

De praktijk van de eerste en tweede correctie van het CSE | Technische rapportage van de landelijke enquête van 2016 in vergelijking met 2011

Hans Kuhlemeier en Paul van der Molen



De praktijk van de eerste en tweede correctie van het CSE

Technische rapportage van de landelijke
enquête van 2016 in vergelijking met 2011

Hans Kuhlemeier en Paul van der Molen
Met medewerking van Ger Limpens

Colofon

Aan dit rapport hebben ook meegewerkt:

Ger Limpens

Hugo Gitsels

Stefan Boom

Michelle Eikelboom

Jeroen Loeffen

Opmaak: Media Support

Foto omslag: Ron Steemers

© Stichting Cito Instituut voor Toetsontwikkeling Arnhem (2016)

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

Inhoud

1	Inleiding	6
2	Methode van onderzoek	10
2.1	Populatie, steekproef en respons	10
2.2	vragenlijst	11
2.3	Codering van de antwoorden op de open vragen	12
2.4	Statistische analyse	13
2.5	Representativiteit	14
2.6	Interpretatie van de resultaten	15
3	Resultaten	18
3.1	Tijdbesteding en ervaren belasting	18
3.1.1	Tijdbesteding aan de eerste correctie	18
3.1.2	Tijdbesteding aan de tweede correctie	21
3.1.3	Tijdbesteding aan de eerste en tweede correctie tezamen	26
3.1.4	Verschillen tussen de tijdbesteding aan de eerste en tweede correctie	30
3.1.5	Ervaren belasting van de eerste en tweede correctie	35
3.1.6	Toelichtingen van docenten bij de vraag naar de ervaren belasting van de eerste en tweede correctie	39
3.2	Schoolbeleid en tevredenheid	42
3.2.1	Het schoolbeleid met betrekking tot de eerste en tweede correctie	42
3.2.2	Toelichtingen van docenten bij de vragen naar de facilitering van de eerste en tweede correctie	45
3.2.3	Tevredenheid over het schoolbeleid	47
3.3	Volledigheid van de tweede correctie	49
3.3.1	Wel of geen tweede correctie	50
3.3.2	Volledigheid van de tweede correctie volgens de eerste corrector	50
3.3.3	Volledigheid van de tweede correctie volgens de tweede corrector	51
3.3.4	Niet-integrale tweede correctie	52
3.3.5	Gefaseerde volledige tweede correctie	53
3.3.6	Toelichtingen van docenten bij de stellingvragen naar de volledigheid van de tweede correctie	54
3.3.7	Percentage gecorrigeerde kandidaten en vragen	58
3.3.8	Percentage integraal nakijkende tweede correctoren	62
3.3.9	Mening over de noodzaak van integrale tweede correctie	65
3.3.10	Oorzaken van niet-integrale tweede correctie	67
3.3.11	Toelichtingen van docenten bij de vraag naar de oorzaken van onvolledige tweede correctie	68
3.3.12	Maatregelen ter bevordering van integrale tweede correctie	70
3.3.13	Toelichtingen van docenten bij de vraag naar maatregelen ter bevordering van volledige tweede correctie	71
3.4	Kwaliteit van de eerste en tweede correctie	73
3.4.1	Toepassing van het correctievoorschrift	74
3.4.2	Zorgvuldigheid van de correctie	74
3.4.3	Soepelheid van de correctie	79
3.4.4	Verbeteringsmaatregelen	83

4	Samenvatting, discussie en aanbevelingen	92
4.1	Samenvatting	93
4.2	Discussie en aanbevelingen	103
	Literatuur	112
	Bijlage 1: Resultaten flankerende enquêtes NVvW, NVON en Levende Talen	114

1 Inleiding

1 Inleiding

Over de correctie van de centraal schriftelijke examens in het voortgezet onderwijs is al decennia veel te doen. In 2011 en 2013 publiceerde Cito drie studies naar het functioneren van de eerste en tweede correctie. In het hoofdonderzoek hebben in totaal dertig onafhankelijke correctoren het examenwerk van 803 examenkandidaten van 127 scholen opnieuw nagekeken (Kuhlemeier, Van Rijn & Kremers, 2011). Daarbij zijn de scores van de derde correctoren vergeleken met die van de 'eigen' docent. In het tweede onderzoek hebben panels van getrainde correctoren een selectie van vragen en examenwerken nog een keer nagekeken (Kuhlemeier, Gitsels, Boom & Van de Kerkhof, 2011). Het panelonderzoek was bedoeld als check op de resultaten van het hoofdonderzoek. Daarnaast is nagegaan in hoeverre verschillen in soepelheid tussen correctoren samenhangen met de aard van de vragen en het beoordelingsmodel. Het derde onderzoek was een landelijke enquête naar de praktijk van de eerste en tweede correctie waaraan in totaal 3695 docenten (respons: 62%) hebben meegedaan (Kuhlemeier & Kremers, 2012). Hierna vatten we de belangrijkste conclusies van het onderzoek kort samen:

- De directe invloed van de tweede corrector op de uiteindelijke scores blijkt klein. De conclusie is dat de huidige procedure waarbij de ene docent de ander adviseert en controleert geen afdoende oplossing biedt voor mogelijke bevoordeling of benadeling.
- Eerste correctoren kennen gemiddeld 6% hogere scores toe dan 'onafhankelijke' derde correctoren (voor het eerste tijdvak). Zou de huidige eerste correctie worden vervangen door de derde correctie uit het onderzoek, dan zouden de percentages onvoldoendes aanzienlijk toenemen. Het verschil tussen de eerste en derde correctie is niet voor elk examen gelijk. Bij wiskunde stijgt het percentages onvoldoendes van 23% naar 28%, bij Engels van 27% naar 35%, bij biologie van 15% naar 25%, bij Nederlands van 20% naar 30%, bij geschiedenis van 23% naar 53% en bij tehatex van 22% naar 54%.
- Een toegeeflijke beoordeling leidt niet tot ongelijkheid als alle kandidaten in dezelfde mate beoordeeld worden. Er zijn echter grote verschillen tussen scholen in de soepelheid van de beoordeling. Op de ene school profiteren leerlingen daar veel meer van dan op de andere school. De geconstateerde verschillen in soepelheid betekenen dat kandidaten met eenzelfde vaardigheidsniveau op de ene school gemakkelijker een voldoende behalen dan op de andere school.
- Bij geschiedenis, tehatex en Nederlands blijken er docenten te zijn die hun kandidaten onverklaarbaar toegeeflijk beoordelen zonder dat de tweede correctie dat corrigeert.
- Verschillen tussen correctoren in soepelheid hangen samen met de aard van de vraag en het beoordelingsmodel. Zo blijken verschillen tussen correctoren vooral voor te komen als het antwoordmodel onvolledig is (dat wil zeggen de kandidaten geven dan antwoorden die niet in het antwoordmodel voorkomen) en het antwoordmodel alleen een voorbeeld geeft van volledig juiste antwoorden (en dus geen steun biedt bij het herkennen van en het toekennen van punten aan minder juiste of geheel foute antwoorden).
- Uit het vragenlijstonderzoek blijkt dat een derde van de docenten de tweede correctie volledig uitvoert. Tijdsdruk wordt als de belangrijkste oorzaak van het niet volledig uitvoeren van de tweede correctie gezien.

De uitkomsten van het onderzoek kregen in de landelijke media veel aandacht. Staatssecretaris Dekker van OCW heeft de Tweede Kamer in maart 2013 geïnformeerd over de resultaten van de drie onderzoeken. De samenvatting met de belangrijkste resultaten (Kuhlemeier & Kremers, 2013) leidde in de Tweede Kamer tot een stevig debat over onder meer de noodzaak van integrale tweede correctie. Daarop verzocht de staatssecretaris de VO-raad pilots uit te voeren om zo beter zicht te krijgen op mogelijke oplossingsrichtingen. Tevens vroeg hij de Inspectie om bij de examens van 2013 extra toe te zien op het naleven van de eisen rond de tweede correctie. Daarnaast kondigde hij aan de besturen en schoolleiders, samen met de VO-raad, aan te

spreken op het uitvoeren van de regels van correctie van de eindexamens en het faciliteren van de tweede correctie.

Op verzoek van de staatssecretaris heeft de VO-raad in het voorjaar van 2014 op tien scholen een pilot met omkering van de correctievolgorde uitgevoerd (Oberon, 2014). Daarbij werd het werk van de kandidaten eerst door een corrector van een andere school, de gecommiteerde, gecorrigeerd en pas daarna door de eigen docent. Uit de pilot bleek onder meer dat de tweede correctie nu wel integraal werd uitgevoerd en dat onder de deelnemende docenten voldoende steun bestond voor landelijke invoering (ondanks het gegeven dat de beschikbare periode voor het corrigeren wat korter en de werkdruk dus wat hoger was). Daarop besloot de staatssecretaris in april 2015 tot omkering van de correctievolgorde. Dit betekent dat het examenwerk eerst door een docent van een andere school en dan pas door de 'eigen' docent wordt nagekeken. De maatregel zou vanaf 2016 ingaan. De omkering zou naar verwachting tot een zorgvuldiger beoordeling leiden omdat de tweede corrector geen houvast meer heeft aan de resultaten van de eerste correctie en dus alles nakijkt. De (oorspronkelijk) eerste corrector doet een zorgvuldige eindcontrole, omdat het om 'zijn' of 'haar' leerlingen gaat. Er wordt dus beide kanten integraal naar het gemaakte werk gekeken. De omkering van de correctievolgorde leidt aldus de staatssecretaris niet tot meer werkdruk aangezien het corrigeren immers al onderdeel is van de taken van een docent.

Het besluit tot omkering van de correctievolgorde bleek in het veld op veel weerstand te stuiten. In zijn brief van juni 2015 aan de Tweede Kamer gaf de staatssecretaris aan bereid te zijn om het besluit nog een jaar aan te houden. Eerst wilde hij bezien of de huidige werkwijze met de inzet van alle betrokken partijen toch goed ingevuld kon worden. Daarbij stelde hij wel de uitdrukkelijke voorwaarde dat de betrokken partijen zich hieraan committeren. Hij gaf daarbij aan te verwachten dat de VO-raad haar leden oproept om de docenten in staat te stellen de correctie volledig uit te laten voeren. Van de bonden verwachtte hij dat zij hun achterban oproepen hier consciëntieus invulling aan te geven en zo nodig collegiaal oplossingen te zoeken en - waar onvermijdelijk - het gesprek met de schoolleiding aan te gaan. Tegelijkertijd werd het CvTE nogmaals gevraagd de beoordelingsvoorschriften tegen het licht te houden en zaken die wellicht tot verwarring kunnen leiden aan te passen. Daarnaast werd het CvTE verzocht te bezien in hoeverre correctie-intensieve vakken in de planning naar voren gehaald kunnen worden. Tot slot kondigde hij aan de correctie van de centrale examens in 2016 opnieuw door Cito te laten onderzoeken en hierover te rapporteren. Als zou blijken dat de tweede correctie nu wel integraal wordt uitgevoerd zou hij afzien van omkering. Mocht het gewenste resultaat uitblijven, dan zou hij de eerder gemelde omkering alsnog in het schooljaar 2016-2017 in procedure brengen.

Na de aanhouding van het besluit tot omkering van de correctievolgorde hebben de betrokken partijen verschillende maatregelen genomen ter bevordering van een zorgvuldige en integrale tweede correctie:

- De VO-raad vroeg scholen met klem de organisatie van de correctie opnieuw te bezien zodat docenten optimaal in de gelegenheid worden gesteld om de correctie integraal uit te voeren.
- De bonden riepen hun leden op ervoor te zorgen dat de uitvoering van de correctie aan de vereiste kwaliteit voldoet en boden hun hulp aan bij het zoeken naar oplossingen voor knelpunten.

- VO-Raad, FvOV, CNV onderwijs en AOb lanceerden het digitaal Meldpunt Correctie Centrale Examens. Via deze website konden scholen voorbeelden van goed georganiseerde correctie delen, knelpunten doorgeven en advies krijgen.
- VO-Raad, FvOV, CNV onderwijs en AOb publiceerden de 'Adviestabel opsturen examenwerk naar de eerste en tweede corrector'. Daarin staat per vak per examen aangegeven op welke dag de tweede corrector het examenwerk van de eerste corrector moet ontvangen en op welke dag het werk weer terug moet zijn bij de eerste corrector.
- CvTE heeft de wens van de staatssecretaris gehonoreerd door arbeidsintensieve vakken waar mogelijk en zinvol in de planning naar voren te halen. Zo meldde de staatssecretaris in 2015 naar aanleiding van Kamervragen dat de vwo-examens Nederlands, tehatex en kunst al op de eerste drie dagen van de eerste week waren ingeroosterd en geschiedenis, biologie en aardrijkskunde in de tweede week. In het havo waren Nederlands en biologie in de eerste week ingeroosterd en geschiedenis, filosofie en tehatex in de tweede week.
- Cito heeft gewerkt aan de verbetering van de open vragen en bijbehorende antwoordmodellen. Meer dan voorheen worden - onder strikte geheimhoudingscondities - al tijdens de constructiefase antwoorden van kandidaten verzameld. Het doel is te komen tot antwoordmodellen die de mogelijke (geheel of gedeeltelijk juiste) antwoorden zo goed mogelijk dekken. Door de antwoorden te analyseren krijgt de examenmaker kwalitatieve informatie ter verdere verbetering van het antwoordmodel. Op basis van deze informatie kunnen aan het beoordelingsmodel uitleg en voorbeelden worden toegevoegd voor het beoordelen van veel voorkomende antwoorden die geheel of gedeeltelijk fout zijn. Daarnaast onderzoekt Cito de mogelijkheden om correctie-intensieve open vragen zonder verlies van betrouwbaarheid en validiteit om te werken tot minder bewerkelijke open vragen of gesloten vragen.

