010

© *Stichting Cito Instituut voor Toetsontwikkeling Arnhem 2011. Auteursrecht voorbehouden.*

Niets uit deze uitgave mag zonder voorafgaande schriftelijke toestemming van Stichting Cito Instituut voor Toetsontwikkeling worden openbaar gemaakt en/of verveelvoudigd door middel van druk, fotokopie/fotografie, scanning, computersoftware of andere elektronische verveelvoudiging of openbaarmaking, microfilm, geluidskopie, film- of videokopie of op welke wijze dan ook.

1	Inleiding	4
2	De procedure van standaardbepaling	5
2.1	Het doel van de standaardbepaling	5
2.2	Werkwijze	5
2.2.1	Keuze van de standaardbepalingsmethode	5
2.2.2	Selectie van de opgaven	6
2.2.3	Selectie van panelleden	6
2.2.4	Familiarisatie en training met het ERK	7
2.2.5	Het beoordelingsproces	7
3	Resultaten	9
3.1	Lezen	10
3.1.1	Beoordelaarsovereenstemming en beoordelaars- betrouwbaarheid	10
3.1.2	ERK-niveaus van Programma I en II	12
3.1.3	De bepaling van grensscores bij relevante ERK-niveaus	14
3.1.4	Samenvatting en conclusies Lezen	16
3.2	Luisteren	19
3.2.1	Beoordelaarsovereenstemming en beoordelaars- betrouwbaarheid	19
3.2.2	ERK-niveaus van Programma I en II	21
3.2.3	De bepaling van grensscores bij relevante ERK-niveaus	23
3.2.4	Samenvatting en conclusies Luisteren	26
4	Samenvatting en conclusies	28
Bijlage 1	Samenstelling van de beoordelaarsgroepen Lezen en Luisteren	30
Bijlage 2	Materiaal bij de familiarisatie	31
Bijlage 3	Aanbiedingsvolgorde van de opgaven in de standaardbepaling	39
Bijlage 4	Scoreformulieren Lezen en Luisteren	45
Bijlage 5	Grensscores volgens beoordelaars en scoreverdeling grenskandidaten B1 en B2 voor Lezen	50
Bijlage 6	Grensscores volgens beoordelaars en scoreverdeling grenskandidaten B1 en B2 voor Luisteren	52
Referenties		54

# 1 Inleiding

In 2003 publiceerde de Raad van Europa een concept handleiding voor het relateren van taalexamens aan het Common European Framework for Languages (Council of Europe, 2003). In 2009 verscheen een hernieuwde versie (Council of Europe, 2009). In deze Handleiding worden verschillende methodes beschreven om de claim op een koppeling van een examen met het Gemeenschappelijk Europees Referentiekader voor Moderne Vreemde Talen (ERK) te onderbouwen. De Handleiding heeft als doel om belanghebbenden mogelijke procedures aan de hand te doen waarmee men evidentie kan aanvoeren voor de relaties van een examen met het ERK..

In de Handleiding worden drie fases van koppeling onderscheiden:

- *Specificatie*: nagaan op welke toelichtende ERK-schalen en descriptoren toetsen en examens betrekking hebben. Een eerste inschatting te geven op het niveau van de toetsen en examens.
- *Standaardisatie*: via een in de Handleiding omschreven procedure experts vragen om de examenopgaven te classificeren in termen van ERK (er ERK-niveaus aan te koppelen). Op basis daarvan kan worden bepaald welke score een kandidaat bij een bepaald examen zou moeten behalen om te kunnen zeggen dat hij/zij een bij dat examen relevant ERK-niveau beheerst).
- *Empirische validatie*: psychometrische validering van de gegevens verzameld in de standaardisatieprocedure. Deze validatie onderscheidt verder interne validatie van een examensysteem door middel van psychometrische onderbouwing en externe validatie door het leggen van relaties tussen een examen en vergelijkbare andere examens die reeds een relatie hebben gelegd met het ERK.

In de eerste fase is de claim op koppeling aan het ERK alleen gebaseerd op een inhoudelijke beschrijving. Deze fase (met varianten) is door een aantal instituten in Europa en daarbuiten gevolgd bij het relateren van hun examens aan het ERK. Het Cito heeft deze procedure al eerder toegepast op de centrale examens leesvaardigheid van VMBO tot VWO voor zeven moderne vreemde talen. (Noijons e.a. 2006, 2007).

De tweede fase – standaardisatie – leidt tot een versterkte claim op een koppeling, namelijk gebaseerd op basis van specificatie en standaardisatie. Ook deze fase is door het Cito voor de centrale examens leesvaardigheid uitgevoerd.

In de derde fase wordt nagegaan of de eerdere claims worden bevestigd door empirische validatie, waarin de relaties tussen verschillende examens binnen één systeem empirisch worden onderzocht en examens gerelateerd kunnen worden aan andere examens.

Voor de Staatsexamens NT2 is al eerder onderzoek uitgevoerd naar de relatie tussen de vaardigheden Spreken en Schrijven en het ERK. Uit dit onderzoek bleek dat voor beide vaardigheden Programma I correspondeert met ERK-niveau B1 en Programma II met ERK-niveau B2. Van dit onderzoek is verslag gedaan in Kuijper e.a. (2004a, 2004b, 2005a, 2005b) en Bechger e.a. (2008). Voor Lezen en Luisteren is de specificatiefase al uitgevoerd (Kuijper, 2009). Deze specificatiefase wees uit dat zowel voor Lezen als voor Luisteren Programma I teksten en opgaven zich in het niveau bereik B1 bevinden, terwijl de opgaven en teksten voor Programma II in het B2-niveau bereik liggen.

In het huidige verslag worden de activiteiten en resultaten beschreven met betrekking tot de standaardisatiefase voor de Staatsexamens Luisteren en Lezen. De interne validatie van de Staatsexamens is geen onderwerp van deze studie. Wel worden kalibratiegegevens uit de Staatsexamens gebruikt om grensscores voor de diverse niveaus op de Staatsexamens en de 500-schalen vast te leggen.

In hoofdstuk 2 wordt de gevolgde procedure beschreven. Hoofdstuk 3 bevat de resultaten van de standaardbepaling voor Lezen en Luisteren. Het rapport wordt in hoofdstuk 4 afgesloten met een samenvatting en conclusies.

## 2 De procedure van standaardbepaling

In dit hoofdstuk wordt verslag gedaan van de werkwijze van de standaardbepaling. De resultaten worden beschreven in hoofdstuk 3.

### 2.1 Het doel van de standaardbepaling

Het doel van de standaardbepaling is een relatie te leggen tussen de lees- en luisterexamens van het Staatsexamen NT2 enerzijds en het ERK anderzijds. Deze relatie wordt allereerst bepaald door voor afzonderlijke opgaven het niveau te bepalen dat minimaal nodig is om de opgave goed te maken. Dit leidt tot een beschrijving van de opgaven in de examens in termen van ERK-niveaus.

De tweede stap bestaat eruit om op grond van deze niveaubepaling aanwijzingen te vinden voor een grensscore die onderscheid maakt tussen kandidaten die een ERK-niveau wel of niet hebben bereikt. Overigens hoeft deze grensscore niet samen te vallen met de cesuur die in de Staatsexamens bepaalt of een kandidaat voor het examen slaagt. Het is in principe mogelijk om zowel een zak-slaagcesuur te hebben als één of meerdere grensscores die informatie geven over het ERK-niveau van de kandidaten.

### 2.2 Werkwijze

In navolging van de Handleiding (Council of Europe, 2009) hebben we tijdens deze standaardisatiefase de volgende stappen ondernomen.

1. Familiarisatie. In deze stap moeten de beoordelaars bekend raken met het ERK .
2. Training met voorbeeldopgaven.
3. Standaardbepaling. Deze stap bestaat uit twee substappen. In de eerste substap bepalen de beoordelaars onafhankelijk van elkaar het niveau van de opgaven. Nadat deze oordelen zijn geanalyseerd, worden in een tweede ronde de opgaven waarover de minste overeenstemming is bereikt gezamenlijk besproken. Na deze bespreking geeft iedere beoordelaar aan of hij bij zijn eerste beoordeling blijft, of dat hij, gezien de argumentatie tijdens de bespreking, tot een ander oordeel is gekomen.
4. Data-analyse ter validering van de standaarden. In de data-analyse worden de resultaten van de standaardbepaling geanalyseerd en gerelateerd aan de empirische data van kandidaten die het Staatsexamen NT2 hebben afgelegd.

In de volgende paragrafen wordt de gevolgde werkwijze beschreven.

#### 2.2.1 Keuze van de standaardbepalingmethode

Er zijn vele standaardbepalingprocedures. In de Handleiding van de Raad van Europa worden twee hoofdcategorieën onderscheiden: de examinee centred (EC) en de test centred (TC) methoden. De eerste categorie wordt gebruikt voor productieve vaardigheden, waarbij de prestaties van kandidaten kunnen worden geobserveerd en geclassificeerd. De test centred methode wordt gebruikt voor de receptieve vaardigheden. Hier hebben we geen zicht op leerlingprestaties, zoals bij de productieve vaardigheden, maar moeten we een oordeel geven over opgaven, waarbij moet worden aangegeven welk niveau een kandidaat moet hebben om een bepaalde opgave goed te kunnen maken. Tabel 2-1 geeft een overzicht van de methoden die in de Handleiding worden vermeld.

Tabel 2-1 Methoden van Standaardbepaling

Method	Section	Class
Tucker-Angoff	6.3.	T-C
The Yes-No Method	6.4.1.	T-C
The Extended Tucker-Angoff Method	6.4.2.	T-C
The Contrasting Groups Method	6.5.1.	E-C
The Borderline Group Method	6.5.2.	E-C

The Body of Work Method	6.6.	E-C
The Item-descriptor Matching Method	6.7.1.	T-C
The Basket Method	6.7.2.	T-C
The Bookmark Method	6.8.	IRT
A Cito Variation on the Bookmark Method	6.9.	IRT

Bron: Council of Europe. Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual, pag 61.

In dit onderzoek is gebruik gemaakt van de Basket Method (verder te noemen mandjesmethode). In deze mandjesmethode wordt iedere beoordelaar gevraagd om elke opgave in een 'mandje' te plaatsen waarvan de naam overeenkomt met het laagste niveau dat nodig geacht wordt om de opgave goed te maken. Er zijn zes mandjes, genaamd A1, A2, B1, B2, C1 en C2.

Als een opgave geplaatst wordt in mandje B1, dan betekent dit dat volgens de beoordelaar iemand op het niveau van B1 de opgave goed dient te kunnen beantwoorden. Dit impliceert dat beheersing van deze opgave ook wordt verondersteld bij alle personen met een hoger niveau (B2 en hoger). Echter, van iemand met het niveau A2 (of lager) kan niet worden verwacht dat hij deze opgave beheerst. Deze methode van standaardbepaling is ontwikkeld voor het project DIALANG (2002). De methode is gekozen omdat zij goed werkbaar is en betrouwbare en bruikbare resultaten oplevert. Daarnaast is deze methode reeds eerder door het Cito gebruikt, o.a. in het koppelingsonderzoek van de Centrale examens leesvaardigheid in het voortgezet onderwijs (Noijons e.a., 2006, 2007). De mandjesmethode wordt toegelicht in paragraaf 2.4 en 2.5. Voor een uitgebreidere beschrijving van de mandjesmethode verwijzen we naar de Handleiding, (Council of Europe, 2009, pag. 75-77).

### 2.2.2 Selectie van de opgaven

In de standaardbepaling zijn dezelfde lees- en luisteropgaven voor het Staatsexamen NT2 gebruikt als in de specificatiefase. Alle opgaven uit deze examens zijn in de standaardbepaling opgenomen. Aan elke vaardigheid zijn 20 opgaven toegevoegd uit de Profieltoets NT2. De helft van deze opgaven waren op NT2-niveau 1, de andere helft op NT2-niveau 2. Deze niveaus zijn grofweg te vergelijken met de ERK niveaus A1 en A2. Deze opgaven zijn meegenomen in de standaardbepaling om de beoordelaars ook de mogelijkheid te geven opgaven van lagere niveaus te beoordelen, waardoor met name de B1 opgaven beter kunnen worden onderscheiden van opgaven die bedoeld waren voor de A1 en A2 niveaus. Tabel 2-2 geeft een overzicht van de toetsen waaruit de opgaven voor de standaardbepaling afkomstig zijn.

Tabel 2-2 Herkomst lees- en luisteropgaven in de standaardbepaling

Examen	Versie	Aantal opgaven
STEX Lezen I	Examen mei 2006	40
STEX Lezen II	Examen maart 2006	40
Lezen Profieltoets NT2-niveau 1	Januari 2008 Variant A	10
Lezen Profieltoets NT2-niveau 2	Januari 2008 Variant B	10
<b>Totaal leesopgaven</b>		<b>100</b>
STEX Luisteren I	Examen januari 2006	40
STEX Luisteren II	Examen maart 2006	40
Luisteren Profieltoets NT2-niveau 1	2006-2 Variant A	10
Luisteren Profieltoets NT2-niveau 2	2006-2 Variant B	10
<b>Totaal luisteropgaven</b>		<b>100</b>

### 2.2.3 Selectie van panelleden

Bij de selectie van beoordelaars is ervan uitgegaan dat de afnemers van de Staatsexamens NT2 in ruime mate vertegenwoordigd moeten zijn. Het merendeel van de beoordelaars is betrokken bij NT2-opleidingen, gericht op één van de beide Staatsexamenprogramma's. Daarnaast is het Certificaat Nederlands als Vreemde Taal (CNaVT) benaderd om een deelnemer naar de standaardsettingbijeenkomsten af te vaardigen. Dit was voor Luisteren mogelijk. Voor Lezen kon het CNaVT geen medewerker afvaardigen. Wel waren er twee panelleden aanwezig die tot voor kort hebben meegewerkt aan de constructie van de examens van het CNaVT. Tot slot is met de

Staatsexamencommissie afgesproken om bij zowel Lezen als Luisteren één lid van de commissie aan de standaardbepaling te laten deelnemen. Verder nam voor Lezen een vertegenwoordiger van het Cito en voor Luisteren een vertegenwoordiger van Bureau ICE deel aan de standaardbepaling. Een ander selectie criterium was dat deelnemers aan de panels goed op de hoogte moesten zijn van ERK en zijn niveaus. Bijlage 1 bevat de samenstelling van de panels Lezen en Luisteren. In totaal hebben 11 beoordelaars deelgenomen aan de standaardbepaling lezen en 10 aan de standaardbepaling Luisteren.

#### **2.2.4 Familiarisatie en training met het ERK**

Zowel voor Lezen als voor Luisteren zijn er twee bijeenkomsten georganiseerd met een tussentijd van 14 dagen.

De eerste bijeenkomst van elke vaardigheid begon met een korte inleiding over het doel en de werkwijze van de standaardbepaling.

Vervolgens vond een familiarisatie plaats door middel van 2 sorteeroefeningen. In deze oefeningen werden de niveaubeschrijvingen van de Globale Schaal van het ERK (Nederlandse Taalunie, 2006, pag. 26) en die van de Schaal Algemene Leesvaardigheid (Nederlandse Taalunie, 2006, pag. 66) in willekeurige volgorde aan de panelleden voorgelegd. Hen werd gevraagd aan elke uitspraak het juiste niveau toe te kennen. Deze fase van familiarisatie werd afgesloten met een discussie over de onderscheidende kenmerken van de verschillende niveaus. Tijdens deze discussie werd een korte powerpoint gebruikt en een formulier uitgereikt met 'vuistregels', beide gebaseerd op eerder onderzoek naar de ERK-niveaus van de centrale examens leesvaardigheid (Noijons e.a., 2006, 2007). Het materiaal is opgenomen in bijlage 2.

Na deze familiarisatiefase volgde een trainingsfase. Tijdens deze training werden de panelleden 7 voorbeeldopgaven uit de Profieltoets en een voorbeeldexamen van de Staatsexamens NT2 voorgelegd, variërend van niveau A1 tot B2. Elk panellid werd gevraagd aan elke opgave een ERK-niveau toe te kennen door de volgende vraag te beantwoorden:

*'Geef bij elke opgave aan welk niveau (A1, A2, B1, B2, C1, C2) minimaal vereist is om de opgave goed te kunnen maken.'*

Vervolgens volgde er voor elke opgave een discussie waarin panelleden hun keuze toe konden lichten en waarin getracht werd tot overeenstemming te komen. Deze training was ervoor bedoeld om samen met de familiarisatie te komen tot een meer gemeenschappelijke standaard voor het feitelijke beoordelingsproces dat op de training zou volgen. De opgaven die in deze training werden gebruikt werden uiteraard niet meegenomen in de standaardbepaling.

#### **2.2.5 Het beoordelingsproces**

Na de familiarisatie en de training begon de uiteindelijke de standaardbepaling, waarbij de panelleden moesten aangeven op welk ERK-niveau opgaven zich bevonden.

De beoordelaars kregen de teksten van Staatsexamens Programma I en II en die van de Profieltoets door elkaar aangeboden, met bij elke tekst de vragen in de volgorde van deze examens. Bijlage 3 bevat een overzicht van de aanbiedingsvolgorde van de opgaven Lezen en Luisteren.

Het beoordelingsproces bestond uit twee stappen.

##### **Stap 1**

Op dag 1 beantwoordden de panelleden voor alle 100 opgaven dezelfde vraag als in de trainingsfase.

*'Geef bij elke opgave aan welk niveau (A1, A2, B1, B2, C1, C2) minimaal vereist is om de opgave goed te kunnen maken.'*

De panelleden gaven het niveau aan op een optisch leesbaar antwoordformulier. Dit formulier is opgenomen in bijlage 4.

In tegenstelling tot in de trainingsfase werkten de panelleden nu onafhankelijk van elkaar. De tijd om de 100 opgaven van een niveau-indicatie te voorzien was ruim voldoende. De gemiddelde tijd die de panelleden nodig hadden was bij Lezen 2 uur en 16 minuten. De snelste beoordelaar had 1 uur en 26 minuten nodig, de langzaamste 3 uur en 22 minuten. Bij Luisteren was de gemiddelde beoordelingstijd

2 uur en 38 minuten. Hier was de snelste beoordelaar na 2 uur en 5 minuten klaar, terwijl de langzaamste 3 uur en 30 minuten nodig had.

#### Stap 2

Op de tweede dag werden de Staatsexamenopgaven waarover op grond van een tussentijdse analyse de minste overeenstemming was gezamenlijk besproken. Deelnemers gaven hun argumenten voor hun niveautoekenning aan de opgaven en lichtten in een discussie hun motieven toe. Na afloop van elke discussie over een opgave werd elk van de panelleden afzonderlijk gevraagd tot welke niveautoekenning hij op basis van de discussie was gekomen. Dit kon hetzelfde oordeel zijn als tijdens de eerste zitting of een bijstelling hiervan. Ook in de tweede zitting was het niet de bedoeling tot een gezamenlijke uitspraak te komen. Bij de discussie werd gebruik gemaakt van een databestand met een overzicht van de niveautoekenningen van de eerste bijeenkomst. Wanneer de niveautoekenning van een panellid verschilde van die van de eerste bijeenkomst werd deze wijziging in het databestand doorgevoerd.