In dit rapport wordt verslag gedaan van het nieuwe onderzoek naar de correctiepraktijk dat de staatssecretaris in het vooruitzicht stelde. Een eerste doel is een antwoord te geven op de volgende, onderling samenhangende onderzoeksvragen:

- Hoeveel tijd kosten de eerste en tweede correctie? En in hoeverre ervaren docenten de correctie als belastend?
- Hoe wordt de eerste en tweede correctie op school gefaciliteerd? En in hoeverre zijn docenten tevreden over dit schoolbeleid?
- In hoeverre voeren docenten de tweede correctie integraal uit? Waarom wordt de tweede correctie niet altijd integraal uitgevoerd? En hoe denken docenten over mogelijke oplossingen?
- Wat vinden docenten van de voorgestelde omkering van de correctievolgorde en andere (vermeende) verbeteringen?
- Hoe denken eerste en tweede correctoren over de volledigheid, zorgvuldigheid en objectiviteit van elkaars beoordelingen? En welke maatregelen vinden docenten wenselijk om de kwaliteit van de correctie waar nodig te verbeteren?

Een tweede doel is een antwoord te geven op de vraag in hoeverre de correctiepraktijk veranderd is. Een groot deel van de enquêtevragen is ook in 2011 gesteld. Daardoor kunnen we de gegevens van 2016 vergelijken met die van 2011.

2 Methode van onderzoek

2 Methode van onderzoek

2.1 Populatie, steekproef en respons

In het onderzoek zijn twaalf examenvakken betrokken, te weten Nederlands vwo, biologie vwo, aardrijkskunde vwo, muziek vwo, Duits havo, scheikunde havo, geschiedenis havo, tehatex havo, Engels gt, wiskunde gt, economie gt en beeldende vakken gt. De vakken zijn gelijkmatig verdeeld over de vier typen vakken (talen, bèta, gamma en kunst) en de drie opleidingstypen (vwo, havo en gt). In alle gevallen ging het om examens voor het schooljaar 2015-2016.

Bij de steekproeftrekking is gebruik gemaakt van het populatiebestand van 1042 scholen voor regulier voortgezet onderwijs die zich met een of meer kandidaten voor een of meer van de genoemde examens hadden ingeschreven (exclusief opleidingen voor volwassenen). Voor de twaalf vakken zijn alle ingeschreven vestigingen geselecteerd die zich met drie of meer kandidaten voor het desbetreffende examen hadden ingeschreven. De eerste twee kolommen van tabel 2.1.1 tonen respectievelijk het aantal vestigingen en kandidaten in het populatiebestand. De derde kolom toont het aantal vestigingen in de steekproef.

Tabel 2.1.1 Populatie, steekproef en respons

	Aantal vestigingen populatie	Aantal kandidaten populatie	Aantal vestigingen steekproef	Aantal geretourneerde vragenlijsten	Responspercentage
Nederlands vwo	586	38602	515	299	58
Biologie vwo	555	19741	509	309	61
Aardrijkskunde vwo	565	12635	495	292	59
Muziek vwo	112	1110	101	60	59
Duits havo	578	19347	503	290	58
Scheikunde havo	578	25401	503	337	67
Geschiedenis havo	589	36058	512	297	58
Tehatex havo	234	4052	191	102	53
Engels gt	875	60588	804	405	50
Wiskunde gt	874	53722	804	492	61
Economie gt	855	44914	786	443	56
Beeldende vakken gt	455	9138	425	214	50
Totaal	6856	325308	6148	3540	58

De verzending van de 6148 vragenlijsten (inclusief begeleidende brief) heeft plaatsgevonden op 23 mei 2016. De vragenlijsten zijn verzonden aan alle sectieleiders van de vestigingen die zich met drie of meer kandidaten voor een of meer van twaalf geselecteerde examenvakken hadden ingeschreven voor het komende CSE. De examensecretarissen werden van de verzending op de hoogte gesteld en hen werd gevraagd om het belang van het onderzoek bij de sectieleiders te onderstrepen. De sectieleider is verzocht de vragenlijst te overhandigen aan de docent die de examenklas(sen) van het desbetreffende vak onder zijn of haar hoede had (en die tevens als eerste en/of tweede corrector bij het examen betrokken was).

Ruim twee weken na de verzending bleek 18% van de vragenlijsten geretourneerd te zijn.

Op 13 juni is de non-respondenten een herinneringsbrief gestuurd en ontvingen de examensecretarissen een mail waarin zij van deze verzending op de hoogte gesteld werden.

De sluitingsdatum was 26 juli, enkele dagen nadat de laatste regio in Nederland vakantie kreeg. In totaal zijn er 3540 vragenlijsten (58%) teruggestuurd die bruikbaar waren voor de statistische

analyse en rapportage. Van de 6148 verzonden enveloppen kwamen er 43 ongeopend retour. Houden we rekening met deze opgeheven of bij een fusie betrokken vestigingen, dan bedraagt de respons nog steeds 58% (na afronding op een geheel getal). Na de sluitingsdatum kwamen er nog 61 vragenlijsten binnen die niet meer verwerkt konden worden.

Het behaalde responspercentage van 58% is iets lager dan de 62% uit het vorige onderzoek naar de correctiepraktijk (Kuhlemeier & Kremers, 2013). Een mogelijke oorzaak is dat de vragenlijsten in 2011 na de zomervakantie verstuurd zijn en in 2016 tijdens de drukke periode vlak voor de zomervakantie. Daarnaast is het wat hogere respons destijds behaald na twee keer rappelleren (waarbij de vragenlijsten bij het tweede rappel nogmaals meegestuurd werden), terwijl de krappe planning deze keer slechts één rappel toeliet.

Tijdens het overleg tussen AOb, FvOV, CNV onderwijs, VO-raad en OCW op 28 april 2016 rees de vraag naar het aantal voor deelname benaderde docenten in relatie tot de totale populatie van docenten met examenklassen in Nederland. Ons staan geen gegevens ter beschikking over de precieze aantallen docenten in de onderscheiden examenvakken die lesgeven in het laatste leerjaar van vmbo, havo en vwo. Wel beschikken we over gegevens over de totale populatie van docenten in het voortgezet onderwijs. Op basis daarvan kunnen we met enkele slagen om de arm een grove raming van de omvang van de doelpopulatie opstellen. In 2014 waren er 104.700 personen werkzaam in het voortgezet onderwijs waarvan 71,7% onderwijzend personeel (bron: www.stamos.nl). Ons onderzoek is gericht op docenten van vwo, havo en vmbo gt. Docenten vmbo bb en kb zijn niet in dit onderzoek betrokken omdat de examens binnen deze schooltypen voornamelijk digitaal worden afgenomen en de problematiek van de correctie daar volledig anders is. Uit de gegevens blijkt dat er (afgerond) 56.000 docenten lesgeven binnen de schooltypen vwo, havo en vmbo gt. De schatting is dat ongeveer de helft van de docenten een examenklas heeft, waarmee het aantal docenten met een examenklas komt op ongeveer 28.000. Met een aantal slagen om de arm is de voorlopige conclusie dat 22% van de doelpopulatie van docenten (= $6148 : 28000 * 100$) voor deelname benaderd is en dat bijna 13% (= $3540 : 28000 * 100$) daadwerkelijk heeft meegedaan.

Deze conclusie staat of valt met de juistheid van het aantal van 28.000 docenten met een examenklas. Ter controle hebben we het meest recente bestand van Wolf geraadpleegd. Daaruit blijkt dat in totaal 29.323 docenten scores van kandidaten hebben ingevoerd, waarvan 9620 vwo-docenten, 9905 havo-docenten en 9798 vmbo-docenten. Dit is meer dan de 28.000 docenten waar wij van uitgegaan zijn bij de berekening van de responspercentages. Een waarschijnlijk klein deel van de 29.323 docenten heeft echter examenklassen van verschillende opleidingstypen onder zijn of haar hoede (bijvoorbeeld een vwo- en havo examenklas Nederlands) of heeft behalve een wiskunde B examenklas ook een wiskunde A examenklas. Houden we hiermee rekening, dan lijkt het aantal van 28.000 docenten met een examenklas niet ver van de waarheid.

2.2 vragenlijst

De vragenlijst bestaat uit 34 vragen waarvan 11 in rubriek A over de eerste correctie, 19 in rubriek B over de tweede correctie en 4 in rubriek C over meer algemene aspecten van de correctiepraktijk. In totaal moest de respondent 72 keer een alternatief aankruisen of een getal invullen, waarvan 16 keer in rubriek A, 34 keer in rubriek B en 22 keer in rubriek C. Bij 9 vragen bestond de mogelijkheid om het gegeven antwoord nader toe te lichten.

Rubriek A bevat 11 vragen die de respondent in de hoedanigheid van eerste corrector moest beantwoorden. De vragen betroffen de volgende onderwerpen:

- het aantal beoordeelde kandidaten, de tijdbesteding en de ervaren belasting van de eerste correctie;
- de wijze waarop de eerste correctie op school gefaciliteerd wordt;
- de tevredenheid over het schoolbeleid met betrekking tot de eerste correctie;
- het oordeel (van de eerste corrector) over de volledigheid, zorgvuldigheid en soepelheid waarmee de tweede corrector heeft nagekeken.

Rubriek B bevat 19 vragen die de respondent in de hoedanigheid van tweede corrector moest beantwoorden. Een deel van de vragen van rubriek B is identiek aan die uit rubriek A, zij het dat ze nu gesteld zijn vanuit het perspectief van de tweede corrector. De vragen betroffen de volgende onderwerpen:

- het aantal beoordeelde kandidaten, de tijdbesteding en de ervaren belasting van de tweede correctie;
- de wijze waarop de tweede correctie op school gefaciliteerd wordt;
- de tevredenheid over het schoolbeleid met betrekking tot de tweede correctie;
- het oordeel (van de tweede corrector) over de volledigheid, zorgvuldigheid en soepelheid waarmee de eerste corrector heeft nagekeken;
- de volledigheid van de tweede correctie.

Rubriek C bevat vier samengestelde matrixvragen over meer algemene aspecten van de correctiepraktijk. De vragen betreffen de mening van de respondenten over volgende onderwerpen:

- de redenen waarom docenten de tweede correctie niet volledig (kunnen) uitvoeren;
- de maatregelen die nodig zijn om te bevorderen dat alle docenten de tweede correctie integraal (kunnen) uitvoeren;
- de wenselijkheid van bepaalde veranderingen van of aanvullingen op de huidige correctieprocedures.

Een groot deel van de vragen is in (vrijwel) dezelfde formulering ook in 2011 gesteld. Hierdoor kan de correctiepraktijk van 2016 worden vergeleken met die van 2011.

2.3 Codering van de antwoorden op de open vragen

De vragenlijst bevat acht vragen waarbij respondenten hun antwoord desgewenst nader konden toelichten. Een groot aantal respondenten heeft hiervan gebruik gemaakt. Hun antwoorden zijn geanalyseerd met behulp van de methode van inhoudsanalyse (Krippendorff, 1980; Weber, 1990). Voor elke vraag is er een afzonderlijk categorieënsysteem ontwikkeld met behulp waarvan de antwoorden geclassificeerd zijn. Daarbij zijn we als volgt te werk gegaan. Allereerst heeft de eerste auteur van dit rapport - na bestudering van (per vraag) ongeveer honderd antwoorden - een eerste versie van het categorieënsysteem ontwikkeld. Dat bevatte ook regels voor het coderen van de antwoorden en illustratieve voorbeelden van 'echte' toelichtingen van respondenten. Vervolgens hebben twee collega's het categorieënsysteem uitprobeerd op een selectie van ongeveer tweehonderd antwoorden en aan de hand van de resultaten bijgesteld. Daarna konden de beide collega's beginnen met daadwerkelijke classificeren van de antwoorden. Op het moment van het schrijven van dit rapport was het coderen van de antwoorden voltooid. Voor het uitgebreide verslag van de inhoudsanalyse wordt verwezen naar Loeffen en Eikelboom (in voorbereiding). In dit rapport volstaan we met het beschrijven van de belangrijkste resultaten.

2.4 Statistische analyse

De gegevens zijn geanalyseerd met behulp van het statistische pakket SPSS 20.0.2. Voor de analyse zijn, afhankelijk van het meetniveau van de variabele, verschillende analysetechnieken gebruikt.

Bij de variabelen op intervalniveau zijn de verschillen tussen jaren (2016 versus 2011) en tussen examens geanalyseerd met univariate variantieanalyse (UVA) of multivariate variantieanalyse (MVA) voor herhaalde metingen. MVA voor herhaalde metingen is toegepast als de dataset per respondent twee metingen bevatte, zoals het geval is bij de tijdbesteding aan en de zorgvuldigheid en soepelheid van de eerste en tweede correctie. In de analyse is de correctieprocedure (eerste versus tweede correctie) opgevat als een binnen-subjecten factor en het jaar en het examen als tussen-subjecten factoren. Hiermee kunnen zowel de hoofdeffecten van correctieprocedure, jaar en examen worden geschat als de interacties tussen deze drie variabelen. We lichten de begrippen hoofd- en interactie-effect toe aan de hand van de onafhankelijke variabele jaar. Een significant hoofdeffect van jaar betekent dat het gemiddelde voor bijvoorbeeld de tijdbesteding aan de correctie in 2016 verschilt van 2011. Een significant interactie-effect betekent dat het verschil tussen beide jaren in bijvoorbeeld de tijdbesteding verschilt van examen tot examen en/of voor de eerste correctie anders is dan voor de tweede correctie.