### **3 Resultaten**

In dit hoofdstuk worden de resultaten van de standaardbepaling gepresenteerd. In paragraaf 3.1 komen de resultaten van de standaardbepaling Lezen aan de orde, op de resultaten van de standaardbepaling Luisteren wordt in paragraaf 3.2 ingegaan.

Beide paragrafen hebben dezelfde opbouw. Achtereenvolgens wordt ingegaan op

- de beoordelaarsovereenstemming en –betrouwbaarheid,
- de niveau-indelingen van de opgaven voor Programma I en Programma II,
- de bepaling van grensscores bij relevante ERK-niveaus voor beide examenprogramma's.

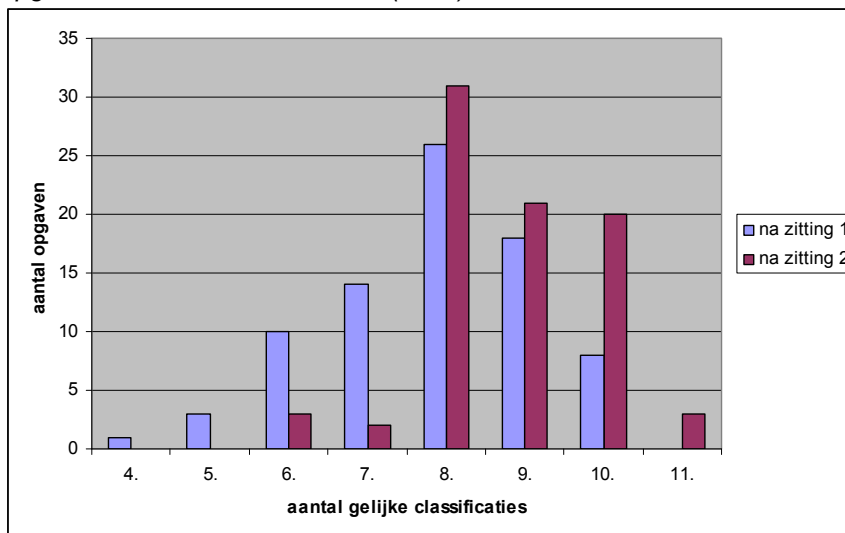
### 3.1 Lezen

#### 3.1.1 Beoordelaarsovereenstemming en beoordelaarsbetrouwbaarheid

Een voldoende mate van beoordelaarsovereenstemming en beoordelaarsbetrouwbaarheid is een voorwaarde om later uitspraken te kunnen doen over het niveau van de opgaven in een examen en voor het vaststellen van grensscores.

Inzicht in de beoordelaarsovereenstemming kan worden verkregen door na te gaan hoeveel beoordelaars een opgave in hetzelfde 'mandje' hebben geplaatst, dat wil zeggen hetzelfde minimumniveau hebben aangegeven dat nodig is om een opgave goed te maken. De beoordelaars konden elke opgave plaatsen in één van de 6 ERK-niveaus (A1 t/m C2). Als de beoordelaars een grote mate van onenigheid zouden hebben over het niveau van een opgave, zouden de oordelen bijvoorbeeld gelijk verdeeld zijn over deze niveaus. In dat geval zouden vijf niveaus door 2 beoordelaars en een zesde niveau door één beoordelaar zijn toegekend. Als de beoordelaars het allemaal eens zijn over het niveau dat nodig is om een opgave goed te maken, zouden 11 beoordelaars hetzelfde niveau aan een opgave toekennen. Figuur 3.1-1 laat zien in welke mate de beoordelaars het na zitting 1 en zitting 2 eens zijn over de niveautoekenning aan opgaven van Programma I en II. De x-as geeft aan hoeveel beoordelaars het een zijn over het niveau van een opgave. Op de y-as kan worden afgelezen om hoeveel opgaven het daarbij gaat.

*Figuur 3.1-1 Mate van overeenstemming van panelleden (n=11) over het vereiste niveau van de lees opgaven in Staatsexamen I en II (k=80)*



Figuur 3.1-1 laat allereerst zien dat de tweede zitting tot meer overeenstemming heeft geleid. Na de tweede zitting waren er drie opgaven die van 6 beoordelaars hetzelfde niveau kregen toegekend. Over 2 opgaven waren 7 beoordelaars het eens. Over de overige 75 van de 80 opgaven waren 8 of meer beoordelaars het eens over de niveautoekenning. De afwijkende beoordelaars plaatsten de opgaven altijd in een niveau dat direct lag naast het door de meerderheid toegekende niveau.

Tabel 3.1-1 geeft een overzicht van de beoordelaarsbetrouwbaarheid, beoordelaarsovereenstemming en resultaten van een variantieanalyse. De beoordelaarsbetrouwbaarheid geeft aan in hoeverre de rangorde van de oordelen over de beoordelaars consistent is, terwijl de beoordelaarsovereenstemming aangeeft in hoeverre opgaven op dezelfde wijze worden geclassificeerd. Beide maten lopen van 0 bij minimale overeenstemming/betrouwbaarheid tot 1 bij een maximale overeenstemming/betrouwbaarheid. De variantieanalyse geeft aan in welke mate de variantie in de oordelen is toe te schrijven aan de opgaven, de beoordelaars aan interactie- en erroreffecten (residu).

*Tabel 3.1-1 Variantiecomponenten, beoordelaarsovereenstemming en beoordelaarsbetrouwbaarheid Lezen*

Beoordelaarsbetrouwbaarheid	.95
Beoordelaarsovereenstemming bij 11	.94

beoordelaars (Rho2)	
Geschatte bijdrage van variantiecomponenten:	
Opgaven	60%
Beoordelaars	4%
Residu (interactie, error)	36%

Zowel de beoordelaarsbetrouwbaarheid als de beoordelaarsovereenstemming zijn hoog. De variantie-analyse wijst uit dat met 60% de meeste variantie verklaard wordt door de opgaven, terwijl de factor beoordelaar slechts 4% van de variantie verklaart. Dit is bevredigend, hoewel de 36% residuvariantie er op wijst dat er geen volledige overeenstemming is.

In tabel 3.1-2 wordt een overzicht gegeven van de absolute verschillen tussen de niveau-indeling door de 11 beoordelaars. Het gemiddelde (Gem.) geeft per beoordelaar de gemiddelde score over alle 80 STEX-opgaven. Delta is het absolute verschil tussen de gemiddelde score van elke beoordelaar en het totaal gemiddelde (Gem-totaal). Bij deze berekening zijn de waarden 1-6 toegekend aan de niveaus A1-C2. Een groot verschil kan sterk afwijkende beoordelaars detecteren. Tabel 3.1-2 geeft verder voor elke beoordelaar de Rbt-waarde. Naar analogie met de Rit is de Rbt de correlatie tussen de beoordelaarscores en de beoordelaargemiddelden.

Tabel 3.1-2 Gemiddelde scores en Rbt's van de beoordelaars over alle STEX-opgaven Lezen.

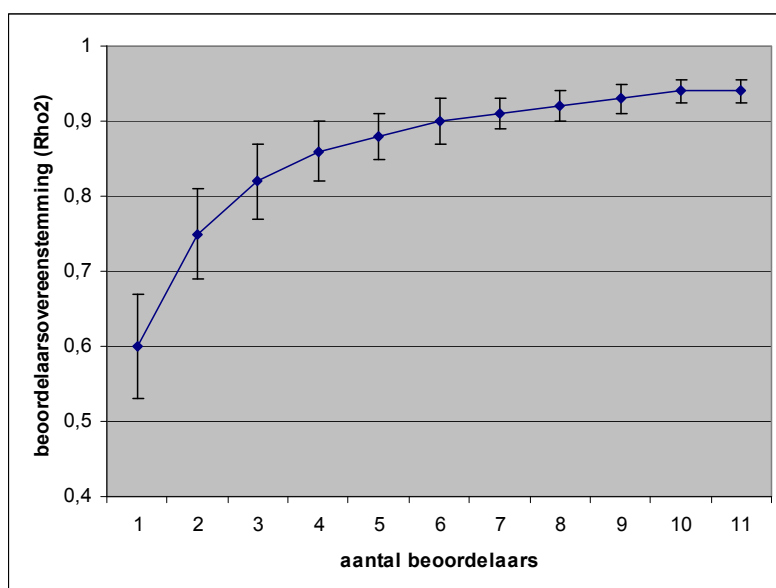
	Beoordelaars											Gem-totaal
	1	2	3	4	5	6	7	8	9	10	11	
Gem*.	3,49	3,51	3,79	3,61	3,50	3,86	3,79	3,66	3,80	3,53	3,54	3,64
Delta*	0,16	0,13	0,14	0,03	0,14	0,22	0,14	0,02	0,16	0,12	0,11	0,12
Rbt*	0,86	0,90	0,83	0,86	0,87	0,78	0,77	0,75	0,71	0,83	0,83	0,82

\* Afgerond op twee decimalen

Het blijkt dat de gemiddelde niveau-indeling voor alle beoordelaars ligt tussen B1 (3) en B2 (4). De gemiddelde delta over alle beoordelaars bedraagt 0,12 van een niveau, terwijl de grootste delta 0,22 bedraagt. Ook de Rbt geeft hoge waarden aan.

Het panel dat voor de standaardsetting werd gebruikt bestond uit 11 beoordelaars. Figuur 3.2-1 geeft weer hoe de beoordelaarsovereenstemming zich verhoudt tot het aantal panelleden. In deze figuur wordt de geschatte beoordelaarsovereenstemming per aantal beoordelaars met een ander panel van dezelfde omvang en samenstelling en de 90%-betrouwbaarheidsintervallen van deze schattingen weergegeven.

Figuur 3.1-2 Relatie tussen het aantal panelleden en de beoordelaarsovereenstemming (Rho2)



Figuur 3.1-2 laat zien dat de geschatte beoordelaarsovereenstemming vanaf 10 beoordelaars niet meer toeneemt en het betrouwbaarheidsinterval van de schattingen blijft vanaf dit aantal gelijk. Het aantal beoordelaars dat we bij deze standaardbepaling hebben gebruikt blijkt op grond hiervan voldoende te zijn.

Bovenstaande resultaten wijzen erop dat de beoordelaarsovereenstemming en –betrouwbaarheid voldoende hoog zijn om in de volgende paragrafen een beschrijving te geven van het niveau waarop de opgaven in Staatsexamens NT2 zich bevinden en een advies uit te brengen over de grensscores op beide examens voor de ERK-niveaus die voor beide programma's relevant zijn.

### 3.1.2 ERK-niveaus van Programma I en II

Tijdens de standaardbepaling hebben de beoordelaars niveaus toegekend aan de opgaven van Staatsexamen Programma I en II. Tabel 3.1-3 en 3.1-4 geven het resultaat weer van deze niveautoekenningen. In deze tabellen staat voor elke beoordelaar het aantal opgaven dat op een bepaald niveau is ingedeeld.

3.1-3 Niveautoekenning aan de opgaven van Staatsexamen Lezen Programma I Lezen (k=40)

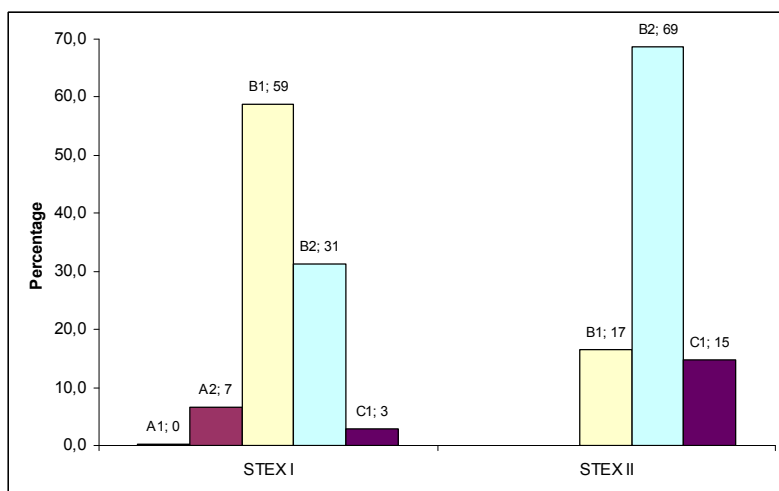
	Beoordelaars											Gem.
	1	2	3	4	5	6	7	8	9	10	11	
A1	0	0	0	0	1	0	0	0	0	0	0	<b>0,1</b>
A2	3	2	1	2	6	2	1	2	1	8	1	<b>2,6</b>
B1	27	28	23	29	21	18	22	22	16	24	29	<b>23,5</b>
B2	10	10	16	9	11	17	13	15	21	6	10	<b>12,5</b>
C1	0	0	0	0	1	3	4	1	2	2	0	<b>1,2</b>
C2	0	0	0	0	0	0	0	0	0	0	0	<b>0</b>

Tabel 3.1-4 Niveautoekenning aan de opgaven van Staatsexamen Lezen Programma II Lezen (k=40)

	Beoordelaars											Gem.
	1	2	3	4	5	6	7	8	9	10	11	
A1	0	0	0	0	0	0	0	0	0	0	0	<b>0</b>
A2	0	0	0	0	0	0	0	0	0	0	0	<b>0</b>
B1	10	10	2	2	8	4	5	10	2	6	9	<b>6,2</b>
B2	28	27	28	34	29	24	27	22	36	28	27	<b>28,2</b>
C1	2	3	10	4	3	12	8	8	2	6	3	<b>5,5</b>
C2	0	0	0	0	0	0	0	0	0	0	0	<b>0</b>

Uit tabellen 3.1-3 en 3.1-4 blijkt dat er een duidelijk verschil is tussen de niveautoekenningen aan opgaven van Programma I en II. Voor het merendeel van de Programma I opgaven heeft men volgens het beoordelaarspanel minimaal B1 nodig om deze opgaven goed te kunnen maken. Voor programma II heeft men voor het merendeel van de opgaven een vaardigheid minimaal op B2-niveau nodig. Aan elke opgave van programma I en II is door 11 beoordelaars een niveau toegekend. Figuur 3.1-3 geeft een overzicht van de percentages over de verschillende niveaus van deze beoordelingen.

*Figuur 3.1-3 verdeling van de beoordelingen van de leesopgaven over de ERK niveaus voor STEX I en STEX II. In percentages. (totaal 80 opgaven beoordeeld door 11 beoordelaars)*



Figuur 3.1-3 geeft het duidelijke niveauverschil tussen beide examens weer. De opgaven van Programma I worden in 59% van de gevallen als B1 ingeschat, terwijl de opgaven van Programma II in 69% van de beoordelingen B2 krijgen toegekend.

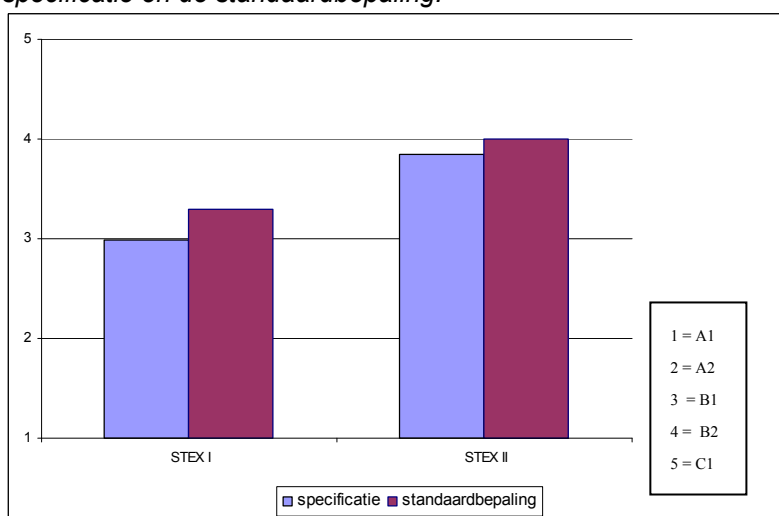
Wanneer we aan de niveaus A1 t/m C2 een waarde toekennen van 1-6, kunnen we per beoordelaar een gemiddelde niveauscore voor Programma I en Programma II berekenen. Tabel 3.1-5 geeft hier een overzicht van voor beide programma's.

Tabel 3.1-5 Gemiddeld vereist minimumniveau voor Programma I en II Lezen

	Beoordelaars											Gem.
	1	2	3	4	5	6	7	8	9	10	11	
Programma I	3,2	3,2	3,4	3,2	3,1	3,5	3,5	3,4	3,6	3,1	3,2	<b>3,3</b>
Programma II	3,8	3,8	4,2	4,1	3,9	4,2	4,1	4,0	4,0	4,0	3,8	<b>4,0</b>

De opgaven voor Programma I liggen gemiddeld iets boven B1. De opgaven van Programma II liggen gemiddeld op B2. Ook in de eerder uitgevoerde specificatiefase (Kuijper, 2009) is een niveau-indicatie van de opgaven uitgevoerd., Figuur 3.1-4 geeft de gemiddelde niveau-inschatting weer voor Programma I en II in de specificatiefase en de standaardbepaling.

Figuur 3.1-4 Gemiddelde niveautoekenning van de leesopgaven Programma I en II volgens de specificatie en de standaardbepaling.



Zowel volgens de specificatie als de standaardbepaling ligt Lezen I in het B1-niveaubereik en Lezen II in het B2-niveaubereik, waarbij in de standaardbepaling het niveau van de opgaven in beide examens iets hoger wordt ingeschat. De verschillen tussen de resultaten van beide methodieken zijn echter gering.

### 3.1.3 De bepaling van grensscores bij relevante ERK-niveaus

In de vorige paragraaf is beschreven dat de Leesexamens van Staatsexamen I en II zich bevinden op respectievelijk het niveaubereik B1 en B2 van het ERK.

Op grond van de resultaten van de standaardsetting kunnen we een grensscore vaststellen die op een examen nodig is om een voor dat examen relevant ERK-niveau te behalen. We kunnen deze grensscores vervolgens vergelijken met de cesuren die nu bij de Staatsexamens worden gehanteerd. Het is belangrijk te benadrukken dat het gezien de functie van de Staatsexamens NT2 niet noodzakelijk is om de huidige cesuren te veranderen op grond van de resultaten die in deze paragraaf besproken worden. De Staatsexamens kennen immers een functionele cesuur die gebaseerd is op het doel van de examens: het toetsen van een startkwalificatie voor de opleidingen en beroepen waarvoor beide programma's bedoeld zijn. De Staatsexamens NT2 zijn niet in eerste instantie bedoeld om een bepaald ERK-niveau te meten. De resultaten die in deze paragraaf besproken maken het mogelijk om scores te relateren aan een ERK niveau, maar hoeven er niet per se toe te leiden de huidige cesuren te veranderen.

Bij het bepalen van een grensscore voor een bepaald ERK-niveau hanteren we twee criteria.