Naast variabelen op intervalniveau bevat de vragenlijst een aantal ordinale variabelen, zoals de tevredenheid met het schoolbeleid. De verschillen tussen jaren (2016 versus 2011), examens en tussen de eerste en tweede correctie zijn hier geanalyseerd met ordinale logistische regressieanalyse (OLR; Agresti, 2002). De analyse is uitgevoerd met de zogeheten *Polytomous Logit Universal Model* (PLUM; McCullagh, 1980), onderdeel van het statistische pakket SPSS 20.0.2. Ordinale logistische regressie is een geëigende techniek voor de analyse van de invloed van nominale onafhankelijke variabelen op een afhankelijke variabele met verschillende ordinale categorieën zoals nooit, in een enkel geval, regelmatig, vaak en zeer vaak. PLUM voorziet in regressiegewichten voor de verschillende niveaus (*thresholds*) van de afhankelijke variabele en de effecten van de onafhankelijke variabelen. De *'goodness-of-fit'*-test geeft aan hoe goed of slecht het model bij de gegevens past. De Wald-test laat zien in hoeverre een regressiegewicht significant afwijkt van nul of van een ander regressiegewicht.

In rubriek A was ook een zogenoemde doorverwijsvraag opgenomen. Hierbij moest de respondent Nee aankruisen als hij of zij geen eerste corrector was en Ja als dat wel het geval was. Bij het schonen van de gegevens van rubriek A bleek deze doorverwijsvraag niet volledig naar behoren gewerkt te hebben. Hierna leggen we uit hoe we daar in de analyse mee omgegaan zijn. De eerste vraag was een zogeheten doorverwijsvraag waarbij de respondent Nee moest aankruisen als hij of zij geen eerste corrector was en Ja als dat wel het geval was. Degenen die naar eigen zeggen geen eerste corrector waren, is gevraagd de overige vragen van rubriek A over te slaan en direct verder te gaan met het beantwoorden van de vragen van rubriek B over de praktijk van de tweede correctie. In de statistische analyse zijn respondenten die de eerste vraag van rubriek A oversloegen of daar Nee aankruisten maar toch een of meer van de overige vragen over de praktijk van de eerste correctie beantwoordden als eerste correctoren beschouwd (door hun antwoord op de vraag of zij eerste corrector waren van Nee naar Ja te hercoderen). Tegelijkertijd zijn respondenten die deze eerste vraag met Ja beantwoordden maar desondanks geen van de overige vragen over de praktijk van de eerste correctie beantwoordden niet als eerste corrector beschouwd (door hun antwoord op de eerste vraag van Ja naar Nee te hercoderen). Zie ook verderop in paragraaf 2.6.

2.5 Representativiteit

De relatief hoge respons van 58% betekent niet noodzakelijkerwijs dat de gegevens representatief zijn voor de examenpraktijk in Nederland. Zo zouden docenten die de tweede correctie niet volledig volgens de wettelijke voorschriften uitvoeren in het onderzoek ondervertegenwoordigd kunnen zijn. Een veronderstelling is dat respondenten die de vragenlijst onmiddellijk beantwoorden en retourneren minder geneigd zijn tot het overtreden van de wettelijke voorschriften met betrekking tot de correctie dan degenen die later reageren. Ter controle is nagegaan in hoeverre er een relatie is tussen de responssnelheid en de antwoorden op de vragen. Als indicator voor de responssnelheid is de week van binnenkomst gebruikt. De correlaties tussen de week van binnenkomst en de antwoorden op 82 variabelen over de correctiepraktijk varieerden van -0,09 tot 0,10 met een mediaan van 0,00 en een standaarddeviatie van 0,04. Er zijn dus geen aanwijzingen dat regelgetrouwe docenten in de respons over- of ondervertegenwoordigd zijn. De conclusie is dat niet kan worden aangetoond dat de respons niet representatief is voor de populatie van docenten in het voortgezet onderwijs.

Wellicht ten overvloede zijn in tabel 2.5.1 de vragen met de grootste negatieve en grootste positieve correlaties met de responssnelheid weergegeven. De resultaten zijn behoudens één uitzondering strijdig met de verwachting dat latere respondenten minder regelgetrouw zijn. De grootste negatieve correlaties beschrijven we als volgt. Hoe later de vragenlijst geretourneerd wordt, hoe minder de respondenten van mening zijn dat een (deel van de respondenten de tweede correctie niet volledig uitvoert omdat een) steekproef voldoende garantie biedt, hoe vaker men alsnog alles volledig nakijkt, hoe meer men zich kan verenigen met de uitspraak 'Ik heb het werk van alle kandidaten volledig nagekeken' en hoe minder kandidaten men bij de tweede correctie heeft nagekeken. De grootste positieve correlaties kunnen als volgt worden beschreven. Hoe later de vragenlijst is teruggestuurd, hoe vaker de respondenten alles alsnog volledig nakeken, hoe sterker men van mening is dat de tweede correctie integraal moet worden uitgevoerd, hoe meer kandidaten er beoordeeld zijn en hoe meer tijd men in de correctie investeerde.

Tabel 2.5.1 Uitspraken met de grootste negatieve en grootste positieve correlatie met de responssnelheid

	Correlatie
De grootste negatieve correlaties (< -0,05)	
Een steekproef van examenvragen biedt voldoende garantie (0 = Nee; 1 = Ja)	-0,09
Een steekproef van kandidaten biedt voldoende garantie (0 = Nee; 1 = Ja)	-0,07
Als de tweede corrector een deel heeft nagekeken en tot de conclusie komt dat de eerste corrector goed heeft nagekeken, is het niet nodig om alles na te kijken (1 = Helemaal mee oneens; 4 = Helemaal mee eens)	-0,07
Ik heb het examenwerk van alle kandidaten volledig nagekeken (1 = Klopt helemaal; 4 = Klopt helemaal niet)	-0,07
Aantal nagekeken kandidaten tweede correctie	-0,06
De grootste positieve correlaties (> 0,05)	
Ik heb eerst een deel van de werken/vragen nagekeken en daarna heb ik alsnog alle werken/vragen volledig nagekeken (omdat de resultaten van het eerste deel daartoe aanleiding gaven) (1 = Klopt helemaal; 4 = Klopt helemaal niet)	0,06
De tweede correctie moet integraal worden uitgevoerd (d.w.z.: alle examenwerken en per kandidaat alle vragen)? (1 = Helemaal mee oneens; 4 = Helemaal mee eens)	0,06
Aantal uren per dag	0,07
Aantal minuten per kandidaat tweede correctie	0,08
Aantal minuten per kandidaat eerste correctie	0,09
Aantal minuten per kandidaat eerste en tweede correctie gezamenlijk	0,10

2.6 Interpretatie van de resultaten

De interpretatie van de antwoorden over de praktijk van de eerste correctie uit rubriek A wordt enigszins bemoeilijkt door een fout in de formulering van de eerste twee vragen van de vragenlijst. De vragenlijsten en de begeleidende brief waren geadresseerd aan de docent die in het laatste leerjaar les gaf aan de kandidaten van het desbetreffende examen. Het was de bedoeling dat de respondent de vragen over de correctiepraktijk voor het huidige schooljaar zou beantwoorden. In rubriek A van de vragenlijst is de respondenten echter abusievelijk gevraagd of zij **vorig jaar** eerste corrector waren in plaats van dit jaar. De mogelijkheid bestaat dat deze formuleringsfout invloed heeft gehad op de antwoorden op de vragen over de praktijk van de eerste correctie uit deel 1 van de vragenlijst. In vergelijking met het vorige onderzoek uit 2011 bleken namelijk minder respondenten eerste corrector te zijn (respectievelijk 96% versus bijna 100%). Dit verschil is begrijpelijk als men bedenkt dat de sectieleider verzocht is de vragenlijst te overhandigen aan de docent die vorig jaar in het laatste leerjaar les gaf en een deel van deze docenten in 2016 geen eerste corrector was (en daardoor de eerste vraag waarheidsgetrouw met Nee beantwoordde). Als gevolg hiervan ontvingen wij in de eerste twee weken na de verzending regelmatig mails waarin docenten ons vroegen of zij de vragen van rubriek A voor het vorige of voor het huidige schooljaar moesten beantwoorden.

Er zijn verschillende maatregelen genomen om de invloed van deze fout te minimaliseren. Allereerst hebben wij de examensecretarissen reeds op de dag na de verzending van de vragenlijst attent gemaakt op de formuleringsfout en hun gevraagd de docenten hiervan op de hoogte te stellen. Drie weken later hebben wij alle 6148 voor deelname benaderde docenten een brief gestuurd met het verzoek de vragen over de praktijk van de eerste en tweede correctie te beantwoorden voor het huidige in plaats van het vorige schooljaar.

Ter bestudering van de eventuele invloed van de fout op de antwoorden zijn wij nagegaan in hoeverre de responsnelheid samenhangt met het percentage docenten dat de eerste vraag oversloeg en het percentage docenten dat deze vraag met Nee beantwoordde (en dus dit jaar geen eerste corrector was). De veronderstelling was dat deze percentages vóór de verzending van de brief waarin de non-respondenten op de fout geattendeerd werden hoger zouden zijn dan in de weken daarna. Een uitgevoerde significantietoetsing laat echter zien dat het veronderstelde effect van de responsnelheid op het percentage ontbrekende waarnemingen en het percentage docenten dat geen eerste corrector was niet aantoonbaar is. In de eerste drie weken na de verzending hebben niet meer docenten de eerste vraag overgeslagen of aangegeven geen eerste corrector te zijn dan in de weken daarna ($p = 0,461$). Evenzo hebben vroege retourneerders niet meer vragen van rubriek A overgeslagen dan latere retourneerders ($p = 0,199$). Als de formuleringsfout een effect gehad heeft op de resultaten, dan doet de vertekening zich bij alle weken in ongeveer dezelfde mate voor. Daar komt bij dat op het moment van de verzending van de brief waarin de non-respondenten op de fout geattendeerd werden nog slechts 18% van de benaderde docenten de vragenlijst geretourneerd had. De conclusie is dat de formuleringsfout bij een deel van de vroege retourneerders voor enige verwarring heeft gezorgd, maar dat de validiteit van de resultaten daardoor niet wezenlijk is aangetast. Uiteraard zijn hiermee niet alle interpretatieproblemen uit de wereld. Uit de gegevens blijkt dat een groot deel van de respondenten die aangaven dit schooljaar geen eerste corrector te zijn toch de overige vragen van rubriek A heeft ingevuld. Vooral bij de vroege retourneerders onder hen is niet duidelijk of zij de vragen van rubriek A voor het huidige dan wel het vorige schooljaar beantwoordden. De vraag is echter hoe erg dit laatste is als we bedenken dat de correctiepraktijk in 2016 voor de meeste vakken sterk zal lijken op die van het vorige schooljaar.

3 Resultaten

3 Resultaten

3.1 Tijdbesteding en ervaren belasting

In deze paragraaf besteden we aandacht aan de tijd die docenten besteden aan de eerste correctie, de tweede correctie en de eerste en tweede correctie samen. We rapporteren over de verschillen tussen examens en vergelijken de tijdbesteding in 2016 met die in 2011. En we gaan na hoe belastend de docenten de correctie ervaren in 2016 en 2011.

Dit hoofdstuk bevat een aantal tabellen met percentages optellend tot 100%. We maken de lezer erop attent dat dit laatste niet altijd het geval is. Dat de kolompercentages soms tot 99% of 101% optellen, is een gevolg van afronding op een geheel getal (en niet van onvolkomenheden in de data).

3.1.1 Tijdbesteding aan de eerste correctie

De respondenten is gevraagd of zij het examen als eerste corrector hebben nagekeken. In 2016 was 96% van de respondenten eerste corrector en in 2011 was dat bijna 100% (voor een mogelijke verklaring van dit verschil zie paragraaf 2.6). Tabel 3.1.1 geeft informatie over de correctielast en de tijdbesteding aan de eerste correctie in 2016 en 2011. Daarbij staat N voor het aantal docenten, Gem voor het gemiddelde, Se-gem voor de standaardfout van het gemiddelde en Stddev voor de standaarddeviatie. De percentielpunten P10, P25, P75 en P90 symboliseren het percentage respondenten met een even hoog of lager aantal kandidaten, uren, minuten of dagen.

Tabel 3.1.1 Correctielast en tijdbesteding aan de eerste correctie per jaar

	Jaar	N	Min	Max	Gem	Se-Gem	Stddev	P10	P25	P75	P90
Aantal nagekeken kandidaten	2016	3378	1	228	38,84	0,40	23,32	15,00	23,00	55,00	70,00
	2011	3637	1	240	38,45	0,40	23,87	14,00	22,00	51,00	70,00
Aantal uren	2016	3364	1	260	21,32	0,32	18,54	5,00	10,00	28,00	40,00
	2011	3597	1	160	22,27	0,32	19,37	6,00	10,00	30,00	45,00
Aantal minuten per kandidaat	2016	3361	1	300	38,22	0,52	30,14	10,67	18,00	48,39	72,00
	2011	3591	1	360	39,87	0,52	31,41	12,00	20,00	50,53	75,73
Aantal dagen	2016	3387	1	20	4,48	0,04	2,42	2,00	3,00	5,00	8,00
	2011	3616	0	20	4,29	0,04	2,45	2,00	3,00	5,00	7,00

In 2016 hebben de eerste correctoren (uit het onderzoek) het werk van 39 examenkandidaten nagekeken en daar in totaal 21 klokuren aan besteed in ruim 4 dagen. De gemiddelde correctietijd per kandidaat was 38 minuten.

Verschillen tussen docenten in 2016

De correctielast en de tijdbesteding aan de eerste correctie blijken sterk van docent tot docent te verschillen. De percentielscores geven het percentage respondenten met een even hoog of lager aantal kandidaten, uren, minuten of dagen. De 10% met de minste correctielast (P10) keek het werk van hooguit 15 kandidaten na en besteedde daar tot 5 klokuren aan in hooguit

2 dagen. De 10% met de zwaarste correctielast (P90) keek het werk van minimaal 70 kandidaten na en besteedde daar minimaal 40 uur aan in 8 dagen of meer.

Verschillen tussen 2016 en 2011

De verschillen tussen jaren en examens zijn op significantie getoetst met behulp van univariate variantieanalyse. De toetsing is voor elk aspect van de tijdbesteding afzonderlijk uitgevoerd (dat wil zeggen aantal nagekeken kandidaten, aantal uren, aantal minuten per kandidaat en aantal dagen). De afhankelijke variabele was de tijdbesteding aan de eerste correctie en de onafhankelijke variabelen jaar en examens. Vanwege het zeer grote aantal waarnemingen is het 1%-significantieniveau als criterium gehanteerd. Het hoofdeffect van jaar geeft bij één van de vier aspecten van de tijdbesteding significantie te zien, maar het hoofdeffect van examens en de interactie van jaar en examens waren telkens wel significant ($p < 0,01$). In 2016 corrigeerden de eerste correctoren evenveel kandidaten als in 2011. Bovendien besteedden zij evenveel uren aan de eerste correctie als in 2011. De gemiddelde correctietijd per kandidaat was in beide jaren gelijk. Wel was de periode waarbinnen de eerste correctie werd uitgevoerd in 2016 wat langer dan in 2011: 4,48 dagen in 2016 versus 4,29 dagen in 2011.

Verschillen tussen examens

De significante interactie van jaar met examens betekent dat het verschil in tijdbesteding tussen 2016 en 2011 niet voor elk examen gelijk is. Tabel 3.1.2 laat zien in welke mate de tijdbesteding aan de eerste correctie verschilt van examen tot examen en van jaar tot jaar. Te zien is onder meer dat docenten Nederlands in 2016 de meeste kandidaten beoordeelden, maar dat geschiedenisdocenten de meeste uren aan de eerste correctie besteedden.

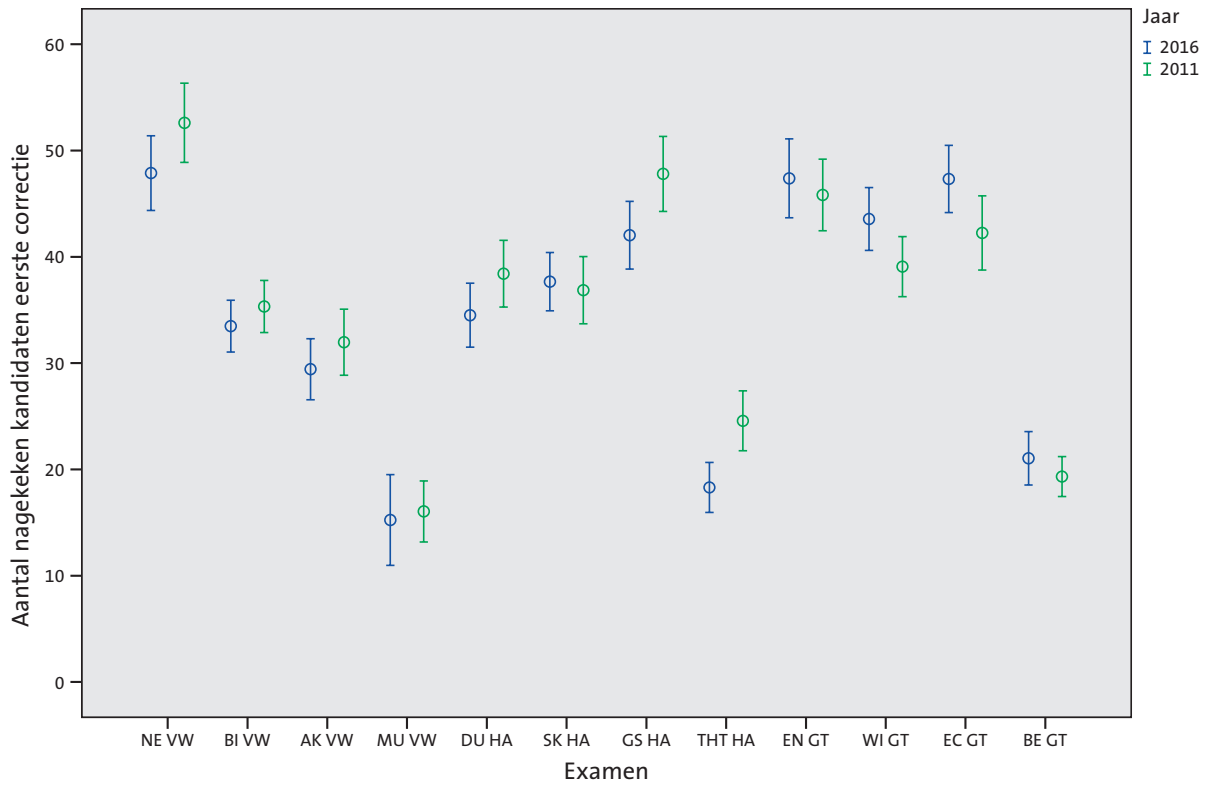
Tabel 3.1.2 Correctielast en tijdbesteding aan de eerste correctie per jaar per examen

	Jaar	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Aantal nagekeken kandidaten	2016	48	33	29	15	35	38	42	18	47	44	47	21
	2011	53	35	32	16	38	37	48	25	46	39	42	19
Aantal uren	2016	35	23	24	12	9	25	36	21	13	16	21	16
	2011	32	25	25	15	13	25	40	23	15	16	19	15
Aantal minuten per kandidaat	2016	47	45	56	57	17	42	57	75	18	25	29	52
	2011	39	47	56	61	22	44	52	61	22	28	31	54
Aantal dagen	2016	7	5	5	3	3	5	7	4	3	4	4	4
	2011	6	5	5	3	3	4	6	4	3	4	4	3

Verschillen tussen jaren per examen

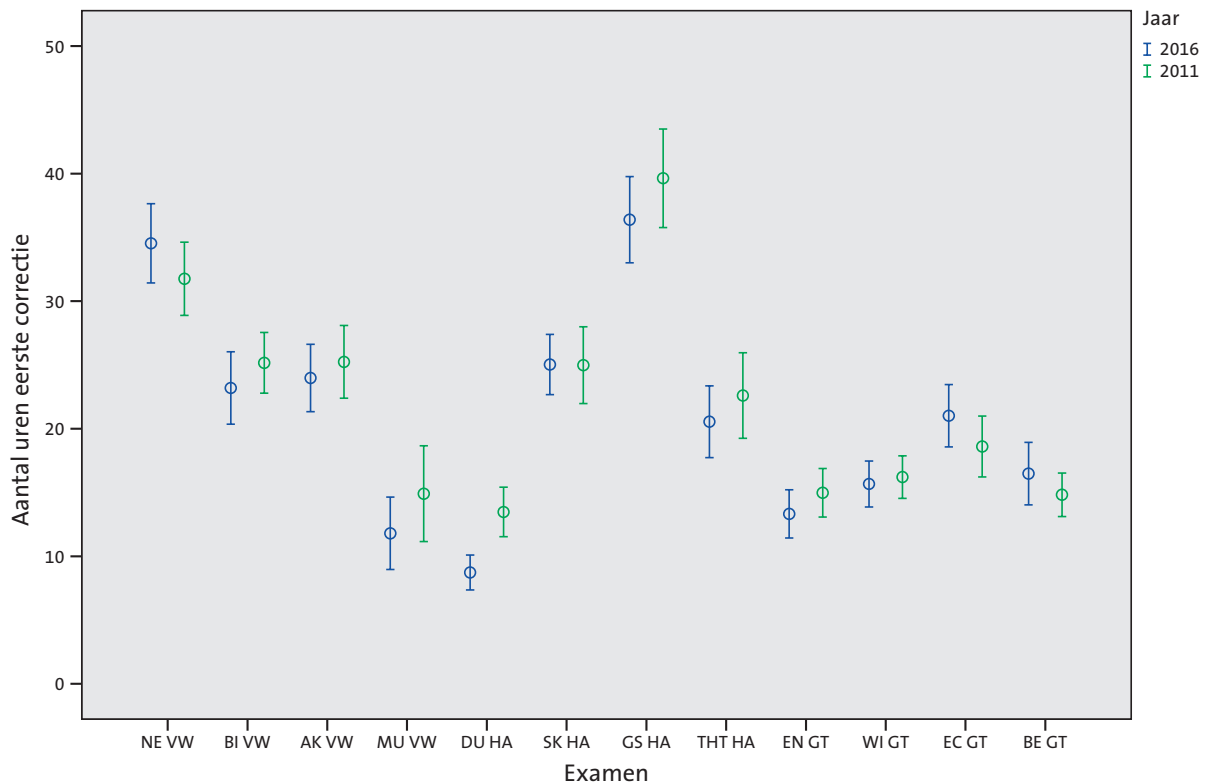
Hiervoor constateerden we dat het verschil tussen de tijdbesteding aan de eerste correctie in 2016 en 2011 niet onafhankelijk is van het specifieke examen. Figuur 1 en 2 tonen voor elk examen en jaar respectievelijk het gemiddeld aantal beoordeelde kandidaten en de gemiddelde correctietijd in uren (voor de eerste correctie). De gemiddelden zijn aangeduid met een cirkel. De verticale lijnen rond het gemiddelde symboliseren het 99%-betrouwbaarheidsinterval. Dat geeft aan binnen welke grenzen het gemiddelde met 99% zekerheid ligt (dit wil zeggen dat bij herhaalde steekproeftrekking uit dezelfde populatie 99 van de 100 gevonden waarden in dit interval terug te vinden zouden zijn). Als de betrouwbaarheidsintervallen van twee gemiddelde elkaar niet overlappen, mag worden aangenomen dat de gemiddelden verschillend zijn op 1%-niveau.

Inspectie van de correctielast in figuur 1 laat zien dat de verschillen tussen 2016 en 2011 bij de meeste examens vrij klein zijn. Een duidelijke uitzondering lijkt tehatex waar docenten in 2016 minder kandidaten moesten beoordelen dan in 2011.



Figuur 1 Gemiddeld aantal nagekeken kandidaten voor de eerste correctie per examen per jaar

Vergelijken we in figuur 2 het aantal uren in 2016 en 2011 per examen, dan lijkt er alleen bij Duits sprake van een betekenisvol verschil. Docenten Duits hadden in 2016 gemiddeld vier uur minder voor de eerste correctie nodig dan in 2011.



Figuur 2 Gemiddelde aantal uren voor de eerste correctie per examen per jaar

3.1.2 Tijdbesteding aan de tweede correctie

Wel of geen tweede corrector

De respondenten is gevraagd of ze het werk van een andere school als tweede corrector hadden nagekeken. In 2016 zegt 95% dat gedaan te hebben en in 2011 was dat eveneens 95%.

Verschillen tussen examens in 2016

De toetsing van de verschillen tussen examens geeft significantie te zien ($p < 0,001$). Zo zijn docenten geschiedenis vaker (ook) als tweede corrector bij hun examens betrokken dan docenten muziek (zie tabel 3.1.3).

Tabel 3.1.3 Betrokkenheid als tweede corrector per examen per jaar (percentages)

Jaar	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
2016	98	97	95	84	95	97	99	91	93	94	94	90
2011	98	97	95	95	97	97	97	94	95	95	95	85

Redenen waarom docenten geen tweede corrector waren

De 181 docenten die aangaven in 2016 geen tweede corrector te zijn, is gevraagd waarom zij het werk van de andere school niet hebben nagekeken. Men kon daarbij kiezen voor een of meer van de antwoordmogelijkheden die zijn weergegeven in tabel 3.1.4. De redenen 'tijdgebrek', 'geringe vergoeding' of 'vervanging door een collega' worden net als in 2011 nauwelijks genoemd.