1. Het 100%-criterium. Hierbij wordt de grensscore voor het bereiken van een ERK-niveau als volgt bepaald: voor het behalen van ERK-niveau x moet een kandidaat alle opgaven van niveau x en de niveaus daaronder goed maken. Deze procedure leidt tot het aantal opgaven dat een kandidaat goed moet maken op de examens die in de standaardsetting zijn opgenomen (zie paragraaf 2.2). Deze score kan vervolgens worden omgezet in de 500-score.
2. Het 80%-criterium. Hierbij wordt de grensscore voor het bereiken van een ERK-niveau als volgt bepaald: voor het behalen van ERK-niveau x moet een kandidaat alle opgaven onder niveau x en 80% van niveau x goed maken. Deze procedure leidt tot het aantal opgaven dat een kandidaat goed moet maken op de examens die in de standaardsetting zijn opgenomen (zie paragraaf 2.2). Deze score kan vervolgens worden omgezet in de 500-score.

Geen van beide berekeningscriteria is de enig juiste. Ze kunnen worden gezien als een wat mildere en wat strengere wijze van het bepalen van grensscores en verschaffen zo een interval waarin een te bepalen grensscore zich zou kunnen bevinden. In analogie met het koppelingsonderzoek van Spreken en Schrijven aan het ERK (Kuijper e.a. 2004, 2005) ligt het voor de hand dat het bij de Staatsexamens NT2 belangrijker is false negatives te voorkomen dan false positives. Daarom lijkt het hanteren van het 80%-criterium meer voor de hand te liggen dan het 100% criterium.

In deze paragraaf presenteren we de grensscores die uit deze procedures voortvloeien, vergelijken we deze met de huidige cesuur en geven we tevens aan wat de consequenties van de gevonden grensscores zouden zijn voor de zak-slaagpercentages bij de leesexamens.

Tabel 3.1-3 en tabel 3.1-4 vormen de basis voor het bepalen van grensscores voor de ERK-niveaus. In deze tabellen staat voor elke beoordelaar hoeveel opgaven hij in elk ERK-niveau heeft ingedeeld. Sommatie van deze oordelen geeft voor elke beoordelaar aan hoeveel opgaven van de beoordeelde examens een kandidaat goed moet hebben om een bepaald niveau te behalen.

Tabel 3.1-6 en 3.1-7 geven voor elke beoordelaar weer hoeveel opgaven een kandidaat volgens deze beoordelaar goed moet maken om een bepaald niveau te bereiken volgens respectievelijk de 80% en het 100%-criterium. In de laatste kolom staat het gemiddelde over alle beoordelaars.

*Tabel 3.1-6 Aantal goed gemaakte opgaven nodig om ERK-niveaus te halen. Lezen I. Per beoordelaar. Afgerond naar gehele itemscores*

Beoordelaars												
	1	2	3	4	5	6	7	8	9	10	11	Gem.
80%-criterium												
A2	2	2	1	2	6	2	1	2	1	6	1	2
B1	25	24	19	25	24	16	19	20	14	27	24	22
B2	38	38	37	38	37	34	33	36	34	37	38	36
100%-criterium												

A2	3	2	1	2	7	2	1	2	1	8	1	<b>3</b>
B1	30	30	24	31	28	20	23	24	17	32	30	<b>26</b>
B2	40	40	40	40	39	37	36	39	38	38	40	<b>39</b>

Volgens het gemiddelde oordeel zou een kandidaat volgens het 80%-criterium 22 en volgens het 100%-criterium 26 van de 40 opgaven goed moeten maken om het B1-niveau te halen. Het blijkt verder dat de Lezen I duidelijk op B1 niveau functioneert en dat daardoor de enige relevante cesuur voor dit examen de A2/B1 cesuur is. De toets is duidelijk geen meetinstrument voor het A2 niveau, omdat men reeds bij 2, respectievelijk 3 opgaven het A2 niveau bereikt zou hebben. Evenmin lijkt Lezen I geschikt te zijn om te meten of iemand B2 beheerst. Omdat men daarvoor respectievelijk 36 of 39 opgaven goed zou moeten hebben, is Lezen I te gemakkelijk voor B2-kandidaten.

Er is variatie is tussen de beoordelaars. Volgens het 80%-criterium is het verschil tussen de mildste en de strengste beoordelaar enerzijds en het gemiddelde anderzijds respectievelijk -8 en 5 punten. De scores van 9 van de 11 beoordelaars liggen in een interval van -3 en 5 rondom het gemiddelde. Twee beoordelaars wijken sterker af: beoordelaar 9 met een grensscore van 14 opgaven en beoordelaar 6 met een grensscore van 16 om B1 te behalen. Volgens het 100%- criterium is het verschil tussen de mildste en de strengste beoordelaar en het gemiddelde respectievelijk -9 en 6 scorepunten. Wanneer we de twee mildste beoordelaars buiten beschouwing zouden liggen de scores in een interval van -3 en 6 rondom het gemiddelde. De beoordelaars blijken niet allemaal hetzelfde beeld te hebben van wat een B1-kandidaat zou moeten kunnen. Voor een interpretatie van de verschillen wordt in bijlage 5 de spreiding van de beoordelaars vergeleken met de theoretische scoreverdeling van kandidaten met de vaardigheid van de gevonden grensscore.

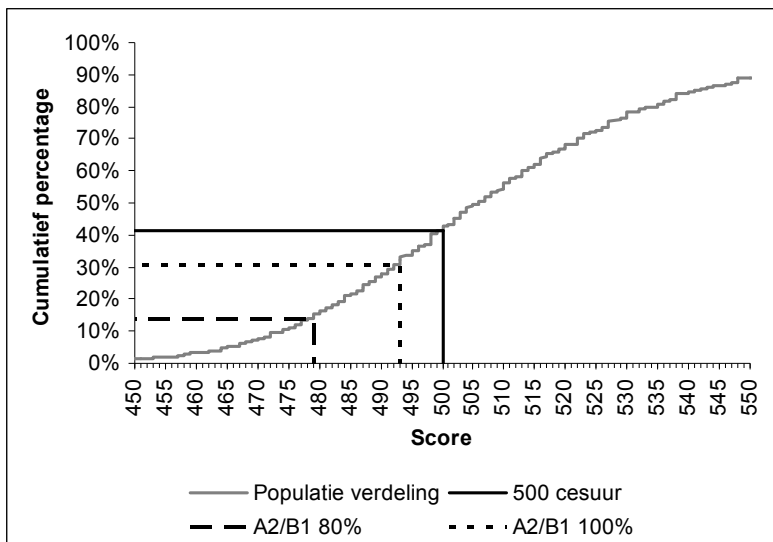
*Tabel 3.1.7 Aantal goed gemaakte opgaven nodig om ERK-niveaus te halen. Lezen II. Per beoordelaar. Afgerond naar gehele itemscores*

Beoordelaars												
	1	2	3	4	5	6	7	8	9	10	11	Gem.
80%-criterium												
A2	0	0	0	0	0	0	0	0	0	0	0	<b>0</b>
B1	8	8	2	2	6	3	4	8	2	11	7	<b>6</b>
B2	32	32	24	29	31	23	27	28	31	30	31	<b>29</b>
C1	40	39	38	39	39	38	38	38	40	39	39	<b>39</b>
100%-criterium												
A2	0	0	0	0	0	0	0	0	0	0	0	<b>0</b>
B1	10	10	2	2	8	4	5	10	2	14	9	<b>7</b>
B2	38	37	30	36	37	28	32	32	38	34	37	<b>35</b>
C1	40	40	40	40	40	40	40	40	40	40	40	<b>40</b>

Volgens het gemiddelde oordeel zou een kandidaat volgens het 80%-criterium 29 en volgens het 100%-criterium 35 van de 40 opgaven goed moeten maken om het B1-niveau te halen. Zoals Lezen I met name geschikt is als meetinstrument voor het B2-niveau, is Lezen II geschikt voor het B2 niveau. Als meetinstrument voor A2 en C1 is het examen niet geschikt. Ook bij Lezen II is er verschil tussen de beoordelaars. Bij de 80%-norm bedraagt het verschil tussen de mildste en de strengste beoordelaar enerzijds en het gemiddelde anderzijds respectievelijk-6 en 3 punten. Bij de 100%-norm is dit -7 respectievelijk 3 punten. Ook hier vallen verschillen tussen de beoordelaars binnen aanvaardbare grenzen, zoals in bijlage 5 wordt toegelicht.

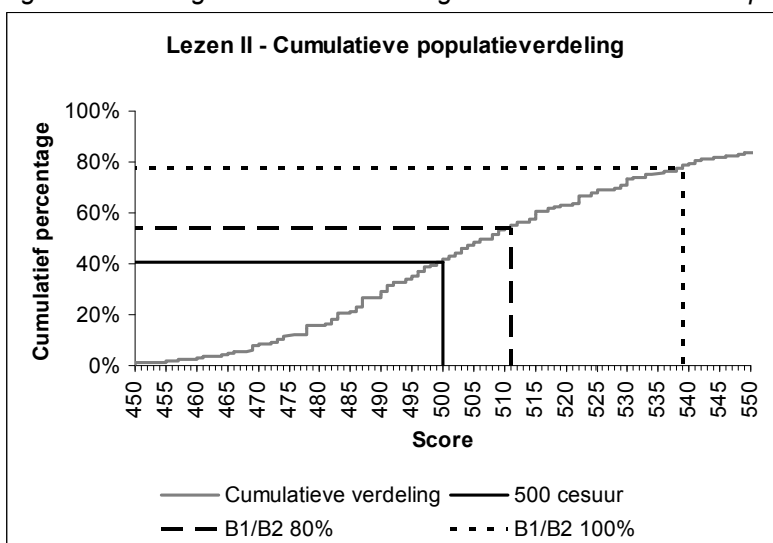
In figuur 3.1 5 en 3.1.6 worden de gemiddelde 80%- en de 100%-grensscores, getransformeerd naar de 500-schaal weergegeven naast de huidige 500-cesuurscore voor respectievelijk B1 bij Lezen I en B2 bij Lezen II. De grafiek geeft tevens de cumulatieve populatieverdeling van 2009 weer. Daardoor zijn voor de beide grensscores en de cesuur de zak- en slaagpercentages af te lezen als we deze op de 2009-populatie zouden toepassen.

*Figuur 3.1-5 B1-grensscores en huidige cesuur en cumulatieve populatieverdeling Lezen I*



Uit figuur 3.1-5 blijkt dat de huidige cesuur van Lezen I boven zowel de 80\*- als 100%-grensscores A2/B1 ligt. Van kandidaten die slagen voor Lezen I kan worden gezegd dat zij het ERK-niveau B1 beheersen. Wanneer de 80% norm van 478/479 als zak-slaaggrens gehanteerd zou worden zou 86% van de kandidaten slagen. Bij de 492/493-grensscore volgens de 100% norm als zak-slaaggrens zou 69% van de populatie slagen.

Figuur 3.1-6 B2-grensscores en huidige cesuur en cumulatieve populatieverdeling. Lezen II



Figuur 3.1-6 laat zien dat beide B2-grensscores boven de huidige cesuur liggen. Volgens het 80%-criterium is de grensscore voor het behalen van B2 510/ 511 en volgens het 100%-criterium zou de grensscore 538/539 moeten zijn. Bij een grensscore van 510/511 zou het percentage geslaagde kandidaten afnemen tot 46%, bij een grensscore van 538/539 tot 33%.

### 3.1.4 Samenvatting en conclusies Lezen

In dit onderzoek zijn de leesexamens programma I en II van het Staatsexamen NT2 gerelateerd aan het ERK. Hiervoor werden opgaven van de leesexamens Programma I en II voorgelegd aan een panel van beoordelaars. Bij dit onderzoek is gebruik gemaakt van de mandjesmethode, een van de standaardsettingsmethodes die beschreven worden in de Manual for relating language examinations to the Common European framework (Council of Europe, 2009).

Het panel van 11 beoordelaars kreeg 40 opgaven voor Lezen I, 40 opgaven voor Lezen II en 20 opgaven uit de Profieltoets NT2 in een random volgorde voorgelegd. De opgaven voor zowel Lezen I



als II vormden een geheel examen. De beoordelaars moesten voor elke opgave de volgende vraag beantwoorden:

*'Geef bij elke opgave aan welk niveau (A1, A2, B1, B2, C1, C2) minimaal vereist is om de opgave goed te kunnen maken.'*

Deze beoordelingstaak werd in twee zittingen uitgevoerd. Tijdens de eerste zitting moesten de beoordelaars onafhankelijk van elkaar deze vraag voor alle 100 opgaven beantwoorden. In een tweede zitting werden de opgaven waarover de minste overeenstemming bestond nogmaals aan de beoordelaars voorgelegd. De beoordelaars konden voor deze opgaven een toelichting op hun eerdere keuze geven. Elke beoordelaar kon na deze bespreking aangeven of hij zijn niveautoekenning al dan niet wilde veranderen. Deze tweede bijeenkomst leidde tot een hogere mate van beoordelaarsovereenstemming.

De mate van overeenstemming tussen de beoordelaars over de niveautoekenning was bevredigend. Voor 93% van de opgaven gold dat minimaal 8 van de 11 beoordelaars het eens waren over het niveau dat nodig was om een opgave goed te maken. De beoordelaarsbetrouwbaarheid en –overeenstemming waren respectievelijk .95 en .94.

Uit de niveautoekenning aan de opgaven bleek dat er duidelijke niveauverschillen waren tussen Lezen I en Lezen II. De gemiddelde niveautoekenning van opgaven van Lezen I lag iets boven B1, die van de opgaven van Lezen II op B2. De resultaten komen in grote mate overeen met de eerdere bevindingen in de specificatiefase van het koppelingsonderzoek (Kuijper, 2009).

Op basis van de gemiddelde niveautoekenningen werden vervolgens de grensscores op de 500-schalen berekend voor B1 op Luisteren I en voor B2 op Luisteren II. Daarvoor werden twee criteria gehanteerd: het 80%-criterium en het 100%-criterium. Bij het 80%-criterium moet een kandidaat om een bepaald niveau te behalen alle opgaven onder dat niveau en 80% van de opgaven van het betreffende niveau goed maken. Iemand haalt dus B1 wanneer hij alle opgaven van A1 en A2 en 80% van de B1 opgaven goed heeft gemaakt. Volgens het 100%-criterium moet een kandidaat alle opgave onder en op het betreffende niveau goed maken. Het laatste criterium leidt vanzelfsprekend tot hogere grensscores. Tabel 3.1-8 geeft een overzicht van de huidige 500-cesuur en de 80%- en 100%-cesuur voor B1 bij Lezen I en B2 bij Lezen II. Kolom 3 geeft de daarbij behorende zak-slaagpercentages weer.

*Tabel 3.1-8 Cesaurscores, B1- respectievelijk B2-grensscores en slaagpercentages, berekend voor de populatie 2009, Lezen I en II*

	Schaalscore	Ruwe toetsscore	Percentage geslaagd
Lezen I			
huidige cesuur	499/500	27/28	59%
B1 grensscore 80%	478/479	21/22	86%
B1 grensscore 100%	492/493	25/26	69%
Lezen II			
huidige cesuur	499/500	26/27	59%
B2 grensscore 80%	510/511	28/29	46%
B2 grensscore 100%	538/539	33/34	22%

Kandidaten die slagen voor Lezen I bevinden zich boven de 80% en 100% A2/B1-grensscores. Voor Programma II ligt de huidige cesuur onder de 80% grensscore van B1/B2 die op 510/511 ligt. Bij 511 hoort een score van 29 op het examen in de standaardsetting. De huidige 500-cesuur correspondeert met een ruwe cesuurscore van 26/27 op het examen dat in de standaardbepaling is gebruikt. Uit de berekeningen van de grensscores voor Lezen I en II bleek verder dat Lezen I en Lezen II niet geschikt zijn om andere niveaus te meten dan respectievelijk B1 en B2. Er was uiteraard variatie tussen de beoordelaars rondom de gemiddelde grensscores. De verdeling van de grensscores op basis van de beoordelaars viel binnen de theoretische scoreverdeling van een kandidaat met de vaardigheid op de gemiddelde grensscores.

Op grond van deze tweede fase koppelingsonderzoek en het eerder uitgevoerde specificatieonderzoek (Kuijper, 2009) kan geconcludeerd worden dat Staatsexamens Lezen Programma I zich op B1 bevindt en Lezen Programma II op B2.

Deze conclusies hoeven niet per se te leiden tot een wijziging in de cesuren. De Staatsexamens kennen immers een functionele cesuur die gebaseerd is op het doel van de examens: het toetsen van een startkwalificatie voor de opleidingen en beroepen waarvoor beide programma's bedoeld zijn. De Staatsexamens NT2 zijn niet in eerste instantie bedoeld om een bepaald ERK-niveau te meten.

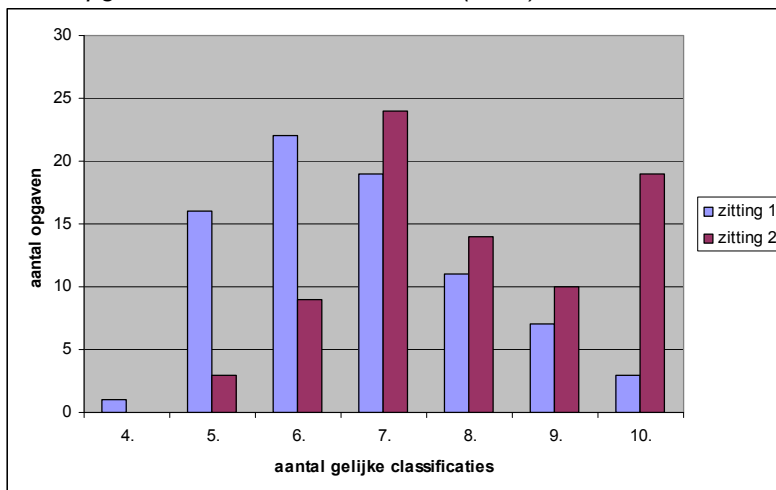
## 3.2 Luisteren

### 3.2.1 Beoordelaarsovereenstemming en beoordelaarsbetrouwbaarheid

Een voldoende mate van beoordelaarsovereenstemming en beoordelaarsbetrouwbaarheid is een voorwaarde om later uitspraken te kunnen doen over het niveau van de opgaven in een examen en voor het vaststellen van grensscores.

Inzicht in de beoordelaarsovereenstemming kan worden verkregen door na te gaan hoeveel beoordelaars een opgave in hetzelfde 'mandje' hebben geplaatst, dat wil zeggen hetzelfde minimumniveau hebben aangegeven dat nodig is om een opgave goed te maken. De beoordelaars konden elke opgave plaatsen in één van de 6 ERK-niveaus (A1 t/m C2). Als de beoordelaars een grote mate van onenigheid zouden hebben over het niveau van een opgave, zouden de oordelen bijvoorbeeld gelijk verdeeld zijn over deze niveaus. In dat geval zouden vier niveaus door 2 beoordelaars en het vijfde en zesde niveau door één beoordelaar zijn toegekend. Als de beoordelaars het allemaal eens zijn over het niveau dat nodig is om een opgave goed te maken, zouden 10 beoordelaars hetzelfde niveau aan een opgave toekennen. Figuur 3.2-1 laat zien in welke mate de beoordelaars het na zitting 1 en zitting 2 eens zijn over de niveautoekenning aan opgaven van Programma I en II. De x-as geeft aan hoeveel beoordelaars het een zijn over het niveau van een opgave. Op de y-as kan worden afgelezen om hoeveel opgaven het daarbij gaat. Het gaat hierbij om 79 opgaven, omdat tijdens de bespreking van de beoordelingen in zitting 2 het panel één van de 40 STEX-I opgaven niet kon beoordelen. Er was te grote onenigheid over het juiste antwoord van deze opgave.