Tabel 3.1.4 Redenen waarom respondenten geen tweede corrector waren per jaar

Reden	Jaar	N	%
Vanwege tijdgebrek	2016	181	1
	2011	167	1
Vanwege de geringe vergoeding	2016	181	0
	2011	167	0
Omdat een collega de tweede correctie van mij heeft overgenomen	2016	181	4
	2011	167	4

Toelichtingen van docenten

Van de 181 docenten die geen tweede corrector waren, hebben er 98 (54%) de moeite genomen om te motiveren waarom zij dat niet waren. De toelichtingen zijn gecodeerd met behulp van een categorieënsysteem. Voor het uitgebreide verslag van de inhoudsanalyse wordt verwezen naar Loeffen en Eikelboom (in voorbereiding). De resultaten van de analyse zijn weergegeven in tabel 3.1.5. In deze tabel staan zes redenen waarom de docent niet bij de tweede correctie betrokken was.

Tabel 3.1.5 Redenen waarom docenten geen tweede corrector waren

	%
Docent was geen tweede corrector/was niet ingedeeld (verder ongespecificeerd)	37
De school had de docent niet ingedeeld	8
DUO had de school niet gekoppeld aan een andere school	33
De docent had geen werk van de eerste corrector ontvangen	9
De docent nam deel aan de pilot omkering	1
Persoonlijke omstandigheden van de docent	5
Weet niet/geen mening	3
Overig on-topic	4
Totaal	100

Bij vijf van deze zes redenen lijkt het onwaarschijnlijk dat de tweede correctie niet toch door een andere docent is uitgevoerd. Als een docent geen tweede corrector was omdat hij of zij niet was ingedeeld - hetzij door de school hetzij door DUO - is het goed mogelijk dat een andere docent de tweede correctie heeft uitgevoerd. Ook in het geval van de docent die deelnam aan de pilot omkering van de VO-raad is het vrijwel zeker dat de tweede correctie door een andere docent - de gecommitteerde - is uitgevoerd (Oberon, 2014). Van de vijf docenten die vanwege persoonlijke omstandigheden (sterfgeval, zwangerschapsverlof, ouderschapsverlof) afhaakten, geeft er een aan dat de school voor haar een vervanger had gevonden. Slechts bij één van deze vijf docenten kan met een grote mate van zekerheid vermoed worden dat de tweede correctie niet is uitgevoerd. De desbetreffende docent motiveerde dat als volgt: 'Ik had al 19 dagen achter elkaar gewerkt (2 weekenden + 15 werkdagen) en had niet de energie om er nog 7 dagen aan vast te plakken'.

Bij één van de zes redenen zou het kunnen dat de tweede correctie daadwerkelijk achterwege gebleven is. Het betreft de tweede correctoren die aangaven geen werk van de eerste corrector te hebben ontvangen. Mogelijk heeft de eerste corrector het werk wel nagekeken maar niet aan de tweede corrector opgestuurd.

Al met al bevestigen de toelichtingen dat het slechts zeer sporadisch voorkomt dat de tweede correctie helemaal niet plaatsvindt.

Tijdbesteding aan de tweede correctie

De verdeling van de correctielast en de tijdbesteding aan de tweede correctie is weergegeven in tabel 3.1.6. In 2016 heeft de gemiddelde tweede corrector het werk van 40 examenkandidaten nagekeken en daar in totaal 13 klokuren aan besteed in ruim 3 dagen. De gemiddelde tijdsbesteding aan de tweede correctie per kandidaat is 23 minuten.

Tabel 3.1.6 Correctielast en tijdbesteding aan de tweede correctie per jaar

	Jaar	N	Min	Max	Gem	Se-Gem	Stddev	P10	P25	P75	P90
Aantal nagekeken kandidaten	2016	3330	1	200	39,81	0,39	22,42	17,00	25,00	55,00	70,00
	2011	3442	1	200	40,02	0,39	23,08	17,00	24,00	50,00	70,00
Aantal uren	2016	3313	1	105	13,15	0,20	11,33	4,00	6,00	16,00	25,00
	2011	3412	1	100	11,97	0,18	10,75	3,00	5,00	15,00	25,00
Aantal minuten per kandidaat	2016	3305	1	200	22,70	0,32	18,63	6,76	10,91	30,00	43,16
	2011	3401	1	150	20,77	0,31	18,22	6,00	9,23	26,67	40,00
Aantal dagen	2016	3319	1	24	3,16	0,03	1,82	1,00	2,00	4,00	5,00
	2011	3430	1	14	2,77	0,03	1,56	1,00	2,00	3,00	5,00

Verschillen tussen docenten in 2016

De correctielast en tijdbesteding aan de tweede correctie in 2016 verschilt sterk van docent tot docent. De percentielscores geven het percentage respondenten met een even hoog of lager aantal kandidaten, uren, minuten of dagen. De 10% met de minste correctielast (P10) heeft het werk van hooguit 17 kandidaten nagekeken en besteedde daar tot 4 klokuren aan in hooguit 1 dag. De 10% met de zwaarste correctielast (P90) keek het werk van minimaal 70 kandidaten na en besteedde daar minimaal 25 uur aan in minimaal 5 dagen.

Verschillen tussen 2016 en 2011

De verschillen tussen jaren (2016 versus 2011) en examens zijn op significantie getoetst met behulp van univariate variantieanalyse. De afhankelijke variabele was de tijdbesteding aan de tweede correctie en de onafhankelijke variabelen waren jaar en examen. Het hoofdeffect van jaar geeft bij twee van de vier aspecten van de tijdbesteding significantie te zien, maar het hoofdeffect van examen en de interactie van jaar en examen waren telkens wel significant ($p < 0,01$). In 2016 beoordeelden de tweede correctoren evenveel kandidaten als in 2011 (40 versus 40) in evenveel uren (13 versus 12 uur), maar de gemiddelde correctietijd per kandidaat (23 versus 21 minuten) en het aantal dagen (3,16 versus 2,77 dagen) waren in 2016 significant hoger dan in 2011.

Deze algemene verschillen tussen beide jaren zijn evenwel niet eenduidig te interpreteren. De significante interactie tussen jaar en examen betekent immers dat het verschil tussen de tijdbesteding in 2016 en 2011 voor het ene examen anders is dan voor het andere examen (zie tabel 3.1.7).

Verschillen tussen examens

Het hoofdeffect van examen en de interactie tussen jaar en examen waren bij alle vier aspecten van de tijdbesteding significant ($p < 0,01$). Dit betekent dat er verschillen tussen examens zijn maar dat het verschil niet voor elk jaar hetzelfde is. Tabel 3.1.7 laat zien in welke mate de tijdbesteding aan de tweede correctie verschilt van examen tot examen en van jaar tot jaar.

Tabel 3.1.7 Correctielast en tijdbesteding aan de tweede correctie per jaar per examen

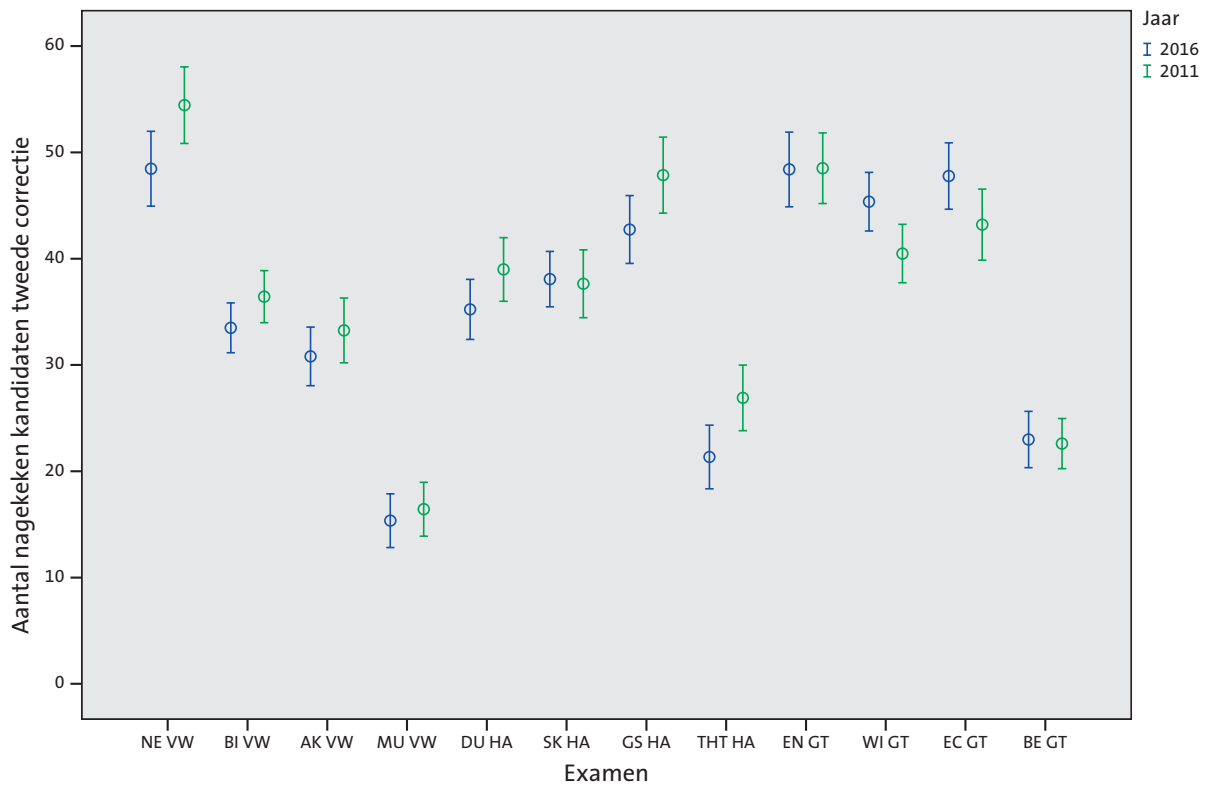
	Jaar	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	ENG T	WI GT	EC GT	BE GT
Aantal nagekeken kandidaten	2016	48	33	31	15	35	38	43	21	48	45	48	23
	2011	54	36	33	16	39	38	48	27	49	40	43	23
Aantal uren	2016	18	12	15	8	6	14	20	14	9	13	13	12
	2011	15	13	12	9	8	12	17	14	10	10	12	11
Aantal minuten per kandidaat	2016	24	24	31	33	11	24	31	49	13	18	18	33
	2011	18	23	25	36	13	22	23	34	14	17	19	32
Aantal dagen	2016	4	3	3	2	2	3	4	4	3	3	3	3
	2011	4	3	3	2	2	3	3	3	2	3	3	3

De verschillen tussen de examens zijn groot. In 2016 beoordeelden docenten Nederlands, Engels en economie bijvoorbeeld meer kandidaten dan docenten muziek, besteedden docenten geschiedenis veel meer uren aan de correctie dan docenten Duits en Engels en was de gemiddelde correctietijd per kandidaat bij tehatex veel groter dan bij Duits.

Verschillen tussen jaren per examen

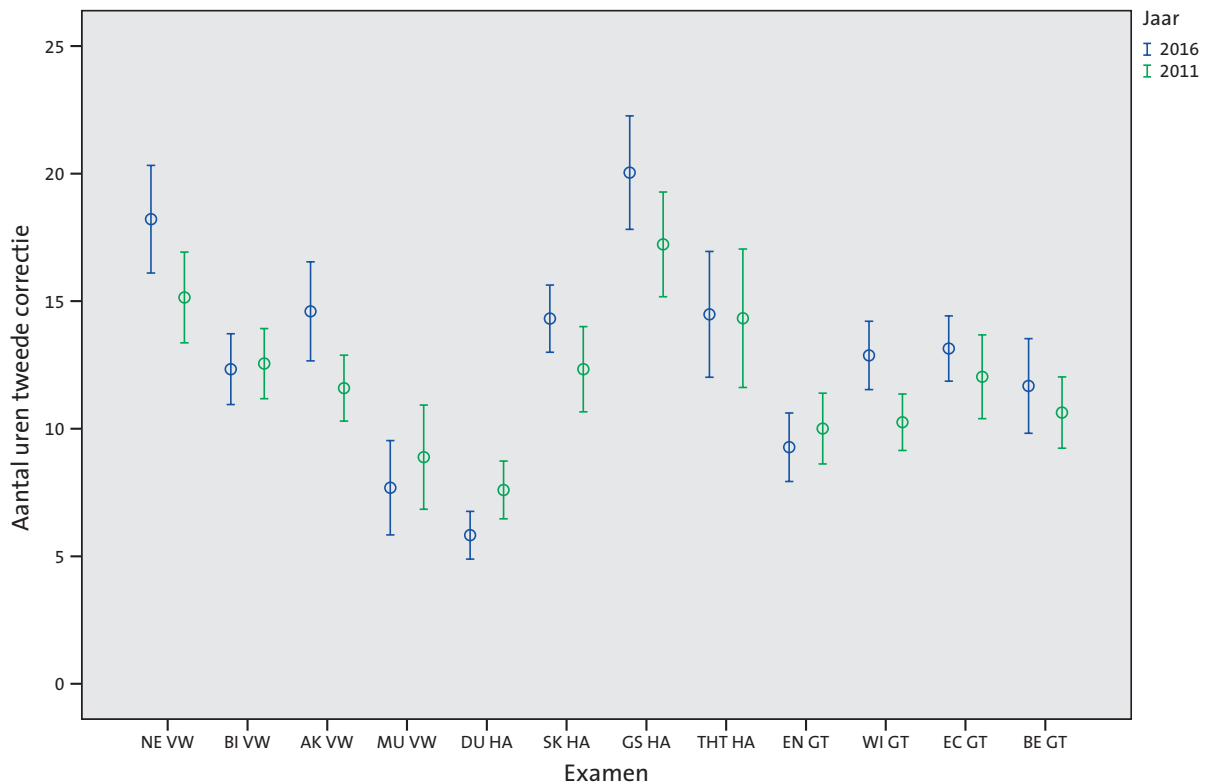
Hiervoor constateerden we dat het verschil tussen de tijdbesteding aan de tweede correctie in 2016 en 2011 niet onafhankelijk is van het specifieke examen. Figuur 3 en 4 tonen voor elk examen en jaar respectievelijk het gemiddeld aantal beoordeelde kandidaten en de gemiddelde correctietijd in uren (voor de tweede correctie). De gemiddelden zijn aangeduid met een cirkel. De verticale lijnen rond het gemiddelde symboliseren het 99%-betrouwbaarheidsinterval. Dat geeft aan binnen welke grenzen het gemiddelde met 99% zekerheid ligt. Als de betrouwbaarheidsintervallen van twee gemiddelden elkaar niet overlappen, mag worden aangenomen dat de gemiddelden verschillend zijn.