*Figuur 3.2-1 Mate van overeenstemming van panelleden (n=10) over het vereiste niveau van de luisteropgaven in Staatsexamen I en II (k=79)*



Figuur 3.2-1 laat allereerst zien dat de tweede zitting tot meer overeenstemming heeft geleid. Na de tweede zitting waren er 3 opgaven die van 5 beoordelaars hetzelfde niveau kregen toegekend. Over 9 opgaven waren 6 beoordelaars het eens. Over de overige 67 van de 79 opgaven waren 7 of meer beoordelaars het eens over de niveautoekenning. Wanneer beoordelaars het niet volledig eens waren lagen de niveau-inschattingen altijd in twee naast elkaar liggende niveaus.

Tabel 3.2-1 geeft een overzicht van de beoordelaarsbetrouwbaarheid, beoordelaarsovereenstemming en resultaten van een variantieanalyse. De beoordelaarsbetrouwbaarheid geeft aan in hoeverre de rangorde van de oordelen over de beoordelaars consistent is, terwijl de beoordelaarsovereenstemming aangeeft in hoeverre opgaven op dezelfde wijze worden geclassificeerd. Beide maten lopen van 0 bij minimale overeenstemming/betrouwbaarheid tot 1 bij een maximale overeenstemming/betrouwbaarheid. De variantieanalyse geeft aan in welke mate de variatie in de oordelen is toe te schrijven aan de opgaven, de beoordelaars aan interactie- en erroreffecten (residu).

Tabel 3.2.-1 Variantiecomponenten, beoordelaarsovereenstemming en beoordelaarsbetrouwbaarheid Luisteren

Beoordelaarsbetrouwbaarheid	.92
Beoordelaarsovereenstemming bij 10 beoordelaars (Rho2)	.91
Geschatte bijdrage van variantiecomponenten:	
Opgaven	50%
Beoordelaars	8%
Residu (interactie, error)	43%

Zowel de beoordelaarsbetrouwbaarheid als de beoordelaarsovereenstemming zijn hoog. De variantie-analyse wijst uit dat met 50% de meeste variantie verklaard wordt door de opgaven, terwijl de factor beoordelaar slechts 8% van de variantie verklaart. Dit is bevredigend, hoewel de 43% residuvariantie er op wijst dat er geen volledige overeenstemming is.

In tabel 3.2-2 wordt een overzicht gegeven van de absolute verschillen tussen de niveau-indeling door de 10 beoordelaars. Het gemiddelde (Gem.) geeft per beoordelaar de gemiddelde score over alle 79 STEX-opgaven. Delta is het absolute verschil tussen de gemiddelde score van elke beoordelaar en het totaal gemiddelde (Gem-totaal). Bij deze berekening zijn de waarden 1-6 toegekend aan de niveaus A1-C2. Een groot verschil kan sterk afwijkende beoordelaars detecteren. Tabel 3.2-2 geeft verder voor elke beoordelaar de Rbt-waarde. Naar analogie met de Rit is de Rbt de correlatie tussen de beoordelaarscores en de beoordelaargemiddelden.

Tabel 3.2-2 Gemiddelde scores en Rbt's van de beoordelaars over alle STEX-opgaven Luisteren.

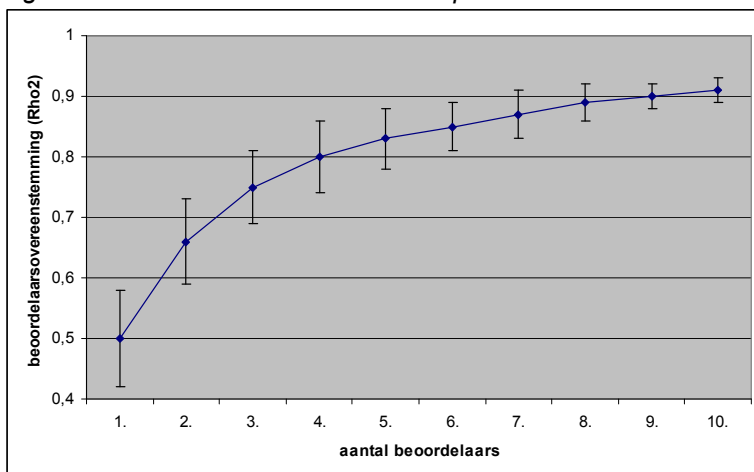
	Beoordelaars										Gem-totaal
	1	2	3	4	5	6	7	8	9	10	
Gem.*	3,34	3,33	3,66	3,41	3,66	3,63	3,73	3,68	3,37	3,62	3,54
Delta*	0,20	0,21	0,12	0,14	0,12	0,09	0,19	0,14	0,18	0,08	0,14
Rbt*	0,80	0,78	0,83	0,81	0,64	0,81	0,76	0,79	0,81	0,61	0,76

\* Afgerond op twee decimalen

Het blijkt dat de gemiddelde niveau-indeling voor alle beoordelaars ligt tussen B1 (3) en B2 (4). De gemiddelde delta over alle beoordelaars bedraagt 0,14 van een niveau, terwijl de grootste delta 0,21 bedraagt. De verschillen tussen beoordelaars zijn dus gering. Ook de Rbt geeft hoge waarden aan.

Het panel dat voor de standaardsetting werd gebruikt bestond uit 10 beoordelaars. In figuur 3.2-2 staat weergegeven hoe de beoordelaarsovereenstemming zich verhoudt tot het aantal panelleden. In deze figuur wordt de geschatte beoordelaarsovereenstemming per aantal beoordelaars met een ander panel van dezelfde omvang en samenstelling en de 90%-betrouwbaarheidsintervallen van deze schattingen weergegeven.

Figuur 3.2-2 Relatie tussen het aantal panelleden en de beoordelaarsovereenstemming (Rho2)



Figuur 3.2-2 laat zien dat de geschatte beoordelaarsovereenstemming vanaf 10 beoordelaars nauwelijks meer zal toenemen evenmin als het betrouwbaarheidsinterval van de schattingen. Het aantal beoordelaars dat we bij deze standaardbepaling hebben gebruikt blijkt op grond hiervan voldoende te zijn.

Bovenstaande resultaten wijzen erop dat de beoordelaarsovereenstemming en –betrouwbaarheid voldoende hoog zijn om in de volgende paragrafen een beschrijving te geven van het niveau waarop de opgaven in de Staatsexamens NT2 zich bevinden en een advies uit te brengen over de grensscores op beide examens voor de ERK-niveaus die voor beide programma's relevant zijn.

### 3.2.2 ERK-niveaus van de opgaven in Programma I en II

Tijdens de standaardbepaling hebben de beoordelaars niveaus toegekend aan de opgaven van Staatsexamen Programma I en II.

Tabel 3.2-3 en 3.2-4 geven het resultaat van deze niveautoekenningen weer. In deze tabellen zien we voor elke beoordelaar het aantal opgaven dat op een bepaald niveau is ingedeeld. Voor Programma I is de beoordeling van één opgave niet in de analyses opgenomen, omdat de beoordelaars ten gevolge van onenigheid over het juiste antwoord het niveau van deze opgaven niet konden inschatten.

Tabel 3.2-3 Niveautoekenning aan de opgaven van Staatsexamen Programma I Luisteren (k=39)

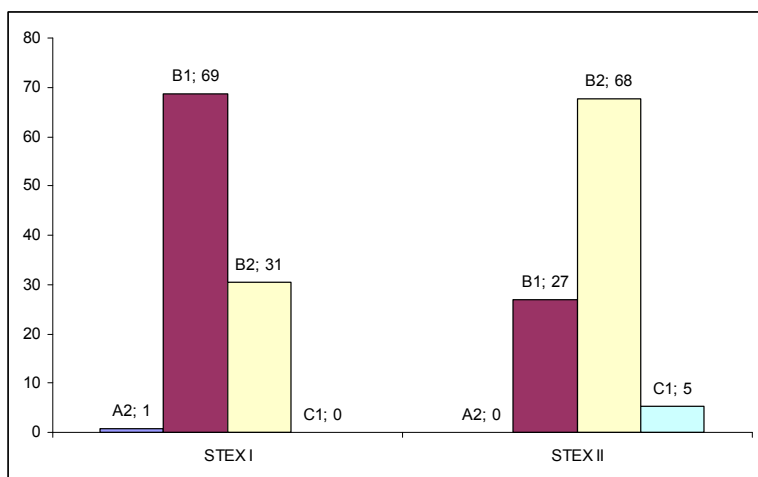
	Beoordelaars										Gem.
	1	2	3	4	5	6	7	8	9	10	
A1	0	0	0	0	0	0	0	0	0	0	<b>0</b>
A2	1	0	0	0	1	0	0	0	1	0	<b>0,3</b>
B1	33	36	27	34	20	23	20	22	33	20	<b>26,8</b>
B2	5	3	12	5	18	16	19	17	5	19	<b>11,9</b>
C1	0	0	0	0	0	0	0	0	0	0	<b>0</b>
C2	0	0	0	0	0	0	0	0	0	0	<b>0</b>

Tabel 3.2-4 Niveautoekenning aan de opgaven van Staatsexamen Programma II Luisteren (k=40)

	Beoordelaars										Gem.
	1	2	3	4	5	6	7	8	9	10	
A1	0	0	0	0	0	0	0	0	0	0	<b>0</b>
A2	0	0	0	0	0	0	0	0	0	0	<b>0</b>
B1	17	18	5	14	7	10	4	7	16	10	<b>10,8</b>
B2	23	21	30	25	31	26	33	29	23	30	<b>27,1</b>
C1	0	1	5	1	2	4	3	4	1	0	<b>2,1</b>
C2	0	0	0	0	0	0	0	0	0	0	<b>0</b>

Uit tabellen 3.2-3 en 3.2-4 blijkt dat er een duidelijk verschil is tussen de niveautoekenningen aan opgaven van Programma I en II. Voor het merendeel van de Programma I opgaven heeft men volgens het beoordelaarspanel minimaal B1 nodig om deze opgaven goed te kunnen maken. Voor programma II heeft men voor het merendeel van de opgaven een vaardigheid minimaal op B2-niveau nodig. Aan elke opgave van programma I en II is door 10 beoordelaars een niveau toegekend. Figuur 3.2-3 geeft een overzicht van de percentages over de verschillende niveaus van deze beoordelingen.

Figuur 3.2-3 verdeling van de beoordelingen van de leesopgaven over de ERK niveaus voor STEX I en STEX II. In percentages. (totaal 80 opgaven beoordeeld door 10 beoordelaars)



Figuur 3.2-3 geeft het duidelijke niveauverschil tussen beide examens weer. De opgaven van Programma I worden in 69% van de gevallen als B1 ingeschat, terwijl de opgaven van Programma II in 68% van de beoordelingen B2 krijgen toegekend.

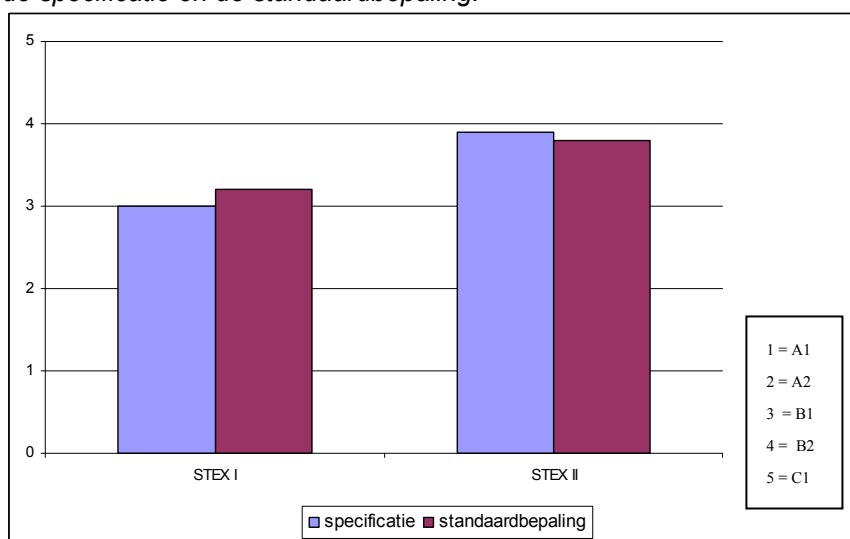
Wanneer we aan de niveaus A1 t/m C2 een waarde toekennen van 1-6 kunnen we per beoordelaar een gemiddelde niveauscore voor Programma I en Programma II berekenen. Tabel 3.2-5 geeft hier een overzicht van voor beide programma's.

Tabel 3.2-5 Gemiddeld vereist minimumniveau voor Programma I en II Lezen

	Beoordelaars										Gem.
	1	2	3	4	5	6	7	8	9	10	
Programma I	3,0	3,0	3,2	3,1	3,4	3,3	3,4	3,4	3,0	3,4	<b>3,2</b>
Programma II	3,6	3,6	4,0	3,7	3,9	3,9	4,0	3,9	3,6	3,8	<b>3,8</b>

De opgaven voor Programma I liggen gemiddeld iets boven B1. De opgaven van Programma II liggen gemiddeld iets onder B2. Ook in de eerder uitgevoerde specificatiefase (Kuijper, 2009) (Kuijper, 2009) is een niveau-indicatie van de opgaven uitgevoerd. Figuur 3.2-4 geeft die de gemiddelde niveau-inschatting weer voor Programma I en II in de specificatiefase en de standaardbepaling.

Figuur 3.2-4 Gemiddelde niveautoekenning van de leesopgaven Programma I en II Luisteren volgens de specificatie en de standaardbepaling.



Zowel volgens de specificatie als de standaardbepaling ligt Luisteren I in het B1-niveaubereik en Luisteren II tegen het B2-niveaubereik aan, waarbij in de standaardbepaling voor Programma I het

niveau iets hoger wordt ingeschat en bij Programma II iets lager dan in de specificatiefase. De verschillen tussen de resultaten van beide methodieken zijn echter gering.

### 3.2.3 De bepaling van grensscores bij relevante ERK-niveaus Luisteren

In de vorige paragraaf is beschreven dat de Luisterexamens van Staatsexamen I en II zich bevinden op respectievelijk het niveaubereik B1 en B2 van het ERK.

Op grond van de resultaten van de standaardsetting kunnen we een grensscore vaststellen die op een examen nodig is om een voor dat examen relevant ERK-niveau te behalen. We kunnen deze grensscores vervolgens vergelijken met de cesuren die nu bij de Staatsexamens worden gehanteerd. Het is belangrijk te benadrukken dat het gezien de functie van de Staatsexamens NT2 niet noodzakelijk is om de huidige cesuren te veranderen op grond van de resultaten die in deze paragraaf besproken worden. De Staatsexamens kennen immers een functionele cesuur die gebaseerd is op het doel van de examens: het toetsen van een startkwalificatie voor de opleidingen en beroepen waarvoor beide programma's bedoeld zijn. De Staatsexamens NT2 zijn niet in eerste instantie bedoeld om een bepaald ERK-niveau te meten. De resultaten die in deze paragraaf besproken maken het mogelijk om scores te relateren aan een ERK niveau, maar hoeven er niet per se toe te leiden de huidige cesuren te veranderen.

Bij het bepalen van een grensscore voor een bepaald ERK-niveau hanteren we twee criteria.

1. Het 100%-criterium. Hierbij wordt de grensscore voor het bereiken van een ERK-niveau als volgt bepaald: voor het behalen van ERK-niveau x moet een kandidaat alle opgaven van niveau x en de niveaus daaronder goed maken. Deze procedure leidt tot het aantal opgaven dat een kandidaat goed moet maken op de examens die in de standaardsetting zijn opgenomen (zie paragraaf 2.2). Deze score kan vervolgens worden omgezet in de 500-score
2. Het 80%-criterium. Hierbij wordt de grensscore voor het bereiken van een ERK-niveau als volgt bepaald: voor het behalen van ERK-niveau x moet een kandidaat alle opgaven onder niveau x en 80% van niveau x goed maken. Deze procedure leidt tot het aantal opgaven dat een kandidaat goed moet maken op de examens die in de standaardsetting zijn opgenomen (zie paragraaf 2.2). Deze score kan vervolgens worden omgezet in de 500-score.

Geen van beide berekeningscriteria is het enig juiste. Ze kunnen worden gezien als een wat mildere en wat strengere wijze van het bepalen van grensscores en verschaffen zo een interval waarin een te bepalen grensscore zich zou kunnen bevinden. In analogie met het koppelingsonderzoek van Spreken en Schrijven aan het ERK (Kuijper e.a. 2004, 2005) ligt het voor de hand dat het bij de Staatsexamens NT2 belangrijker is false negatives te voorkomen dan false positives. Daarom lijkt het hanteren van het 80%-criterium meer voor de hand te liggen dan het 100% criterium.

In deze paragraaf presenteren we de grensscores die uit deze procedures voortvloeien, vergelijken we deze met de huidige cesuur en geven we tevens aan wat de consequenties van de gevonden grensscores zouden zijn voor de zak-slaagpercentages bij de luisterexamens.

Tabel 3.2-3 en tabel 3.2-4 vormen de basis voor het bepalen van grensscores voor de ERK-niveaus Luisteren. In deze tabellen staat voor elke beoordelaar hoeveel opgaven hij in elk ERK-niveau heeft ingedeeld. Sommatie van deze oordelen geeft voor elke beoordelaar aan hoeveel opgaven van de beoordeelde examens een kandidaat goed moet hebben om een bepaald niveau te behalen.

Tabel 3.2-6 en 3.2-7 geven voor elke beoordelaar weer hoeveel opgaven een kandidaat volgens deze beoordelaar goed moet maken om een bepaald niveau te bereiken volgens respectievelijk het 80%- en het 100%-criterium. In de laatste kolom staat het gemiddelde over alle beoordelaars.

*Tabel 3.2.6 Aantal goed gemaakte opgaven nodig om ERK-niveaus te halen. Luisteren I. Per beoordelaar. Afgerond naar gehele itemscores*

Beoordelaars											
	1	2	3	4	5	6	7	8	9	10	Gem.
80%-criterium											
A2	1	0	0	0	1	0	0	0	1	0	<b>0</b>
B1	27	29	22	27	17	18	16	18	27	16	<b>22</b>

B2	38	38	37	38	35	36	35	36	38	35	<b>37</b>
100%-criterium											
A2	1	0	0	0	1	0	0	0	1	0	<b>0</b>
B1	34	36	27	34	21	23	20	22	34	20	<b>27</b>
B2	39	39	39	39	39	39	39	39	39	39	<b>39</b>

Volgens het gemiddelde oordeel zou een kandidaat volgens het 80% criterium gemiddeld 22 en volgens het 100% criterium 27 van de 39 opgaven goed moeten maken om het B1-niveau te halen. Het blijkt verder dat de Luisteren I op B1 niveau functioneert. De toets is zeer duidelijk geen meetinstrument voor het A2 niveau, omdat men volgens beide criteria met nul opgaven goed het A2 niveau bereikt zou hebben. Evenmin lijkt Luisteren I erg geschikt te zijn om te meten of iemand B2 beheerst. Omdat men daarvoor respectievelijk 37 of 39 opgaven goed zou moeten hebben, is Luisteren I te gemakkelijk voor B2-kandidaten.