Figuur 3 laat zien dat het verschil tussen het aantal kandidaten dat docenten in 2016 en 2011 als tweede corrector moesten beoordelen bij het ene examen groter is dan bij het andere examen, maar dat dit verschil bij geen enkel examen statistisch betekenisvol is.



Figuur 3 Gemiddeld aantal kandidaten voor de tweede correctie per examen per jaar

Vergelijken we in figuur 4 het aantal uren voor de tweede correctie in 2016 en 2011 per examen, dan lijkt er alleen bij wiskunde sprake van een statistisch significant verschil. Docenten wiskunde deden in 2016 gemiddeld drie uur langer over de tweede correctie dan in 2011. Overigens besteedden ook docenten Nederlands, geschiedenis en aardrijkskunde in 2016 meer tijd aan de tweede correctie dan in 2011, maar hier is het verschil statistisch niet significant.



Figuur 4 Gemiddelde aantal uren voor de tweede correctie per examen per jaar

3.1.3 Tijdbesteding aan de eerste en tweede correctie tezamen

Hoeveel tijd besteden docenten aan de eerste en tweede correctie samen? In 2016 heeft de gemiddelde docent het werk van 80 examenkandidaten nagekeken en daar in totaal 35 klokuren aan besteed in acht dagen (waarvan 4,5 voor de eerste correctie en 3,2 voor de tweede correctie).

Verschillen tussen docenten

Tabel 3.1.8 toont de verdeling van het totaal aantal kandidaten en de totale tijdbesteding in uren en dagen voor de eerste en tweede correctie samen. De correctielast en de tijdbesteding aan de eerste en tweede correctie tezamen blijken zeer ongelijk verdeeld over de docenten. De 10% met de kleinste correctielast (P10) keek het werk van hooguit 34 kandidaten na en besteedde daar tot 10 klokuren aan in hooguit 4 dagen. De 10% met de zwaarste correctielast (P90) keek het werk van minimaal 140 kandidaten na en besteedde daar ten minste 68 uur aan in 13 dagen of meer.

Tabel 3.1.8 Tijdbesteding en correctielast van de eerste en tweede correctie samen

	Jaar	N	Min	Max	Gem	Se-Gem	Stddev	P10	P25	P75	P90
Aantal nagekeken kandidaten eerste en tweede correctie	2016	3241	5	403	79,65	0,79	44,75	34,00	48,00	110,00	140,00
	2011	3427	8	440	79,48	0,78	45,86	33,00	47,00	105,00	140,00
Aantal uren eerste en tweede correctie	2016	3225	3	300	34,89	0,49	27,83	10,00	16,00	45,00	68,00
	2011	3392	2	240	34,69	0,48	27,74	10,00	16,00	45,00	70,00
Aantal minuten eerste en tweede correctie per kandidaat	2016	3216	5	386	60,58	0,79	44,80	18,65	30,03	77,37	116,76
	2011	3376	4	420	60,07	0,77	44,96	19,29	30,00	75,00	116,00
Aantal dagen eerste en tweede correctie	2016	3243	2	35	7,70	0,07	3,88	4,00	5,00	9,00	13,00
	2011	3417	1	27	7,12	0,06	3,64	4,00	5,00	9,00	12,00

Verschillen tussen jaren

De verschillen tussen jaren (2016 versus 2011) en examens zijn op significantie getoetst met behulp van univariate variantieanalyse. Het hoofdeffect van jaar was bij een van de vier toetsingen significant ($p < 0,01$). In 2016 waren de docenten gemiddeld wat langer bezig met de correctie dan in 2011 (7,7 versus 7,1 dagen) maar er was geen significant verschil in het aantal beoordeelde kandidaten, het aantal bestede uren en de gemiddelde correctietijd per kandidaat. Deze bevindingen moeten geïnterpreteerd worden in het licht van het hoofdeffect van examen en de interactie tussen jaar en examen die bij alle vier aspecten van de tijdbesteding significantie te zien gaven ($p < 0,01$). Dit betekent dat de verschillen tussen beide jaren voor het ene examen anders zijn dan voor het andere examen.

Verschillen tussen examens

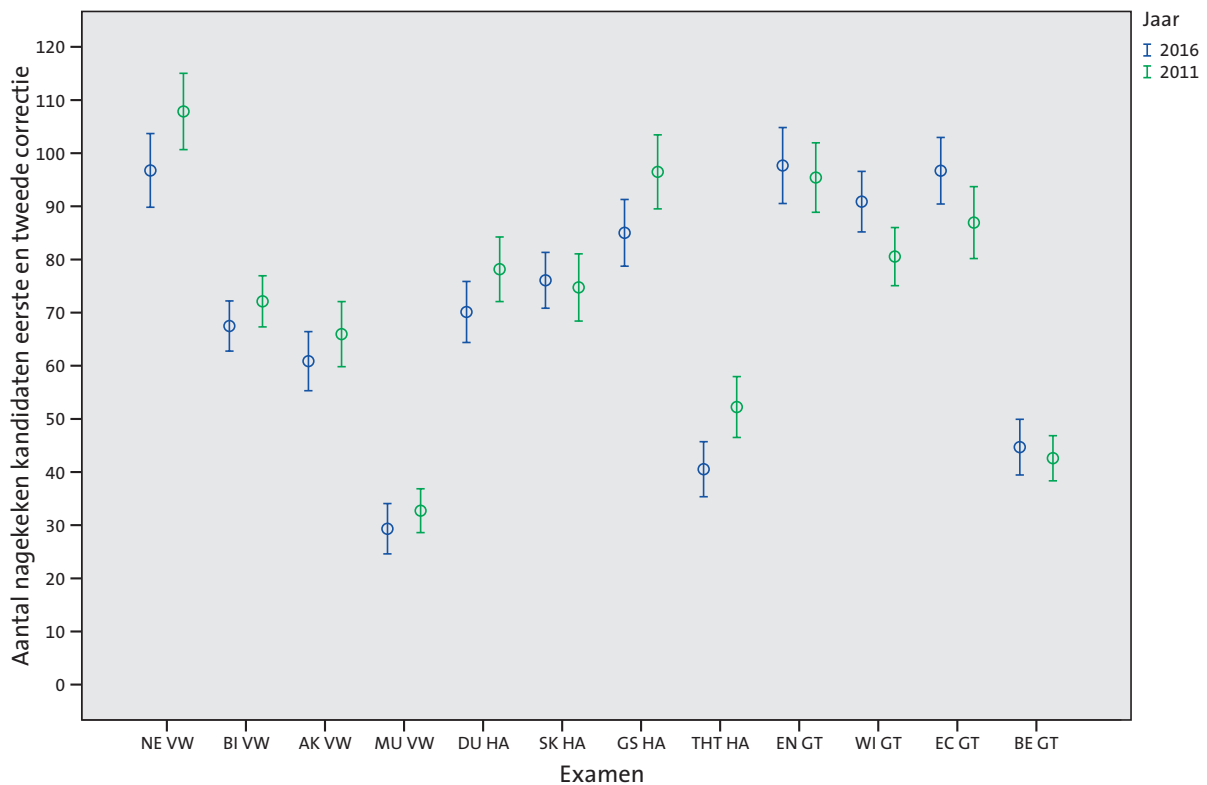
De gemiddelden per examen per jaar voor de eerste en tweede correctie samen zijn weergegeven in tabel 3.1.9. Te zien is onder meer dat aantal beoordeelde kandidaten en de tijdbesteding sterk van examen tot examen en van jaar tot jaar variëren. Zo beoordelen docenten Engels in 2016 ruim drie keer zoveel kandidaten als docenten muziek en maken geschiedenisdocenten bijna vier keer zoveel uren als docenten Duits.

Tabel 3.1.9 Tijdbesteding en correctielast van de eerste en tweede correctie per examen per jaar

	Jaar	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	ENG T	WIG T	EC GT	BEG T
Aantal nagekeken kandidaten	2016	97	67	61	29	70	76	85	41	98	91	97	45
	2011	108	72	66	33	78	75	96	52	95	81	87	43
Aantal uren	2016	53	36	39	19	15	39	57	36	23	29	35	29
	2011	47	38	37	24	21	37	57	38	25	27	31	26
Aantal minuten per kandidaat	2016	71	69	87	90	28	66	88	123	31	42	46	84
	2011	57	69	80	96	35	64	76	95	35	44	49	85
Aantal dagen	2016	11	8	8	5	4	8	11	8	6	7	7	7
	2011	10	7	7	6	5	7	10	7	6	6	6	6

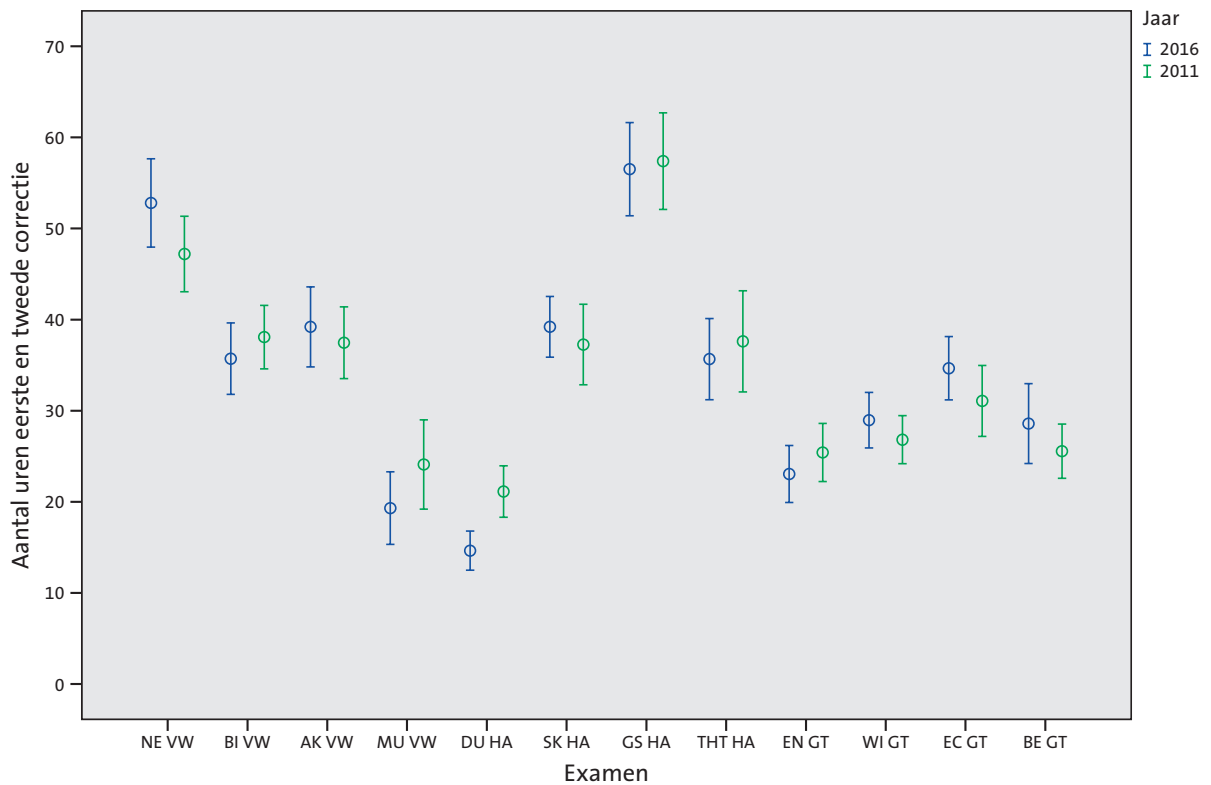
Verschillen tussen jaren per examen

Figuur 5 en 6 tonen voor elk examen en jaar respectievelijk de gemiddelden en de bijbehorende 99%- betrouwbaarheidsintervallen voor het totale aantal beoordeelde kandidaten en de totale correctietijd voor de eerste en tweede correctie samen. Figuur 5 laat onder meer zien dat docenten Nederlands en geschiedenis in 2016 minder kandidaten beoordeelden dan in 2011 en dat docenten wiskunde juist meer kandidaten nakeken, maar tevens dat deze verschillen statistisch gezien verwaarloosbaar klein zijn.