Er is variatie tussen de beoordelaars. Volgens de 80% norm is het verschil tussen de mildste en de strengste beoordelaar enerzijds en het gemiddelde anderzijds respectievelijk -6 en 7 punten. Bij het 80%-criterium kunnen we twee groepen onderscheiden. De eerste groep bestaat uit 5 beoordelaars waarvan de B1-grensscore tussen 16 en 18 ligt en een groep van vier beoordelaars met grensscores van 27 tot 29. Één beoordelaar ligt daartussen met een score van 22, wat ook het gemiddelde is over alle beoordelaars heen. Volgens het 100% criterium is het verschil tussen de mildste en de strengste beoordelaar en het gemiddelde respectievelijk -7 en 9 scorepunten. Weer zijn er twee groepen beoordelaars te herkennen. De eerste groep van 5 beoordelaars geeft grensscores van 20 tot 23, de tweede groep, bestaande uit vier beoordelaars geeft grensscores van 34 tot 36. De tiende beoordelaar komt tot een grensscore van 27, weer gelijk aan het gemiddelde. De beoordelaars blijken niet hetzelfde beeld te hebben van wat een B1-kandidaat zou moeten kunnen. Voor een interpretatie van de verschillen wordt in bijlage 6 de spreiding van de beoordelaars vergeleken met de theoretische scoreverdeling van kandidaten met de vaardigheid van de gevonden grensscore. Bij de 80%-grensscore blijkt de variabiliteit binnen deze theoretische verdeling te liggen, terwijl voor de 100% grensscore één beoordelaar buiten deze scoreverdeling ligt.

*Tabel 3.2-7 Aantal goed gemaakte opgaven nodig om ERK-niveaus te halen. Luisteren II. Per beoordelaar. Afgerond naar gehele itemscores*

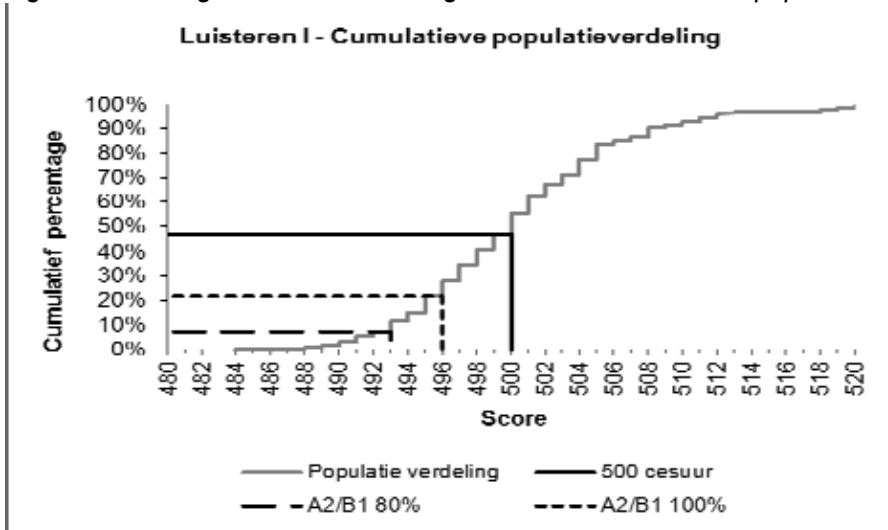
Beoordelaars											
	1	2	3	4	5	6	7	8	9	10	Gem.
80%-criterium											
A2	0	0	0	0	0	0	0	0	0	0	<b>0</b>
B1	14	14	4	11	6	8	3	6	13	8	<b>9</b>
B2	35	35	29	34	32	31	30	30	34	34	<b>32</b>
C1	40	40	39	40	40	39	39	39	40	40	<b>40</b>
100%-criterium											
A2	0	0	0	0	0	0	0	0	0	0	<b>0</b>
B1	17	18	5	14	7	10	4	7	16	10	<b>11</b>
B2	40	39	35	39	38	36	37	36	39	40	<b>38</b>
C1	40	40	40	40	40	40	40	40	40	40	<b>40</b>

Volgens het gemiddelde oordeel zou een kandidaat volgens het 80%-criterium 32 en volgens het 100%-criterium 38 van de 40 opgaven goed moeten maken om het B2-niveau te halen. Zoals Luisteren I met name geschikt is als meetinstrument voor het B1-niveau, is Luisteren II met name geschikt voor het B2 niveau, hoewel het 100%-criterium hiervoor wel een erg hoge grensscore zou vereisen. Als meetinstrument voor B1 en C1 is het examen niet geschikt. Voor B1 is het examen te moeilijk, voor C1 te gemakkelijk. Ook bij Luisteren II is er verschil tussen de beoordelaars, maar dit verschil is duidelijk kleiner dan bij Luisteren I. Volgens de 80% norm is het verschil tussen de mildste en de strengste beoordelaar enerzijds en het gemiddelde anderzijds respectievelijk -3 en 3 punten. Bij de 100%-norm is dit respectievelijk -3 en 2, waarbij aangetekend moet worden dat een grensscore van meer dan 2 boven de 38 niet kan worden gevonden, gezien de maximumscore van 40. Voor een interpretatie van de verschillen wordt in bijlage 6 de spreiding van de beoordelaars vergeleken met de theoretische scoreverdeling van kandidaten met de vaardigheid van de gevonden grensscores.



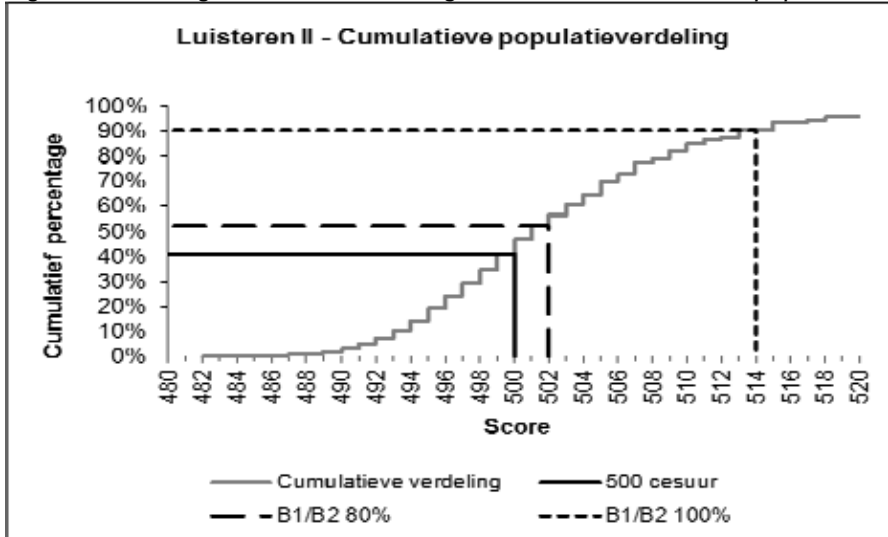
In figuur 3.2.5 en 3.2.6 worden de gemiddelde 80%- en de 100%-grensscores, getransformeerd naar de 500-schaal weergegeven naast de huidige 500-cesuurscore voor respectievelijk B1 bij Luisteren I en B2 bij Luisteren II. De grafiek geeft tevens de cumulatieve populatieverdeling van 2009 weer. Daardoor zijn voor de beide grensscores en de cesuur de zak- en slaagpercentages af te lezen als we deze op de 2009-populatie zouden toepassen.

Figuur 3.2-5 B1-grensscores en huidige cesuur en cumulatieve populatieverdeling Luisteren I



Uit figuur 3.2-5 blijkt dat de huidige cesuur van Luisteren I boven zowel de 80%- als de 100%-grensscores A2/B1 ligt. Deze grensscores bedragen respectievelijk 492/493 en 495/496. Van kandidaten die slagen voor Luisteren I kan worden gezegd dat zij het ERK-niveau B1 beheersen. Bij de 80%-grensscore bedraagt het aantal geslaagde kandidaten 93%, bij de 100%-grensscore zou dit 78% zijn.

Figuur 3.1-6 B2-grensscores en huidige cesuur en cumulatieve populatieverdeling. Luisteren II



Figuur 3.2-6 laat zien dat beide B2-grensscores boven de huidige cesuur liggen. Volgens het 80%-criterium is voor het behalen van B2 de grensscore 501/502 en volgens het 100%-criterium zou de grensscore 513/514 moeten zijn. Bij een grensscore van 502 zou het percentage geslaagde kandidaten afnemen tot 48%, bij een grensscore van 514 tot 10%.

### 3.2.4 Samenvatting en conclusies Luisteren

In dit onderzoek zijn de Luisterexamens programma I en II van het Staatsexamen NT2 gerelateerd aan het ERK. Hiervoor werden opgaven van de luisterexamens Programma I en II voorgelegd aan een panel van 10 beoordelaars. Bij dit onderzoek is gebruik gemaakt van de mandjesmethode, een van de standaardsettingsmethodes die beschreven worden in de Manual for relating language examinations to the Common European framework (Council of Europe, 2009).

Het panel van 10 beoordelaars kreeg 40 opgaven voor Luisteren I, 40 opgaven voor Luisteren II en 20 opgaven uit de Profieltoets NT2 in een random volgorde voorgelegd. De opgaven voor zowel Luisteren I als II vormden een geheel examen. De beoordelaars moesten voor elke opgave de volgende vraag beantwoorden:

*'Geef bij elke opgave aan welk niveau (A1, A2, B1, B2, C1, C2) minimaal vereist is om de opgave goed te kunnen maken.'*

Deze beoordelingstaak werd in twee zittingen uitgevoerd. Tijdens de eerste zitting moesten de beoordelaars onafhankelijk van elkaar deze vraag voor alle 100 opgaven beantwoorden. In een tweede zitting werden de opgaven waarover de minste overeenstemming bestond nogmaals aan de beoordelaars voorgelegd. De beoordelaars konden voor deze opgaven een toelichting op hun eerdere keuze geven. Elke beoordelaar kon na deze bespreking aangeven of hij zijn niveautoekenning al of niet wilde veranderen. Deze tweede bijeenkomst leidde tot een hogere mate van beoordelaarsovereenstemming.

De mate van overeenstemming tussen de beoordelaars over de niveautoekenning was bevredigend. Voor 85% van de opgaven gold dat minimaal 7 van de 10 beoordelaars het eens waren over het niveau dat nodig was om een opgave goed te maken. De beoordelaarsbetrouwbaarheid en –overeenstemming waren respectievelijk .92 en .91.

Uit de niveautoekenning aan de opgaven bleek dat er duidelijke niveaoverschillen waren tussen Luisteren I en Luisteren II. De gemiddelde niveautoekenning van opgaven van Luisteren I lag iets boven B1, die van de opgaven van Luisteren II net onder B2. De resultaten komen in hoge mate overeen met de eerdere bevindingen in de specificatiefase van het koppelingsonderzoek (Kuijper, 2009).

Op basis van de gemiddelde niveautoekenningen werden vervolgens de grensscores op de 500-schalen berekend voor B1 op Luisteren I en voor B2 op Luisteren II. Daarvoor werden twee criteria gehanteerd: het 80%-criterium en het 100% criterium. Bij het 80%-criterium moet een kandidaat om een bepaald niveau te behalen alle opgaven onder dat niveau en 80% van de opgaven van het betreffende niveau goed maken. Iemand haalt dus B1 wanneer hij alle opgaven van A1 en A2 en 80% van de B1 opgaven goed heeft gemaakt. Volgens het 100%-criterium moet een kandidaat alle opgaven onder en op het betreffende niveau goed maken. Het laatste criterium leidt vanzelfsprekend tot hogere grensscores. Tabel 3.2-8 geeft een overzicht van de huidige 500-cesuur en de 80%- en 100%-cesuur voor B1 bij Luisteren I en B2 bij Luisteren II. Kolom 3 geeft de daarbij behorende zakslaagpercentages berekend op de scoreverdeling van 2009 weer.

*Tabel 3.2-8 Cesaurscores, B1- respectievelijk B2-grensscores en slaagpercentages, berekend voor de populatie 2009, Luisteren I en II*

	Schaalscore	Ruwe toetsscore	Percentage geslaagd
<b>Luisteren I*</b>			
huidige cesuur	499/500	30/31**	53%
B1 grensscore 80%	492/493	21/22	93%
B1 grensscore 100%	497/496	26/27	78%
<b>Luisteren II</b>			
huidige cesuur	499/500	29/30	59%
B2 grensscore 80%	501/502	31/32	48%
B2 grensscore 100%	513/514	37/38	10%

\* De beoordelaars hebben aan 39 van de 40 opgaven een niveau toegekend, omdat men van één opgave geen niveau-indicatie konden geven.

\*\* De cesuurscore over alle 40 opgaven lag op 31/32

Kandidaten die slagen voor Luisteren I bevinden zich boven beide A2/B1-grensscores. Voor Programma II ligt de huidige cesuur onder de 80% grensscore van B1/B2 die op 501/502 ligt en de 100%-grensscore van 513/514.

Uit de berekeningen van de grensscores voor Lezen I en II bleek verder dat Luisteren I en Luisteren II niet geschikt zijn om andere niveaus te meten dan respectievelijk B1 en B2.

Op grond van deze tweede fase koppelingsonderzoek en het eerder uitgevoerde specificatie-onderzoek (Kuijper, 2009) kan geconcludeerd worden dat Staatsexamens Luisteren Programma I zich op B1 bevindt en Programma II op B2.

Deze conclusies hoeven niet per se te leiden tot een wijziging in de cesuren. De Staatsexamens kennen immers een functionele cesuur die gebaseerd is op het doel van de examens: het toetsen van een startkwalificatie voor de opleidingen en beroepen waarvoor beide programma's bedoeld zijn. De Staatsexamens NT2 zijn niet in eerste instantie bedoeld om een bepaald ERK-niveau te meten.

## 4 Samenvatting en conclusies

In 2003 publiceerde de Raad van Europa een concept-handleiding voor het relateren van taalexamens aan het ERK (Council of Europe, 2003). In 2009 verscheen een nieuwe versie van deze Handleiding (Council of Europe, 2009). In deze Handleiding worden verschillende methodes beschreven om de claim op een koppeling van een examen met het ERK te onderbouwen. De Handleiding heeft als doel om belanghebbenden mogelijke procedures aan de hand te doen waarmee men evidentie kan aanvoeren voor de relaties van een examen met het Common European Framework of Reference for Languages ( ERK ).

In de Handleiding worden drie fases van koppeling onderscheiden:

- *Specificatie*: nagaan op welke toelichtende ERK -schalen en descriptoren toetsen en examens betrekking hebben. Hiermee kan een eerste inschatting gegeven worden van het niveau van de toetsen en examens.
- *Standaardisatie*: via een in de Handleiding omschreven procedure experts vragen om de examenopgaven te classificeren in termen van ERK (er ERK-niveaus aan te koppelen). Op basis daarvan kan worden bepaald welke score een kandidaat bij een bepaald examen zou moeten behalen om te kunnen zeggen dat hij/zij een bij dat examen relevant ERK -niveau beheerst).
- *Empirische validatie*: psychometrische validering van de gegevens verzameld in de standaardisatieprocedure.

Voor Staatsexamens NT2 is al eerder onderzoek uitgevoerd naar de relatie tussen de vaardigheden Spreken en Schrijven en het ERK. Uit dit onderzoek bleek dat voor beide vaardigheden Programma I correspondeert met ERK-niveau B1 en Programma II met ERK-niveau B2. Van dit onderzoek is verslag gedaan in Kuijper e.a. (2004a, 2004b, 2005a, 2005b) en Bechger e.a. (2008).

Voor Lezen en Luisteren is de specificatiefase al uitgevoerd (Kuijper, 2009). Deze specificatiefase wees uit dat zowel voor Lezen als voor Luisteren Programma I teksten en opgaven zich in het niveaubereik B1 bevinden, terwijl de opgaven en teksten voor Programma II in het B2-niveaubereik liggen.

In dit rapport wordt verslag gedaan van de standaardisatiefase van het koppelingsonderzoek voor Lezen en Luisteren en van de empirische validering van deze procedure.

Bij de standaardisatieprocedure zijn de richtlijnen gevolgd uit de Handleiding van de Raad van Europa. Er is daarbij gebruik gemaakt van de mandjesmethode.

Een panel van 11 beoordelaars voor Lezen en een panel van 10 beoordelaars voor Luisteren hebben aan de opgaven van een examen Lezen respectievelijk Luisteren voor Programma I en een examen voor Programma II een ERK-niveau toegekend.

De mate van beoordelaarsovereenstemming bij de toekenning van niveaus aan de opgaven was zowel voor Lezen als Luisteren bevredigend.

Voor beide vaardigheden bleek dat de examens van Programma I en Programma II een duidelijk niveauverschil lieten zien. Lezen en Luisteren Programma I liggen in het B1-niveaubereik, terwijl de examens voor Programma II zich in het B2-niveaubereik bevinden. Deze resultaten zijn een bevestiging van de resultaten van de specificatiefase (Kuijper, 2009). Beide programma's blijken niet geschikt te zijn om andere niveaus te meten.

Op basis van de niveautoekenningen zijn vervolgens grensscores berekend voor het bereiken van B1 op de examens programma I en B2 voor Programma II, uitgedrukt op de 500-schalen van de vier examens.

Voor elk examen zijn twee grensscores berekend op grond van de volgende criteria:

1. Het 100%-criterium. Hierbij wordt de grenscore voor het bereiken van een ERK-niveau als volgt bepaald: voor het behalen van ERK-niveau x moet een kandidaat alle opgaven van niveau x en de niveaus daaronder goed maken. Deze procedure leidt tot het aantal opgaven dat een kandidaat goed moet maken op de examens die in de standaardsetting zijn opgenomen (zie paragraaf 2.2). Deze score kan vervolgens worden omgezet in de 500-score

2. Het 80%-criterium. Hierbij wordt de grensscore voor het bereiken van een ERK-niveau als volgt bepaald: voor het behalen van ERK-niveau x moet een kandidaat alle opgaven onder niveau x en 80% van niveau x goed maken. Deze procedure leidt tot het aantal opgaven dat een kandidaat goed moet maken op de examens die in de standaardsetting zijn opgenomen (zie paragraaf 2.2). Deze score kan vervolgens worden omgezet in de 500-score.

Geen van beide berekeningscriteria is het enig juiste. Ze kunnen worden gezien als een wat mildere en wat strengere wijze van het bepalen van grensscores en verschaffen zo een interval waarin een te bepalen grensscore zich zou kunnen bevinden. In analogie met het koppingsonderzoek van Spreken en Schrijven aan het ERK (Kuijper e.a. 2004, 2005) ligt het voor de hand dat het bij de Staatsexamens NT2 belangrijker is false negatives te voorkomen dan false positives. Daarom lijkt het hanteren van het 80%-criterium meer voor de hand te liggen dan het 100%-criterium.

Tabel 4-1 geeft een overzicht van de ERK-grensscores, de huidige cesuur en de zakpercentages bij deze scores.

*Tabel 4-1 Cesuurscores, B1- respectievelijk B2-grensscores en slaagpercentages, berekend voor de populatie 2009, Lezen en Luisteren*

	Schaalscore	Ruwe toetsscore	Percentage geslaagd
<b>Lezen I</b>			
huidige cesuur	499/500	27/28	59%
B1 grensscore 80%	478/479	21/22	86%
B1 grensscore 100%	492/493	25/26	69%
<b>Lezen II</b>			
huidige cesuur	499/500	26/27	59%
B2 grensscore 80%	510/511	28/29	46%
B2 grensscore 100%	538/539	33/34	22%
<b>Luisteren I*</b>			
huidige cesuur	499/500	30/31**	53%
B1 grensscore 80%	492/493	21/22	93%
B1 grensscore 100%	497/496	26/27	78%
<b>Luisteren II</b>			
huidige cesuur	499/500	29/30	59%
B2 grensscore 80%	501/502	31/32	48%
B2 grensscore 100%	513/514	37/38	10%

\* De beoordelaars hebben aan 39 van de 40 opgaven een niveau toegekend, omdat men van één opgave geen niveau-indicatie konden geven.

\*\* De cesuurscore over alle 40 opgaven lag op 31/32

Bij de examens voor Programma I liggen beide B1-grensscores (80%- en 100%-criterium) onder de huidige cesuur. Een kandidaat die slaagt voor Luisteren I of Lezen I beheerst dus B1.

Voor Programma II geldt dat de huidige cesuur voor zowel lezen als Luisteren onder de B2-grensscores ligt. Volgens de 80%-norm wordt bij Lezen het B2-niveau bereikt bij een score van 510/511 en bij Luisteren bij een score van 501/502. De 100%-grensscores liggen voor lezen en Luisteren respectievelijk op 538/539 en 513/514.

Men kan niet verwachten dat de beoordelaars het volledig eens zijn. Van een commissie die zich een B1- of B2-kandidaat voorstelt zal men dan ook variatie in de grensscores verwachten. Zelfs een kandidaat die vele malen een toets zou maken komt immers niet altijd tot de zelfde score. Een vergelijking tussen de spreiding in de grensscores op basis van de niveautoekenningen door de beoordelaars en de theoretische scoreverdeling van kandidaatscores op deze grensscores wees uit dat de grensscores van de beoordelaars binnen de theoretische scoreverdelingen van grensscorekandidaten vielen.

Het is belangrijk te benadrukken dat het gezien de functie van de Staatsexamens NT2 niet noodzakelijk is om de huidige cesuren te veranderen grond van de resultaten die in deze paragraaf besproken worden. De Staatsexamens kennen immers een functionele cesuur die gebaseerd is op het doel van de examens: het toetsen van een startkwalificatie voor de opleidingen en beroepen

waarvoor beide programma's bedoeld zijn. De Staatsexamens NT2 zijn niet in eerste instantie bedoeld om een bepaald ERK-niveau te meten. De resultaten die in deze paragraaf besproken maken het mogelijk om scores te relateren aan een ERK-niveau, maar hoeven er niet per se toe te leiden de huidige cesuren te veranderen.

## **Bijlage 1 Samenstelling van de beoordelaarsgroepen Lezen en Luisteren**

<b>Samenstelling beoordelaarspanel Lezen</b>
Coördinator NT2 universiteit
Docent NT2 ROC
Docent NT2 universiteit
Docent NT2 universiteit
Docent NT2 universiteit
Hoofd sectie Nederlands Talencentrum Universiteit
Lid CvE Staatsexamens Nt2
Projectmanager STEX Cito
Toetsdeskundige Cito
Oud medewerker CNaVT
Docent NT2 universiteit

<b>Samenstelling beoordelaarspanel Luisteren</b>
Docent NT2 universiteit
Docent NT2 universiteit
Toetsdeskundige CNaVT
Docent NT2 ROC
Toetsdeskundige Bureau ICE
Lid CvE Staatsexamens Nt2
Coördinator NT2 universiteit
Docent NT2 universiteit
Toetsdeskundige CNaVT
Hoogleraar NT2

## Bijlage 2 Materiaal bij de familiarisatie Lezen en Luisteren

### Lezen

#### Sorteerformulier globale schaal ERK

De onderstaande descriptoren uit de globale schaal staan door elkaar.

Geef per descriptor het bijbehorende ERK-niveau aan.

Niveau	Beschrijving
	Kan zinnen en regelmatig voorkomende uitdrukkingen begrijpen die verband hebben met zaken van direct belang (bijvoorbeeld persoonsgegevens, familie, winkelen, plaatselijke geografie, werk). Kan communiceren in simpele en alledaagse taken die een eenvoudige en directe uitwisseling over vertrouwde en alledaagse kwesties vereisen. Kan in eenvoudige bewoordingen aspecten van de eigen achtergrond, de onmiddellijke omgeving en kwesties op het gebied van directe behoeften beschrijven.
	Kan de belangrijkste punten begrijpen uit duidelijke standaardteksten over vertrouwde zaken die regelmatig voorkomen op het werk, op school en in de vrije tijd. Kan zich redden in de meeste situaties die kunnen optreden tijdens reizen in gebieden waar de taal wordt gesproken. Kan een eenvoudige lopende tekst produceren over onderwerpen die vertrouwd of die van persoonlijk belang zijn. Kan een beschrijving geven van ervaringen en gebeurtenissen, dromen, verwachtingen en ambities en kan kort redenen en verklaringen geven voor meningen en plannen.
	Kan de hoofdgedachte van een ingewikkelde tekst begrijpen, zowel over concrete als over abstracte onderwerpen, met inbegrip van technische besprekingen in het eigen vakgebied. Kan zo vloeiend en spontaan reageren dat een normale uitwisseling met moedertaalsprekers mogelijk is zonder dat dit voor een van de partijen inspanning met zich meebrengt. Kan duidelijke, gedetailleerde tekst produceren over een breed scala van onderwerpen; kan een standpunt over een actuele kwestie uiteenzetten en daarbij ingaan op de voor- en nadelen van diverse opties.
	Kan vrijwel alles wat hij of zij hoort of leest gemakkelijk begrijpen. Kan informatie die afkomstig is uit verschillende gesproken en geschreven bronnen samenvatten, argumenten reconstrueren en hiervan samenhangend verslag doen. Kan zichzelf spontaan, vloeiend en precies uitdrukken en kan hierbij fijne nuances in betekenis, zelfs in complexere situaties, onderscheiden.
	Kan vertrouwde dagelijkse uitdrukkingen en basiszinnen, gericht op de bevrediging van concrete behoeften, begrijpen en gebruiken. Kan zichzelf aan anderen voorstellen en kan vragen stellen en beantwoorden over persoonlijke gegevens zoals waar hij/zij woont, wie hij/zij kent en dingen die hij/zij bezit. Kan op een simpele wijze reageren, aangenomen dat de andere persoon langzaam en duidelijk praat en bereid is om te helpen.
	Kan een uitgebreid scala van veeleisende, lange teksten begrijpen en de impliciete betekenis herkennen. Kan zichzelf vloeiend en spontaan uitdrukken zonder daarvoor aantoonbaar naar uitdrukkingen te moeten zoeken. Kan flexibel en effectief met taal omgaan ten behoeve van sociale, academische en beroepsmatige doeleinden. Kan een duidelijke, goed gestructureerde en gedetailleerde tekst over complexe onderwerpen produceren en daarbij gebruikmaken van organisatorische structuren en verbindingswoorden.



## Sorteerformulier algemene leesvaardigheid ERK

De onderstaande descriptoren van niveaus van algemene leesvaardigheid staan door elkaar. Geef per descriptor het bijbehorende ERK-niveau aan.

Niveau	Beschrijving
	Kan met voldoende begrip directe feitelijke teksten lezen over onderwerpen die betrekking hebben op zijn of haar interessegebied.
	Kan lange, complexe teksten op detailniveau begrijpen, ongeacht of zij betrekking hebben op zijn of haar eigen vakgebied, mits hij of zij moeilijke passages kan herlezen.
	Kan in hoge mate zelfstandig lezen, past zijn of haar leesstijl en -snelheid aan verschillende teksten en doeleinden aan, en maakt selectief gebruik van toepasselijke naslagwerken. Beschikt over een grote actieve leeswoordenschat, maar kan enige moeite hebben met weinig voorkomende idiomatische uitdrukkingen.
	Kan zeer korte, eenvoudige teksten frase voor frase begrijpen door vertrouwde namen, woorden, en elementaire combinaties te herkennen en indien nodig te herlezen.
	Kan korte, eenvoudige teksten begrijpen over vertrouwde zaken van concrete aard, die zijn geschreven in veelgebruikte alledaagse of werkgebonden taal. Kan korte, eenvoudige teksten begrijpen, geschreven met de meest gebruikte woorden, met inbegrip van een aantal woorden uit de gemeenschappelijke internationale woordenschat.
	Kan vrijwel alle vormen van geschreven taal begrijpen en kritisch interpreteren, met inbegrip van abstracte, structureel complexe of zeer spreektaalige literaire en niet-literaire geschriften. Kan een breed scala van lange, complexe teksten begrijpen en daarbij subtiele verschillen in stijl en impliciete en expliciete betekenissen opmerken.

## Sorteerformulier algemene luistervaardigheid ERK

Niveau	Beschrijving
	<p>Kan directe feitelijke informatie verstaan over gewone alledaagse of werkgebonden onderwerpen, en daarbij zowel algemene boodschappen als specifieke details herkennen, mits het gesproken woord helder wordt gearticuleerd in een over het algemeen vertrouwd accent.</p> <p>Kan de hoofdpunten verstaan van heldere spraak in standaardtaal over vertrouwde zaken die regelmatig aan de orde komen op het werk, op school, in de vrije tijd, enzovoort, met inbegrip van korte verhalende teksten.</p>
	<p>Kan genoeg verstaan om uitgebreide betogen te volgen over abstracte en complexe onderwerpen buiten zijn of haar eigen vakgebied, al moet hij of zij misschien af en toe een detail laten bevestigen, vooral wanneer het accent niet vertrouwd is.</p> <p>Kan een breed scala van idiomatische uitdrukkingen en uitdrukkingen uit de spreektaal herkennen, en merkt daarbij registerverschuivingen op.</p> <p>Kan een uitgebreid betoog volgen, ook wanneer het niet duidelijk gestructureerd is en wanneer verbanden slechts worden geïmpliceerd en niet uitdrukkelijk worden benoemd.</p>
	<p>Kan spraak volgen die heel langzaam en zorgvuldig wordt uitgesproken, met lange pauzes om de betekenis op te nemen.</p>
	<p>Heeft geen moeite met het begrijpen van enige vorm van gesproken taal, hetzij in direct contact, hetzij via de media, voortgebracht met de snelheid van de moedertaalspreker.</p>
	<p>Kan gesproken standaardtaal verstaan, hetzij in direct contact, hetzij via de media, over vertrouwde en nietvertrouwde onderwerpen die gewoonlijk worden aangetroffen in het persoonlijke, sociale, academische-of professionele leven. Alleen extreem achtergrondgeluid, onvoldoende structuur in het betoog en/of idiomatisch taalgebruik ondermijnen het begrip van het gesprokene.</p> <p>Kan de hoofdedachten begrijpen van naar inhoud en vorm complexe gesproken tekst in een standaarddialect over concrete en abstracte onderwerpen, met inbegrip van technische besprekingen in het eigen vakgebied.</p> <p>Kan uitgebreide betogen en complexe redeneringen volgen, mits het onderwerp redelijk vertrouwd is en de richting van de discussie wordt aangegeven met expliciete signalen.</p>
	<p>Kan genoeg verstaan om te kunnen voldoen aan behoeften van concrete aard, mits er helder en langzaam wordt gearticuleerd.</p> <p>Kan frasen en uitdrukkingen verstaan die verband houden met zaken van de meest directe prioriteit (elementaire persoons- en familiegegevens, boodschappen doen, plaatselijke geografie, werk), mits er helder en langzaam wordt gearticuleerd.</p>

## Elementen in niveau-uitspraken CEFR Algemene Leesschaal

C2	Kan vrijwel <b>alle vormen</b> van geschreven taal <b>BEGRIJPEN</b> en <b>KRITISCH INTERPRETEREN</b> , met inbegrip van <b>abstracte, structureel complexe</b> of <b>zeer spreektaalige literaire en niet-literaire</b> geschriften. Kan een <b>breed scala</b> van <b>lange, complexe</b> teksten <b>BEGRIJPEN</b> en daarbij subtiele verschillen in stijl en impliciete en expliciete betekenissen <b>OPMERKEN</b>
C1	Kan <b>lange, complexe</b> teksten <b>OP DETAILNIVEAU BEGRIJPEN</b> , ongeacht of zij betrekking hebben op zijn of haar eigen vakgebied, mits hij of zij moeilijke passages kan herlezen
B2	Kan in hoge mate <b>ZELFSTANDIG LEZEN</b> , <b>PAST</b> zijn of haar leesstijl en -snelheid <b>AAN</b> verschillende teksten en doeleinden aan, en <b>MAAKT SELECTIEF GEBRUIK</b> van toepasselijke naslagwerken. <b>BESCHIKT OVER</b> een <b>grote actieve leeswoordenschat</b> , maar <b>KAN ENIGE MOEITE HEBBEN</b> met <b>weinig voorkomende idiomatische</b> uitdrukkingen
B1	Kan met voldoende begrip <b>directe (straightforward) feitelijke</b> teksten <b>LEZEN</b> over onderwerpen die betrekking hebben op zijn of haar interessegebied
A2	Kan <b>korte, eenvoudige</b> teksten <b>BEGRIJPEN</b> over vertrouwde zaken van <b>concrete</b> aard, die zijn geschreven in <b>veelgebruikte alledaagse of werkgebonden</b> taal. Kan <b>korte, eenvoudige</b> teksten <b>BEGRIJPEN</b> , geschreven met de <b>meest gebruikte woorden</b> , met inbegrip van een <b>aantal woorden uit de gemeenschappelijke internationale woordenschat</b> .
A1	Kan <b>zeer korte, eenvoudige</b> teksten frase voor frase <b>BEGRIJPEN</b> door <b>vertrouwde namen, woorden, en elementaire combinaties</b> te <b>HERKENNEN</b> en indien nodig te herlezen.

Rood tekstlengte

lila complexiteit

Blauw vocabulaire

groen tekstsoort

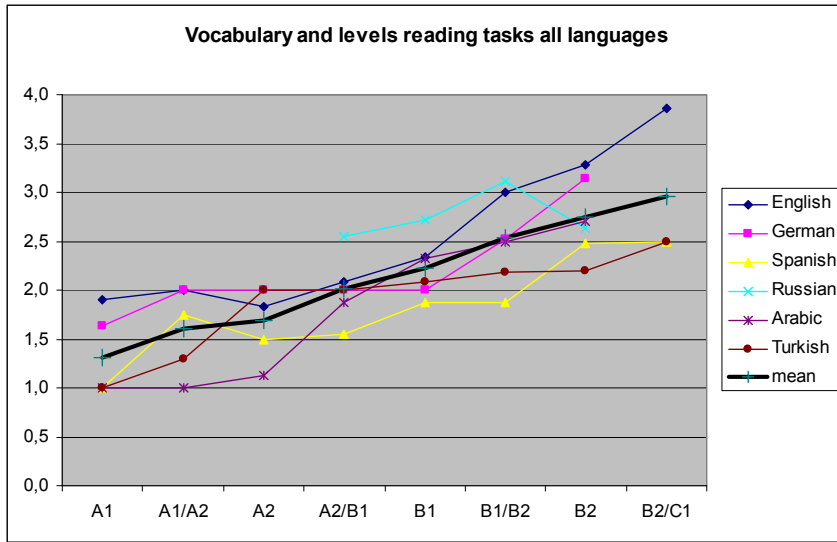


## Aspecten die samenhangen met het niveau

- Tekstsoort
- Cognitieve complexiteit
- Vocabulair
- Grammatica
- Tekstlengte



## Relatie vocabulaire en geschat CEFR-niveau



1 alleen frequent vocabulaire

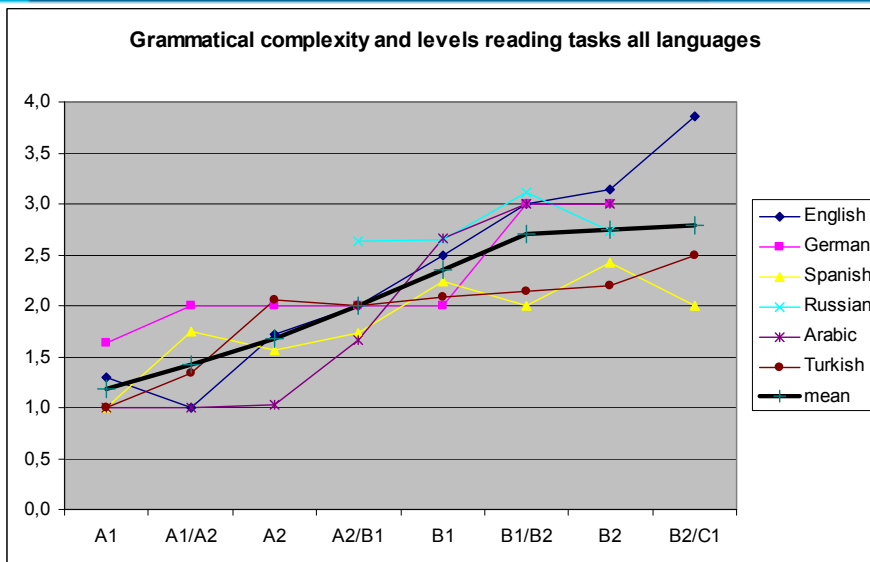
2 voornamelijk frequent vocabulaire

3 tamelijk uitgebreid vocabulaire

4 uitgebreid vocabulaire



## Relatie grammatica en geschat CEFR-niveau



1 alleen eenvoudige structuren

2 overwegend eenvoudige structuren

3 beperkte mate complexe structuren

4 ruime mate complexe structuren





C2	Heeft geen moeite met het BEGRIJPEN van enige vorm van gesproken taal, hetzij in direct contact, hetzij via de media, voortgebracht met de <i>snelheid van de moedertaalspreker</i> .
C1	Kan GENOEG VERSTAAN om <i>uitgebreide</i> betogen te VOLGEN over <i>abstracte en complexe</i> onderwerpen <i>buiten zijn of haar eigen vakgebied</i> , al moet hij of zij misschien af en toe een detail laten bevestigen, vooral wanneer het <i>accent</i> niet vertrouwd is. Kan een <i>breed scala van idiomatische uitdrukkingen en uitdrukkingen uit de spreektaal</i> HERKENNEN, en MERKT daarbij <i>registerverschuivingen</i> OP. Kan een <i>uitgebreid</i> betoog VOLGEN, ook wanneer het <i>niet duidelijk gestructureerd</i> is en wanneer verbanden slechts worden <i>geïmpliceerd</i> en <i>niet uitdrukkelijk</i> worden benoemd.
B2	Kan <i>gesproken standaardtaal</i> VERSTAAN, hetzij in direct contact, hetzij via de media, over <i>vertrouwde en nietvertrouwde onderwerpen die gewoonlijk worden aangetroffen in het persoonlijke, sociale, academische-of professionele leven</i> . Alleen extreem achtergrondgeluid, <i>onvoldoende structuur</i> in het betoog en/of <i>idiomatisch taalgebruik</i> ondermijnen het begrip van het gesprokene. Kan de hoofdgedachten BEGRIJPEN van naar <i>inhoud en vorm complexe</i> gesproken tekst in een standaarddialect over <i>concrete en abstracte</i> onderwerpen, met inbegrip van technische besprekingen in het <i>eigen vakgebied</i> . Kan uitgebreide betogen en <i>complexe</i> redeneringen VOLGEN, mits het onderwerp <i>redelijk vertrouwd</i> is en de richting van de discussie wordt aangegeven met expliciete signalen.
B1	Kan directe <i>feitelijke</i> informatie VERSTAAN over <i>gewone alledaagse of werkgebonden onderwerpen</i> , en daarbij zowel algemene boodschappen als specifieke details herkennen, mits het gesproken woord helder wordt gearticuleerd in een over het <i>algemeen vertrouwd accent</i> . Kan de HOOFDPUNTEN VERSTAAN van <i>heldere spraak</i> in <i>standaardtaal</i> over vertrouwde zaken die regelmatig aan de orde komen op het werk, op school, in de vrije tijd, enzovoort, met inbegrip van <i>korte verhalende teksten</i> .
A2	Kan GENOEG VERSTAAN om te kunnen voldoen aan behoeften van <i>concrete aard</i> , mits er <i>helder</i> en <i>langzaam</i> wordt gearticuleerd. Kan FRASEN EN UITDRUKKINGEN VERSTAAN die verband houden met zaken van de <i>meest directe prioriteit</i> (elementaire persoons-en familiegegevens, boodschappen doen, plaatselijke geografie, werk), mits er <i>helder</i> en <i>langzaam</i> wordt gearticuleerd.
A1	Kan spraak VOLGEN die <i>heel langzaam</i> en <i>zorgvuldig</i> wordt uitgesproken, met <i>lange pauzes</i> om de betekenis op te nemen.