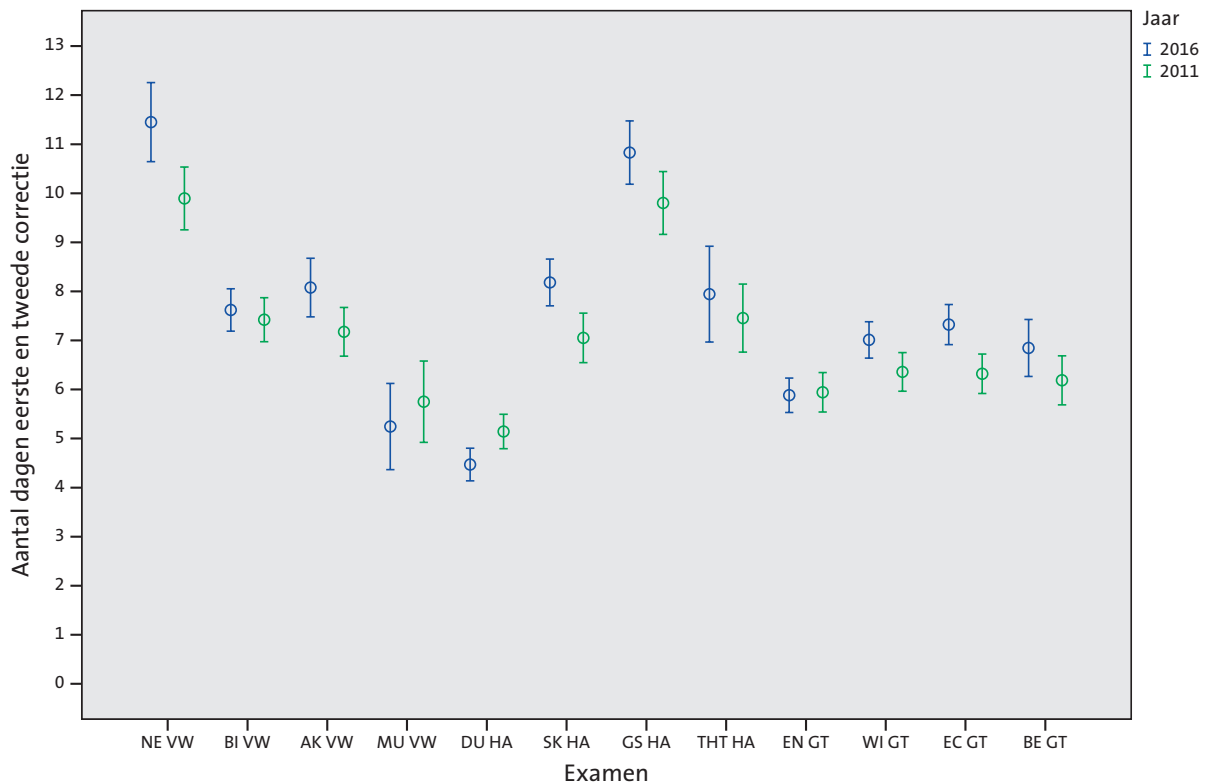
Figuur 5 Gemiddeld aantal kandidaten voor de eerste en tweede correctie samen per examen per jaar

Het verschil tussen aantal bestede uren in 2016 en 2011 is alleen bij het vak Duits significant (zie figuur 6). In 2016 besteedden docenten Duits in zijn totaliteit zes uur minder aan de eerste en tweede correctie dan in 2011. Overigens investeerden docenten Nederlands zes uur meer in de correctie, maar dit verschil is gezien de wat grotere betrouwbaarheidsintervallen statistisch gezien niet aantoonbaar.



Figuur 6 Gemiddeld aantal uren voor de eerste en tweede correctie samen per examen per jaar

Bij Nederlands, scheikunde en economie is er sprake van een significant verschil tussen beide jaren voor wat betreft de lengte van de periode waarin de eerste en tweede correctie werden uitgevoerd (zie figuur 7). Bij deze drie vakken gaat het om ongeveer één dag extra.



Figuur 7 Gemiddeld aantal dagen voor de eerste en tweede correctie samen per examen per jaar

3.1.4 Verschillen tussen de tijdbesteding aan de eerste en tweede correctie

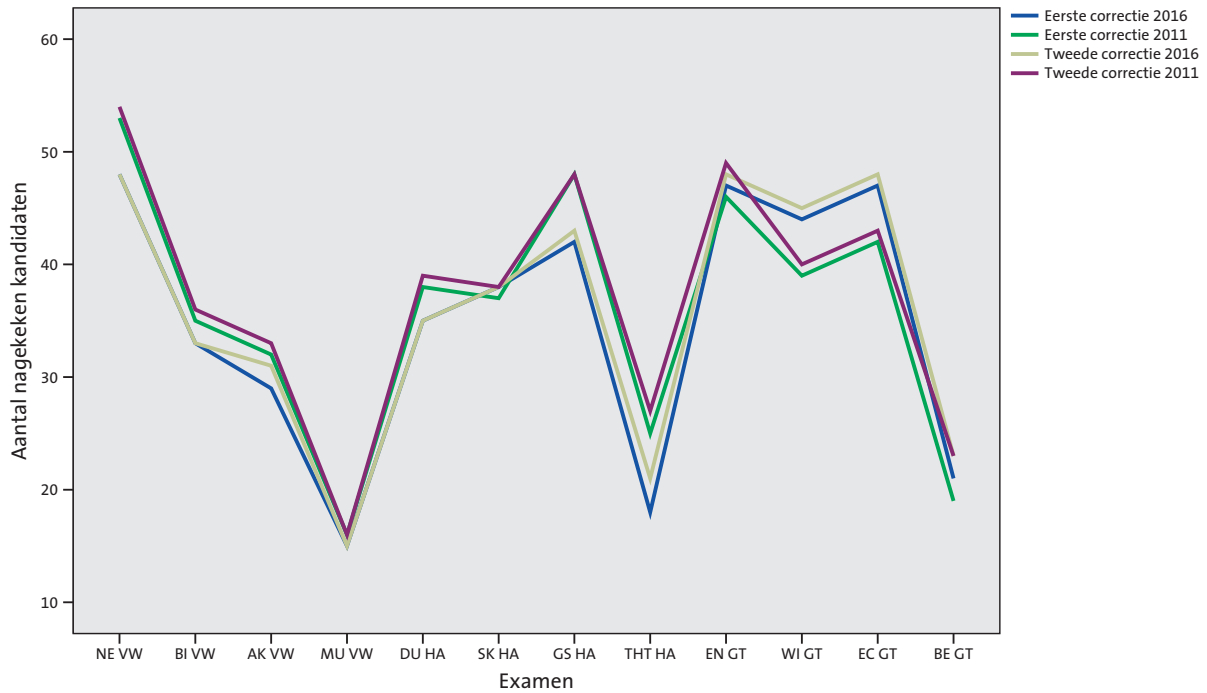
De verschillen tussen de tijdbesteding aan de eerste en tweede correctie zijn geanalyseerd met behulp van een multivariate variantieanalyse volgens een herhaalde-metingen ontwerp. De correctieprocedure (eerste versus tweede correctie) is daarbij gespecificeerd als een binnen-subjecten factor en het jaar (2016 versus 2011) en het examen als tussen-subjecten factoren. De analyse is per aspect van de tijdbesteding afzonderlijk uitgevoerd.

Tabel 3.1.10 Verschillen qua tijdbesteding tussen de eerste en tweede correctie

	Jaar	N	Min	Max	Gem	Se-Gem	Stddev	P10	P25	P75	P90
Aantal nagekeken kandidaten eerste correctie	2016	3378	1	228	38,84	0,40	23,32	15,00	23,00	55,00	70,00
	2011	3637	1	240	38,45	0,40	23,87	14,00	22,00	51,00	70,00
Aantal nagekeken kandidaten tweede correctie	2016	3330	1	200	39,81	0,39	22,42	17,00	25,00	55,00	70,00
	2011	3442	1	200	40,02	0,39	23,08	17,00	24,00	50,00	70,00
Aantal uren eerste correctie	2016	3364	1	260	21,32	0,32	18,54	5,00	10,00	28,00	40,00
	2011	3597	1	160	22,27	0,32	19,37	6,00	10,00	30,00	45,00
Aantal uren tweede correctie	2016	3313	1	105	13,15	0,20	11,33	4,00	6,00	16,00	25,00
	2011	3412	1	100	11,97	0,18	10,75	3,00	5,00	15,00	25,00
Aantal minuten per kandidaat eerste correctie	2016	3361	1	300	38,22	0,52	30,14	10,67	18,00	48,39	72,00
	2011	3591	1	360	39,87	0,52	31,41	12,00	20,00	50,53	75,73
Aantal minuten per kandidaat tweede correctie	2016	3305	1	200	22,70	0,32	18,63	6,76	10,91	30,00	43,16
	2011	3401	1	150	20,77	0,31	18,22	6,00	9,23	26,67	40,00
Aantal dagen eerste correctie	2016	3387	1	20	4,48	0,04	2,42	2,00	3,00	5,00	8,00
	2011	3616	0	20	4,29	0,04	2,45	2,00	3,00	5,00	7,00
Aantal dagen tweede correctie	2016	3319	1	24	3,16	0,03	1,82	1,00	2,00	4,00	5,00
	2011	3430	1	14	2,77	0,03	1,56	1,00	2,00	3,00	5,00

Aantal nagekeken kandidaten

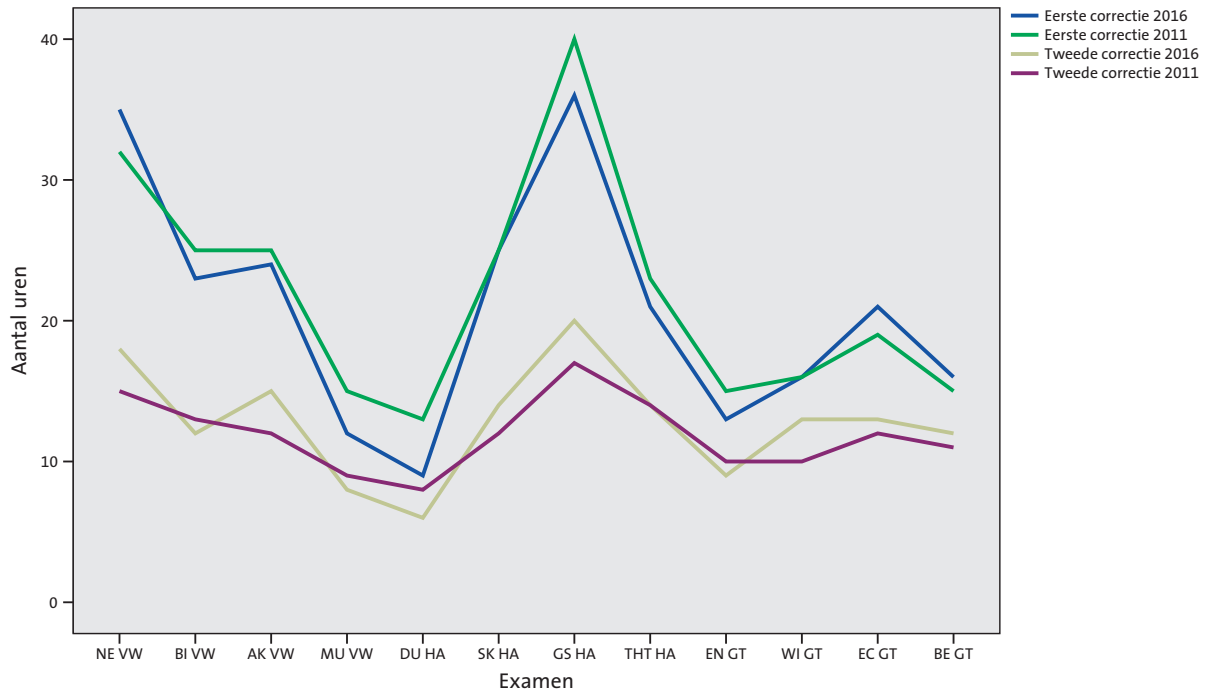
Uit de analyse van de binnen-subjecten effecten komt naar voren dat er sprake is van een significant hoofdeffect van correctieprocedure (eerste versus tweede correctie; $p < 0,001$) en een significante tweeweg interactie van correctieprocedure met examen ($p = 0,003$) maar dat de tweeweg interactie tussen correctieprocedure en jaar en de drieweg interactie tussen correctieprocedure, jaar en examen dat niet zijn (p respectievelijk 0,523 en 0,367). Er is dus een verschil tussen het aantal kandidaten dat bij de eerste en tweede correctie beoordeeld wordt en dat verschil is niet onafhankelijk van het examen, maar het verschil (tussen het aantal nagekeken kandidaten bij de eerste en tweede correctie) hangt niet samen met het jaar en de combinatie van jaar en examen. Figuur 7 geeft inzicht in het gemiddeld aantal nagekeken kandidaten per correctieprocedure per jaar. De figuur bevestigt wat de tabellen en de significantietoetsing al eerder deden vermoeden, namelijk dat de verschillen tussen de examens in de aantallen in 2016 en 2011 nagekeken kandidaten getalsmatig erg klein zijn.



Figuur 8 Verschil tussen het aantal nagekeken kandidaten voor de eerste en tweede correctie per examen per jaar

Aantal uren

Het hoofdeffect van de correctieprocedure en de drie interacties met jaar en examens zijn alle significant ($p < 0,001$). Er is dus een verschil tussen het aantal uren voor de eerste en tweede correctie en dat verschil hangt af van het jaar (2016 versus 2011), het examen en de specifieke combinatie van jaar en examens. Figuur 9 geeft inzicht in het totaal aantal bestede uren per correctieprocedure (eerste versus tweede correctie) per examen per jaar.