tekstlengte   complexiteit   spreektempo   standaardaccent  
vocabulaire   tekstsoort   vertrouwdheid   articulatie   register

## Aspecten die samenhangen met het niveau

	Lezen	Luisteren
<b>Cognitieve complexiteit</b>		
<b>Vocabulaire</b>		
<b>Grammaticale complexiteit</b>		
<b>Register</b>		
<b>Tekstsoort</b>		
<b>Tekstlengte</b>		
<b>Vertrouwdheid</b>		
<b>Spreektempo</b>		
<b>Articulatie</b>		
<b>Accent</b>		

## Schalen voor linguïstische competenties. Bewerking Dutch CEFR-grid

<b>Cognitieve complexiteit</b>	<b>1 alleen concreet</b> <b>2 vooral concreet</b> <b>3 tamelijk abstract</b> <b>4 voornamelijk abstract</b>
<b>Vocabulaire</b>	<b>1 Alleen frequent</b> <b>2 Voornamelijk frequent</b> <b>3 Tamelijk uitgebreid</b> <b>4 Uitgebreid</b>
<b>Grammaticale complexiteit</b>	<b>1 Alleen eenvoudige structuren</b> <b>2 Voornamelijk eenvoudige structuren</b> <b>3 Beperkte mate complexe structuren</b> <b>4 Ruime mate complexe structuren</b>
<b>Spreektempo</b>	<b>1 kunstmatig langzaam</b> <b>2 langzaam</b> <b>3 normaal</b> <b>4 snel</b>
<b>Articulatie</b>	<b>1 kunstmatig duidelijke articulatie</b> <b>2 duidelijke articulatie</b> <b>3 normale articulatie</b> <b>4 soms onduidelijke articulatie</b>



### Bijlage 3 Aanbiedingsvolgorde van de opgaven in de standaardbepaling

#### Lezen

nr in standaardbepaling	oorspronkelijke toets	nr in oorspr toets	naam tekst
1	stex-I	1	opleiding gezondheidszorg
2	stex-I	2	opleiding gezondheidszorg
3	stex-I	3	opleiding gezondheidszorg
4	stex-I	4	opleiding gezondheidszorg
5	stex-I	5	opleiding gezondheidszorg
6	profiel-B	1	werkzaamheden NS
7	profiel-A	12	bereikbaarheid bank
8	stex-II	1	scriptieprijs
9	stex-II	2	scriptieprijs
10	stex-II	3	scriptieprijs
11	profiel-B	7	speeltuin
12	profiel-B	8	speeltuin
13	profiel-B	9	speeltuin
14	stex-I	6	contract/cao
15	stex-I	7	contract/cao
16	stex-I	8	contract/cao
17	stex-I	9	contract/cao
18	stex-I	10	contract/cao
19	stex-II	4	massage op de werkvloer
20	stex-II	5	massage op de werkvloer
21	stex-II	6	massage op de werkvloer
22	stex-II	7	massage op de werkvloer
23	stex-II	8	massage op de werkvloer
24	stex-II	9	massage op de werkvloer
25	stex-I	11	kraamzorg de waarden
26	stex-I	12	kraamzorg de waarden
27	stex-I	13	kraamzorg de waarden
28	stex-I	14	kraamzorg de waarden
29	stex-I	15	kraamzorg de waarden
30	profiel-B	3	België
31	profiel-B	4	België
32	profiel-A	2	aanbieding
33	profiel-A	3	aanbieding
34	stex-II	10	telewerken
35	stex-II	11	telewerken
36	stex-II	12	telewerken
37	stex-II	13	telewerken
38	stex-II	14	telewerken
39	stex-I	16	oliameter
40	stex-I	17	oliameter
41	stex-I	18	oliameter
42	stex-I	19	oliameter
43	stex-I	20	oliameter
44	profiel-B	14	kinderdagverblijf



45	profiel-A	10	tandarts
46	profiel-A	11	tandarts
47	stex-II	15	ziekteverzuim
48	stex-II	16	ziekteverzuim
49	stex-II	17	ziekteverzuim
50	stex-II	18	ziekteverzuim
51	stex-II	19	ziekteverzuim
52	stex-II	20	ziekteverzuim
53	stex-II	21	ziekteverzuim
54	stex-I	21	onderwijs- en examenreglement
55	stex-I	22	onderwijs- en examenreglement
56	stex-I	23	onderwijs- en examenreglement
57	stex-I	24	onderwijs- en examenreglement
58	stex-I	25	onderwijs- en examenreglement
59	stex-II	22	kleinere klassen
60	stex-II	23	kleinere klassen
61	stex-II	24	kleinere klassen
62	stex-II	25	kleinere klassen
63	stex-II	26	kleinere klassen
64	profiel-A	4	belangrijke data
65	profiel-A	5	belangrijke data
66	profiel-A	7	belangrijke data
67	profiel-A	8	belangrijke data
68	profiel-A	9	belangrijke data
69	profiel-B	28	toezichtteam
70	stex-I	26	zoek je weg op de arbeidsmarkt
71	stex-I	27	zoek je weg op de arbeidsmarkt
72	stex-I	28	zoek je weg op de arbeidsmarkt
73	stex-I	29	zoek je weg op de arbeidsmarkt
74	stex-I	30	zoek je weg op de arbeidsmarkt
75	stex-II	27	journalistiek schrijven
76	stex-II	28	journalistiek schrijven
77	stex-II	29	journalistiek schrijven
78	stex-II	30	journalistiek schrijven
79	stex-II	31	journalistiek schrijven
80	stex-I	31	vakgericht communicatie anderstalige
81	stex-I	32	vakgericht communicatie anderstalige
82	stex-I	33	vakgericht communicatie anderstalige
83	stex-I	34	vakgericht communicatie anderstalige
84	stex-I	35	vakgericht communicatie anderstalige
85	profiel-B	18	bloed geven
86	profiel-B	19	bloed geven
87	stex-II	32	nieuwe talenstudies
88	stex-II	33	nieuwe talenstudies
89	stex-II	34	nieuwe talenstudies
90	stex-II	35	nieuwe talenstudies
91	stex-I	36	brandalarm
92	stex-I	37	brandalarm
93	stex-I	38	brandalarm
94	stex-I	39	brandalarm
95	stex-I	40	brandalarm
96	stex-II	36	siemens

97	stex-II	37	siemens
98	stex-II	38	siemens
99	stex-II	39	siemens
100	stex-II	40	siemens

## Luisteren

nr in standaardbepaling	oorspronkelijke toets	nr in oorspr toets	naam tekst
<b>CD A</b>			
1	Profiel A	1	weerbericht
2	Profiel B	1	medicijnen innemen
3	STEX-I	1	zwembad
4	STEX-I	2	zwembad
5	STEX-I	3	zwembad
6	STEX-I	4	zwembad
7	STEX-I	5	zwembad
8	STEX-I	6	zwembad
9	STEX-I	7	zwembad
10	STEX-I	8	zwembad
11	STEX-I	9	zwembad
12	STEX-I	10	zwembad
13	STEX-I	11	zwembad
14	Profiel B	5	peter blom
15	Profiel B	8	doktersassistentie 2
16	STEX-II	1	meerlingen
17	STEX-II	2	meerlingen
18	STEX-II	3	meerlingen
19	STEX-II	4	meerlingen
20	STEX-II	5	meerlingen
21	STEX-II	6	meerlingen
22	STEX-II	7	meerlingen
23	STEX-II	8	meerlingen
24	STEX-II	9	meerlingen
25	STEX-II	10	meerlingen
26	Profiel B	9	doktersassistentie 3
27	Profiel A	4	Op de radio/3 plaatjes
28	Profiel A	7	geen werk gevonden
29	STEX-I	12	cliniclown
30	STEX-I	13	cliniclown
31	STEX-I	14	cliniclown
32	STEX-I	15	cliniclown
33	STEX-I	16	cliniclown
34	STEX-I	17	cliniclown
35	STEX-I	18	cliniclown
36	STEX-I	19	cliniclown
37	Profiel B	11	Monique wonen
38	Profiel A	8	kassa erbij
39	Profiel B	13	Monique wonen
40	Profiel A	9	welke zender
41	STEX-II	11	zeppelin
42	STEX-II	12	zeppelin
43	STEX-II	13	zeppelin
44	STEX-II	14	zeppelin
45	STEX-II	15	zeppelin
46	STEX-II	16	zeppelin
47	STEX-II	17	zeppelin

48	STEX-II	18	zeppelin
49	Profiel A	10	Ziekenhuis
50	Profiel A	11	In een groentewinkel
<b>CD B</b>			
51	Profiel B	16	Monique wonen
52	STEX-II	b1	danser
53	STEX-II	b2	danser
54	STEX-II	b3	danser
55	STEX-II	b4	danser
56	STEX-II	b5	danser
57	STEX-II	b6	danser
58	STEX-II	b7	danser
59	Profiel A	12	Het Postkantoor
60	Profiel A	24	De stad in 2
61	Profiel A	25	De stad in 3
62	STEX-I	b1	etaleuse
63	STEX-I	b2	etaleuse
64	STEX-I	b3	etaleuse
65	STEX-I	b4	etaleuse
66	STEX-I	b5	etaleuse
67	STEX-I	b6	etaleuse
68	STEX-I	b7	etaleuse
69	STEX-I	b8	etaleuse
70	STEX-I	b9	etaleuse
71	STEX-I	b10	etaleuse
72	STEX-I	b11	etaleuse
73	Profiel A	21	bedrijfskantine 4
74	Profiel A	27	De stad in 8
75	STEX-II	b8	secretaresse
76	STEX-II	b9	secretaresse
77	STEX-II	b10	secretaresse
78	STEX-II	b11	secretaresse
79	STEX-II	b12	secretaresse
80	STEX-I	c1	wegenwacht
81	STEX-I	c2	wegenwacht
82	STEX-I	c3	wegenwacht
83	STEX-I	c4	wegenwacht
84	STEX-I	c5	wegenwacht
85	STEX-I	c6	wegenwacht
86	STEX-I	c7	wegenwacht
87	STEX-I	c8	wegenwacht
88	STEX-I	c9	wegenwacht
89	STEX-I	c10	wegenwacht
90	Profiel A	22	bedrijfskantine 5
91	STEX-II	c1	bomenservice
92	STEX-II	c2	bomenservice
93	STEX-II	c3	bomenservice
94	STEX-II	c4	bomenservice
95	STEX-II	c5	bomenservice
96	STEX-II	c6	bomenservice

97	STEX-II	c7	bomenservice
<b>CD C</b>			
98	STEX-II	c10	bomenservice
99	STEX-II	c8	bomenservice
100	STEX-II	c9	bomenservice

# Bijlage 4 Scoreformulieren Lezen en Luisteren



cito



## Standaardbepaling Staatsexamens Lezen aan het Europees Raamwerk voor Talen

Naam beoordelaar: \_\_\_\_\_

Beo nr.: \_\_\_\_\_

Geef bij elke opgave aan welk niveau (A1, A2, B1, B2, C1, C2) minimaal vereist is om de opgave goed te kunnen maken. (Maak zo het hokje zwart ●).

label	Tekst nr.	Opgave nr.	A1	A2	B1	B2	C1	C2
183	1	1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
184	1	2	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
185	1	3	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
186	1	4	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
187	1	5	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1266	2	6	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1249	3	7	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
226	4	8	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
227	4	9	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
228	4	10	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1281	5	11	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1282	5	12	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1283	5	13	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
266	6	14	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
267	6	15	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
269	6	16	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
270	6	17	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
272	6	18	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
240	7	19	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
242	7	20	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
243	7	21	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
244	7	22	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
245	7	23	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
246	7	24	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
767	8	25	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
769	8	26	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
770	8	27	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
772	8	28	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
773	8	29	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1269	9	30	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

label	Tekst nr.	Opgave nr.	A1	A2	B1	B2	C1	C2
1270	9	31	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1237	10	32	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1238	10	33	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
753	11	34	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
754	11	35	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
755	11	36	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
757	11	37	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
758	11	38	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
735	12	39	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
736	12	40	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
737	12	41	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
738	12	42	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
739	12	43	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
1289	13	44	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
932	14	45	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
933	14	46	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
233	15	47	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
234	15	48	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
235	15	49	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
236	15	50	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
237	15	51	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
238	15	52	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
239	15	53	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
148	16	54	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
151	16	55	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
152	16	56	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
153	16	57	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
155	16	58	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
186	17	59	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
187	17	60	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>



MBO10-le-v1.0



label	Tekst nr.	Opgave nr.	A1	A2	B1	B2	C1	C2
188	17	61	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
189	17	62	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
190	17	63	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
747	18	64	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
748	18	65	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
750	18	66	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
751	18	67	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
752	18	68	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
992	19	69	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
741	20	70	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
742	20	71	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
743	20	72	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
746	20	73	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
747	20	74	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
181	21	75	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
182	21	76	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
183	21	77	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
184	21	78	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
185	21	79	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
173	22	80	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
174	22	81	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
175	22	82	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
176	22	83	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
177	22	84	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
989	23	85	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
990	23	86	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
747	24	87	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
748	24	88	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
749	24	89	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
750	24	90	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
227	25	91	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
228	25	92	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
229	25	93	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
230	25	94	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
231	25	95	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
783	26	96	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
784	26	97	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
785	26	98	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
786	26	99	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
787	26	100	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



## Standaardbepaling Staatsexamens Luisteren aan het Europees Raamwerk voor Talen

Naam beoordelaar: \_\_\_\_\_

Beo nr.: \_\_\_\_\_

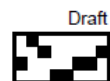
Geef bij elke opgave aan welk niveau (A1, A2, B1, B2, C1, C2) minimaal vereist is om de opgave goed te kunnen maken. (Maak zo het hokje zwart ●).

th-pronr	opgavenr.		eerste zin	A1	A2	B1	B2	C1	C2
	boekje	cd							
		CD A							
43895	1	1	1 Wat is dit?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43909	2	2	1 Een man moet medicijnen gebruiken	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14586	3	3	De opleiding duurt normaal gesproken....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14587	4	4	Module 1 leidt op tot toezichthouder	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14588	5	5	Over onregelmatige werktijden	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14589	6	6	Het is prachtig als, eh, als jonge kinderen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14590	7	7	Op zich lijkt het me alleen maar goed	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14591	8	8	In zijn algemeenheid kan je zeggen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14592	9	9	Dames en heren, Jochem kan nu een aantal vragen uit de zaal beantwoorden.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14593	10	10	Ik wil geen volledige baan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14594	11	11	In de krant lees je vaak over jongeren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14596	12	12	Als de kinderen in het water zijn, horen ze je dan wel?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14597	13	13	Dit is wel een heel sportief beroep.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43906	14	14	5 U hoort een telefoongesprek	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43714	15	15	8 Op hoeveel dagen is de vrouwelijke huisarts bereikbaar?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14429	16	16	Hallo met mevrouw Leijser	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14431	17	17	Goedemiddag, met Meneer Kozar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14432	18	18	Goedemiddag met mevrouw Zeilmans.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14433	19	19	Als ik met mijn dochter Amber en de tweeling	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14434	20	20	Dag, met mevrouw Jager.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14436	21	21	Hallo met mevrouw Martins.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14438	22	22	Met mevrouw Alzar.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14439	23	23	Met meneer van Zuijlen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14440	24	24	Goedemiddag met meneer van Dal.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14441	25	25	Eh met Conny Busser.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43715	26	26	9 Wat vindt Angeline NIET leuk om te doen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43020	27	27	4 U hoort een nieuwsbericht op de radio.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43883	28	28	U hoort een gesprek over het werk.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14098	29	29	Kan elke clown Cliniclown worden?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14099	30	30	Heb je veel contact met andere Cliniclowns?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>





th-pronr	opgavenr.		eerste zin	A1	A2	B1	B2	C1	C2
	boekje	cd							
14100	31	31	Werkt er in elk ziekenhuis een Cliniclown?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14101	32	32	Hoe reageert het verplegend personeel op cliniclowns?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14102	33	33	Je ziet in je werk veel zieke kinderen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14103	34	34	Kun je een voorbeeld geven van wat jullie doen voor de kinderen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14104	35	35	Hou je rekening met de leeftijd van de kinderen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14105	36	36	Wat vind je moeilijk aan je werk?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43779	37	37	11 Wat zegt Monique over haar huis?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43886	38	38	8 U hoort een tekst.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43781	39	39	13 Wat zegt Monique over het onderhoud van haar huis?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43898	40	40	9 Naar welke zender wil de man kijken?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14442	41	41	U zei dat u netcoördinator was.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14443	42	42	En wat wilt u bereiken met uw werk, met Z@ppelin?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14444	43	43	Veel ouders van jonge kinderen hebben bezwaar tegen de inhoud van Zappelin	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14445	44	44	Veel ouders zeggen, dat de grenzen tussen de programmaonderdelen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14447	45	45	Bent u niet bang dat het jeugdjournaal	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14448	46	46	Er wordt wel beweerd dat veel televisie kijken	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14451	47	47	Zijn er ook programma's die u niet geschikt vindt voor kinderen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14452	48	48	Hoe denk je over sportprogramma's voor kinderen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43403	49	49	10 Marco van Doeten moet naar het ziekenhuis	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43023	50	50	11 Een klant staat in een groentewinkel	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
		CD B							
43784	51	1	16 Wat vertelt Monique over de grootte van haar huis?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14132	52	2	Wat voor soort dansgezelschap is introdans?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14133	53	3	Kun je iets vertellen over de leeftijd..	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14135	54	4	Een choreograaf werkt in de studio	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14136	55	5	Kun je iets zeggen over de relatie..	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14138	56	6	Zitten er ook moeilijke kanten aan het dansen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14139	57	7	Hoe is dat nu als een danser moet invallen voor een zieke collega?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14140	58	8	Sommige mensen vergelijken dansen met topsport.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43812	59	9	Ankie vraagt wanneer het postkantoor open is.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43912	60	10	24 Hoeveel geld geeft John aan Karin?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



th-pronr	opgavenr.		eerste zin	A1	A2	B1	B2	C1	C2
	boekje	cd							
43913	61	11	25 Op de markt horen John en Karin een koopman hard hard schreeuwen.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14454	62	12	Zou het een goed idee zijn om te laten zien...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14455	63	13	Welke materialen gebruik jij als etaleur..	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14456	64	14	Denk je dat er genoeg werk is voor etaleurs?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14457	65	15	Heb je nog tips hoe ik het beste met winkelchefs om kan gaan?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14458	66	16	Kun je vertellen hoe de prijs van een etalage berekend wordt?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14459	67	17	Zijn er nog andere dingen waar ik op moet letten?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14460	68	18	Kun je trouwens ook als etaleur in vaste dienst werken?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14461	69	19	Is er een duidelijk verschil tussen de winkels in het westen, de Randstad...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14462	70	20	Er zijn natuurlijk op de hele wereld winkels met etalages	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14463	71	21	Sommige winkels hebben een heel kleine etalage	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14464	72	22	Vind jij het nog steeds leuk om als etaleur te werken?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43893	73	23	21 Welke soep kan de man het best nemen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43918	74	24	Waarom gaan John en Karim met de taxi terug naar huis?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14488	75	25	Kunt u vertellen waarom u op deze functie solliciteert?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14489	76	26	Een belangrijke taak voor de afdelingssecretaresse is ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14490	77	27	Is het voor u belangrijk om zelfstandig te kunnen werken?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14492	78	28	Een belangrijk onderdeel van de functie is het notuleren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14493	79	29	Als u een verslag af hebt, wat gebeurt er dan mee?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14282	80	30	Werken bij de wegwacht, is dat nou heel anders ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14284	81	31	Ik heb gehoord dat jullie een technische helpdesk hebben	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14285	82	32	Gebeuren er wel eens ongelukken?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14286	83	33	Heeft iedere wegwacht een eigen auto?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14287	84	34	Ik heb gehoord dat je als wegwacht veel contact hebt met de meldkamer.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14289	85	35	Als wegwacht werk je alleen. Wat vind je daarvan	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14290	86	36	Kun je uitleggen wat er nou zo leuk is aan je werk?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14291	87	37	Je werkt al langs als wegwacht.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14292	88	38	Je werkt veel buiten. Wat vind je daarvan?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14293	89	39	Kun je iets vertellen over je werktijden?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
43894	90	40	22 Hoeveel leent de man van zijn collega?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14509	91	41	Kun je me iets meer vertellen over de werkzaamheden ...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14510	92	42	Oké, en werk ik in teamverband?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14511	93	43	Het bedrijf werkt veelal voor gemeenten hé?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14513	94	44	Moet ik tijdens een onderzoek ook graven?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14514	95	45	Op dit moment volg ik dezelfde opleiding als jij hebt gedaan.	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14516	96	46	Als ik nou een boom ga beoordelen, waar moet ik dan op letten?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14517	97	47	Als ik een boom ga beoordelen, heb ik dan een lijst met punten....	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
		CD C							
14521	98	1	Tot slot nog een heel praktische vraag	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14518	99	2	Wat kan ik tijdens mijn stage nog meer doen?	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
14520	100	3	Ik heb begrepen dat je ook allerlei instrumenten meeneemt...	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>



## Bijlage 5 Grensscores volgens beoordelaars en scoreverdeling grenskandidaten B1 en B2 voor Lezen

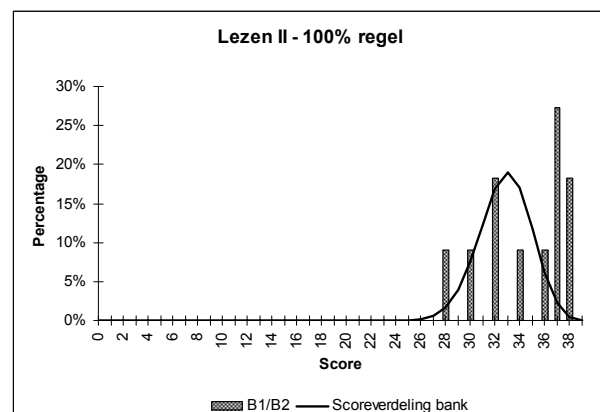
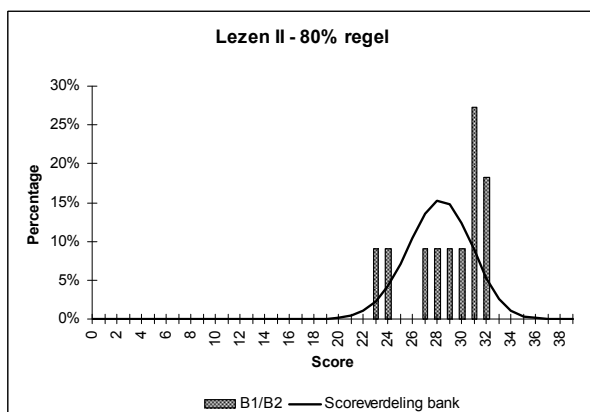
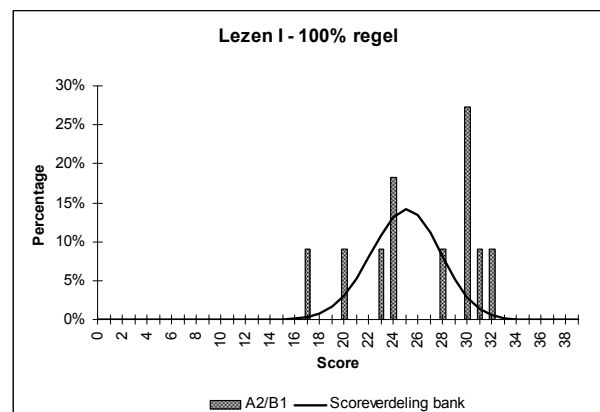
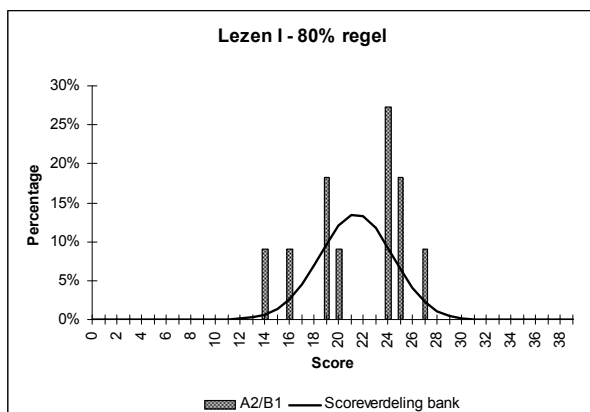
De beoordelaars is gevraagd van elk item aan te geven door de volgende vraag te beantwoorden: *'Geef bij elke opgave aan welk niveau (A1, A2, B1, B2, C1, C2) minimaal vereist is om de opgave goed te kunnen maken.'*

Hiermee wordt de beoordelaars in feite gevraagd zich een typische STEX-kandidaat voor te stellen van een bepaald niveau – dus vaardigheid – en te voorspellen hoe een kandidaat zich gedraagt op de opgaven.

Deze procedure leidt er toe dat de gemiddelden van de niveautoekenningen over de beoordelaars heen wordt omgezet in een toetsscore die nodig is om een bepaald niveau te halen. Voor Lezen I en Lezen II zijn op basis van de standaardsetting de 80% en 100%-grensscores voor respectievelijk A2/B1 en B1/B2 berekend. Voor Lezen I waren dit respectievelijk 21/22 en 25/26, voor Lezen II 28/29 en 34/35, corresponderend met schaalscores 478/479 en 492/493 voor Lezen I en 510/511 en 538/539 voor Lezen II.

Men kan niet verwachten dat de beoordelaars het volledig eens zijn. Van een commissie die zich een B1- of B2-kandidaat voorstelt zal men dan ook variatie in de grensscores verwachten. Zelfs een kandidaat die vele malen een toets zou maken komt immers niet altijd tot de zelfde score. Men kan een theoretische scoreverdeling voor een kandidaat met een bepaalde score berekenen. Dit is de verdeling van scores die een kandidaat volgens ons statistisch model zou halen als hij dezelfde toets vele malen zou maken, overigens zonder dat hij zich de vorige keer zou herinneren. Als het de beoordelaars gelukt is zich een typische STEX-kandidaat voor te stellen met dezelfde vaardigheid, dan verwachten we dat hun scores zich ook ongeveer zo verdelen en dus op de theoretische scoreverdeling van een kandidaat passen.

Onderstaande figuur geeft voor Lezen I en Lezen II de theoretische scoreverdelingen van een kandidaat die horen bij de 80%- en 100% grensscores. De staven geven de grensscores aan van de afzonderlijke beoordelaars. Een staaf van 9% vertegenwoordigt één beoordelaars, een staaf van 18% 2 beoordelaars en een staaf van 27% 3 beoordelaars.



Idealiter zouden we willen dat alle beoordelaars zich op of direct rond de gemiddelde grensscore zouden bevinden. Hoewel dit niet het geval is, valt de verdeling van de beoordelaars in alle gevallen binnen de theoretisch scoreverdeling van een kandidaat met die grensscores. Het beoordelingspanel als geheel is bij Lezen dus redelijk, zij het niet perfect, in staat zich een B1- respectievelijk B2-kandidaat voor te stellen.

## Bijlage 6 Grensscores volgens beoordelaars en scoreverdeling grenskandidaten B1 en B2 voor Luisteren

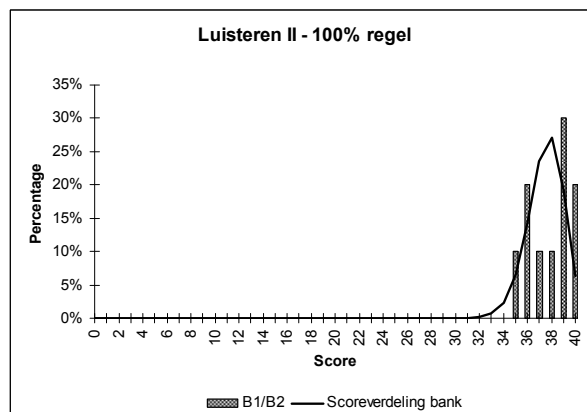
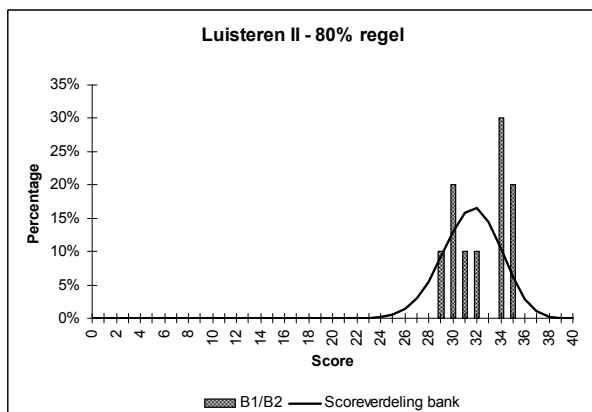
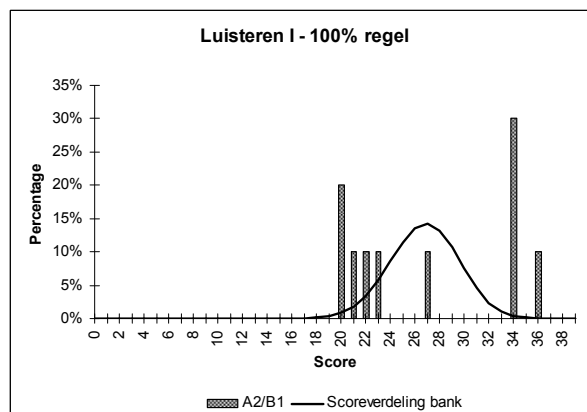
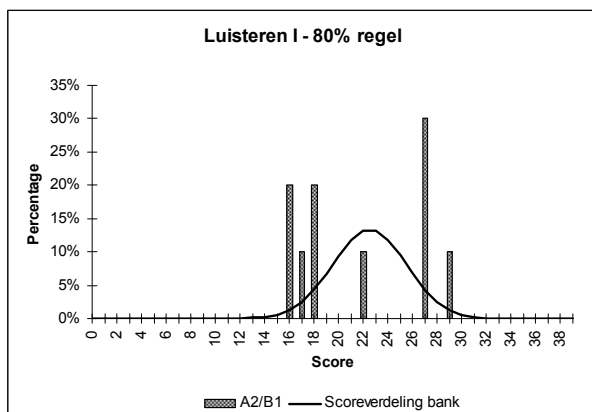
De beoordelaars is gevraagd van elk item aan te geven door de volgende vraag te beantwoorden: *'Geef bij elke opgave aan welk niveau (A1, A2, B1, B2, C1, C2) minimaal vereist is om de opgave goed te kunnen maken.'*

Hiermee wordt de beoordelaars in feite gevraagd zich een typische STEX-kandidaat voor te stellen van een bepaald niveau – dus vaardigheid – en te voorspellen hoe een kandidaat zich gedraagt op de opgaven.

Deze procedure leidt er toe dat de gemiddelde niveautoekenning over de beoordelaars heen wordt omgezet in een toetsscore die nodig is om een bepaald niveau te halen. Voor Luisteren I en Luisteren II zijn op basis van de standaardsetting de 80% en 100%-grensscores berekend. Voor Luisteren I waren dit respectievelijk 21/22 en 26/27, voor Luisteren II 31/32 en 37/38, corresponderend met schaalscores 493 en 496 voor Luisteren I en 502 en 510 voor Luisteren II.

Men kan niet verwachten dat de beoordelaars het volledig eens zijn. Van een commissie die zich een B1- of B2-kandidaat voorstelt zal men dan ook variatie in de grensscores verwachten. Zelfs een kandidaat die vele malen een toets zou maken komt immers niet altijd tot de zelfde score. Men kan een theoretische scoreverdeling voor een kandidaat met een bepaalde score berekenen. Dit is de verdeling van scores die een kandidaat volgens ons statistisch model zou halen als hij dezelfde toets vele malen zou maken, overigens zonder dat hij zich de vorige keer zou herinneren. Als het de beoordelaars gelukt is zich een typische STEX-kandidaat voor te stellen met dezelfde vaardigheid, dan verwachten we dat hun scores zich ook ongeveer zo verdelen en dus op de theoretische scoreverdeling van een kandidaat passen.

Onderstaande figuur geeft voor Luisteren I en Luisteren II de scoreverdelingen die horen bij de 80%- en 100% grensscores. De staven geven de grensscores aan van de afzonderlijke beoordelaars. Een staaf van 10% vertegenwoordigt één beoordelaars, een staaf van 20% 2 beoordelaars en een staaf van 30% 3 beoordelaars.



Idealiter zouden we willen dat alle beoordelaars zich op of direct rond de gemiddelde grensscore zouden bevinden. Hoewel dit niet het geval is, valt de verdeling van de beoordelaars voor Luisteren II 80%-grensscore en die van beide grensscores voor Luisteren II binnen de theoretische scoreverdeling van de grensscores. Alleen bij de 100% grensscore voor Luisteren I valt één beoordelaar buiten deze verwachte verdeling, terwijl drie beoordelaars zich net op de grens ervan bevinden. Het beoordelingspanel als geheel is bij Luisteren dus redelijk, zij het niet perfect, in staat zich een B1-respectievelijk B2-kandidaat voor te stellen. Enige twijfel bestaat hierover bij de 100%-grensscore voor Luisteren I.

## Referenties

Bechger, Timo M., Henk Kuijper; Gunter Maris. Standard Setting in Relation to the Common European Framework of Reference for Languages: The Case of the State Examination of Dutch as a Second Language. In *Language Assessment Quarterly*, Volume 6 Issue 2 2009, 126-150.

Council of Europe (2001), *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press, Cambridge.

Council of Europe (2003), Language Policy Division, *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching and Assessment. Manual Preliminary Pilot Version*. Council of Europe, Strasbourg.

Council of Europe (2009), *Relating Language Examinations to the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR). A Manual*. Council of Europe, Strasbourg.

Kuijper, Henk, Aukje Bergsma, Timo Bechger (2004a). *Staatsexamen NT2. Het gewenste niveau. Deel 1 Behoeftepeiling Programma II*. Cito. Arnhem, Augustus 2004.

Kuijper, Henk, Aukje Bergsma, Timo Bechger (2004b). *Staatsexamen NT2. Het gewenste niveau. Deel 2B Cesuurindicatie Spreken en Schrijven Programma II. Aanvullend onderzoek*. Cito, Arnhem, December 2004.

Kuijper, Henk, Aukje Bergsma, Timo Bechger (2005a). *Staatsexamen NT2. Het gewenste niveau. Deel 3 Behoeftepeiling Programma I*. Cito. Arnhem, Mei 2005.

Kuijper, Henk, Aukje Bergsma, Timo Bechger (2005b). *Staatsexamen NT2. Het gewenste niveau. Deel 4 Cesuurindicatie Spreken en Schrijven Programma I*. Cito, Arnhem, Augustus 2005.

Kuijper, Henk (2009). *De koppeling van de Staatsexamens Nederlands als Tweede Taal Lezen en Luisteren aan het Europees Referentiekader voor Talen. Fase 1 Specificatie*. Cito, Arnhem, november 2009.

Nederlandse Taalunie (2006), *Gemeenschappelijk Europees Referentiekader voor Moderne Vreemde Talen: Leren, Onderwijzen, Beoordelen*. Nederlandse Taalunie, Den Haag.

Noijons, José en Henk Kuijper (2006). *De koppeling van de centrale examens leesvaardigheid moderne vreemde talen aan het Europees Referentiekader*. Cito, Arnhem, januari 2007 (Cito, januari 2006).

Noijons José, Henk Kuijper en Evelyn Reichard (2007). *De koppeling van de centrale examens leesvaardigheid Arabisch, Russisch, Spaans en Turks aan het Europees Referentiekader*. Cito, Arnhem, januari 2007.