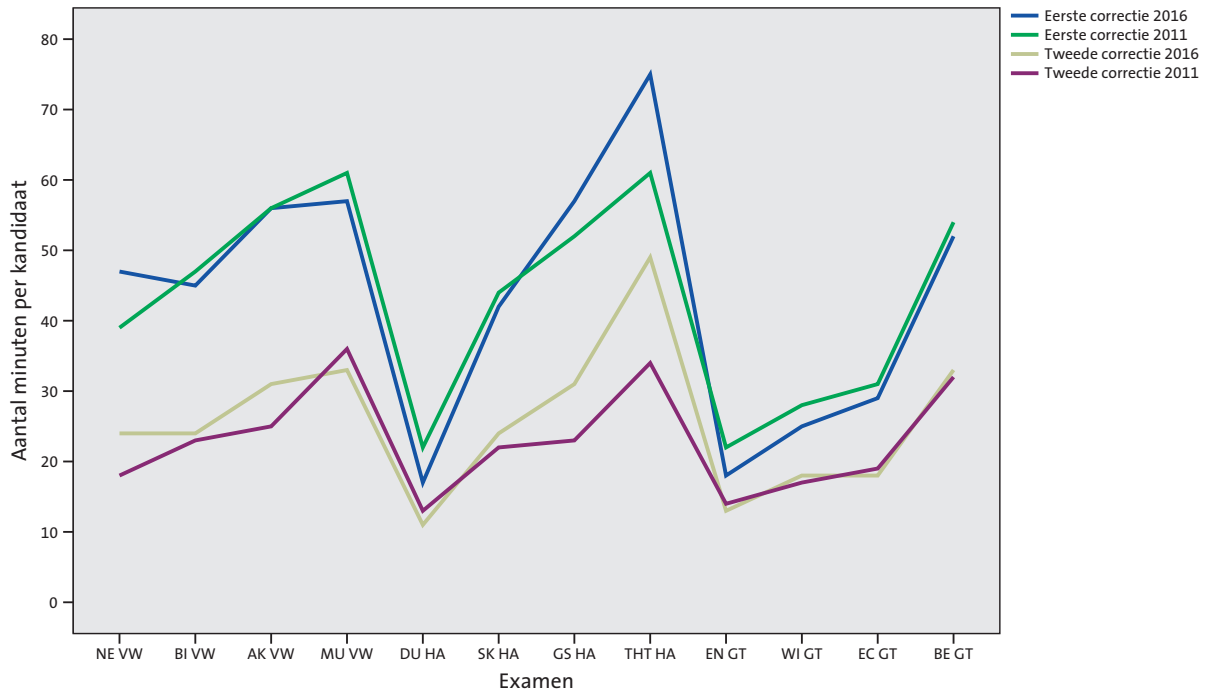
Figuur 9 Verschil tussen het aantal uren voor de eerste en tweede correctie per examen per jaar

Figuur 9 laat zien dat docenten veel meer tijd steken in de eerste correctie dan in de tweede correctie en dat dit verschil veel sterker afhangt van het examen dan van het jaar. In 2016 vergden de eerste en tweede correctie gemiddeld respectievelijk 21 en 13 uur en in 2011 waren de overeenkomstige gemiddelden 22 en 12 uur.

Eerder constateerden we dat de gemiddelde tijdbesteding aan de eerste correctie sinds 2011 met bijna één uur is afgenomen (van 22,27 naar 21,32 uur; zie tabel 3.1.1) en dat de tijdbesteding aan de tweede correctie met ruim één uur is toegenomen (van 11,97 naar 13,15 uur; zie tabel 3.1.6). Aangezien de totale tijdbesteding aan de eerste en tweede correctie tezamen vrijwel gelijk gebleven is (van 34,69 naar 34,89; zie tabel 3.1.8) lijkt het redelijk om te veronderstellen dat er sinds 2011 een kleine verschuiving heeft plaatsgevonden van de tijdbesteding van de eerste naar de tweede correctie.

Aantal minuten per kandidaat

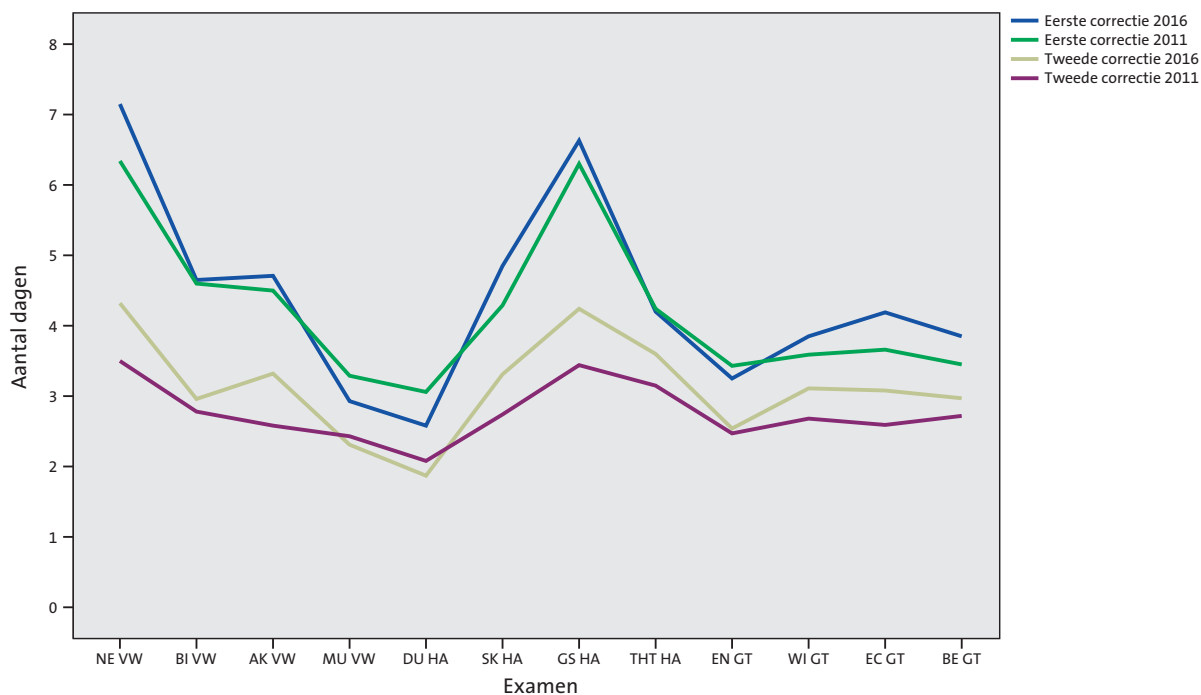
Het hoofdeffect van de correctieprocedure en de beide tweeweg interacties met jaar en examen zijn significant ($p < 0,001$) maar de drieweg interactie van correctieprocedure, lijst en examen is dat niet ($p = 0,143$). De gemiddelde correctietijd per kandidaat is dus niet gelijk voor de verschillende jaren en examens maar hangt niet af van de specifieke combinatie van jaar en examen. Figuur 10 geeft informatie over het verschil tussen het gemiddeld aantal minuten per kandidaat voor de eerste en tweede correctie per examen in 2016 (boven) en 2011 (onder). De figuur laat zien dat docenten per kandidaat veel meer tijd steken in de eerste correctie dan in de tweede correctie en dat dit verschil veel sterker afhangt van het examen dan van het jaar.



Figuur 10 Het gemiddeld aantal minuten per kandidaat voor de eerste en tweede correctie per examen per jaar

Aantal dagen

Het hoofdeffect van de correctieprocedure en de twee- en drieweg interacties met jaar en examens zijn significant ($p < 0,01$). Er bestaat dus een verschil tussen het aantal dagen voor de eerste en tweede correctie en dat verschil hangt af van het jaar, het examen en de combinatie van beide. In 2016 werden de eerste en tweede correctie uitgevoerd in respectievelijk 4,48 en 3,16 dagen en in 2011 ging het om respectievelijk 4,29 en 2,77 dagen. De significante interactie van correctieprocedure en jaar betekent dat het verschil tussen het aantal dagen voor de eerste en tweede correctie in 2016 iets kleiner is dan in 2011, namelijk 1,32 dagen in 2016 tegen 1,52 in 2011. Gegeven de significante drieweg interactie is dit verschil niet onafhankelijk van het examen. Figuur 11 toont het verschil tussen het aantal dagen voor de eerste en tweede correctie per examen in 2016 en 2011. De figuur maakt aannemelijk dat het verschil tussen het aantal dagen voor de eerste en tweede correctie bij een tijdsintensief examen als Nederlands in 2016 ongeveer even groot is als in 2011; bij een ander bewerkelijk examen - geschiedenis - lijkt dat verschil daarentegen kleiner te zijn geworden (aangezien de eerste correctie in 2016 ongeveer evenveel tijd kost als in 2011 en de tweede correctie nu duidelijk meer tijd vergt dan destijds).



Figuur 11 Gemiddeld aantal dagen voor de eerste en tweede correctie per examen per jaar

3.1.5 Ervaren belasting van de eerste en tweede correctie

In het voorgaande constateerden we dat de eerste correctie wat minder en de tweede correctie wat meer tijd kost dan in 2011 maar dat de tijdbesteding aan de eerste en tweede correctie tezamen gelijk gebleven is. In deze paragraaf gaan we na hoe belastend de docenten de correctie vinden en in hoeverre de ervaren belasting sinds 2011 veranderd is.

Ervaren belasting eerste correctie

In 2016 vindt 11% van de docenten de eerste correctie niet belastend, 35% enigszins belastend, 36% belastend en 17% zeer belastend (zie tabel 3.1.11).

Tabel 3.1.11 Ervaren belasting van de eerste correctie per jaar (kolompercentages optellend tot 100%)

	2016	2011
Niet belastend	11	13
Enigszins belastend	35	36
Belastend	36	38
Zeer belastend	17	13

Verschillen tussen jaren

De significantietoetsing van de verschillen tussen jaren laat zien dat docenten de eerste correctie in 2016 als meer belastend ervaren dan in 2011 ($p < 0,001$). Zo vond 17% de eerste correctie in 2016 zeer belastend tegen 13% in 2011.

Verschillen tussen examens

De naar examens en jaar uitgesplitste percentages zijn weergegeven in tabel 3.1.12. De verschillen tussen examens zijn significant, zowel in 2016 als in 2011 ($p < 0,001$). Zo ervaren

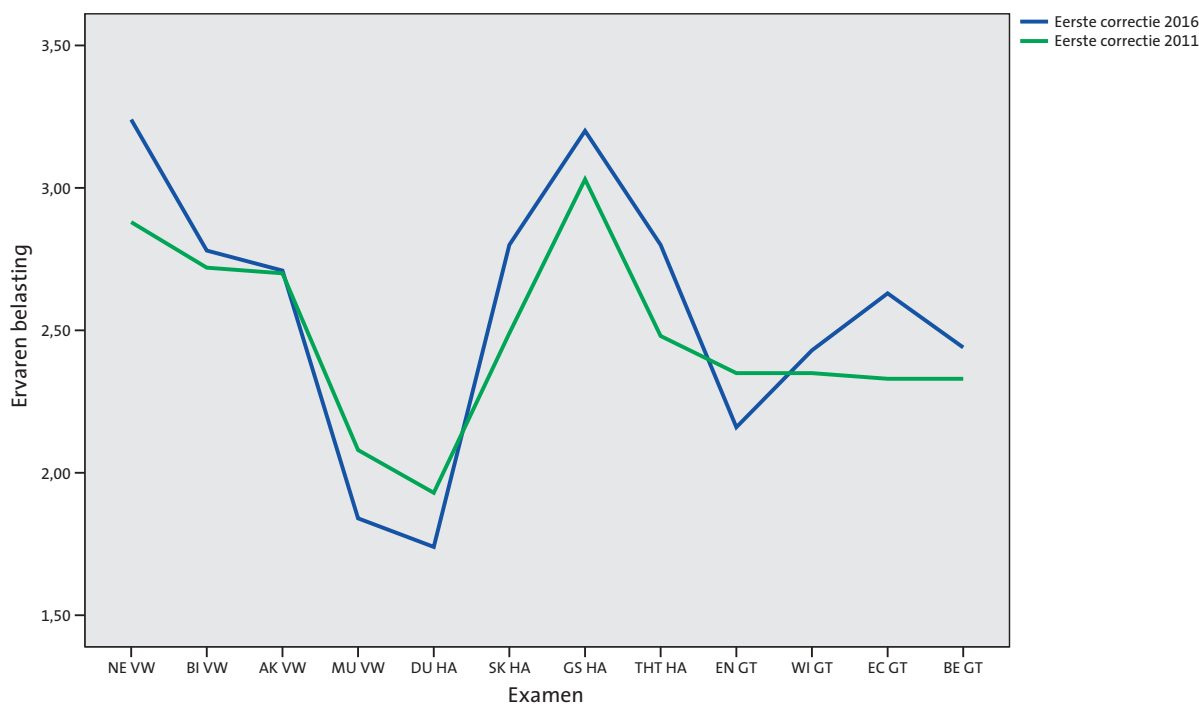
Nederlands en geschiedenis de eerste correctie als een veel grotere belasting dan docenten Duits en Engels. Waarschijnlijk niet toevallig zijn Nederlands en geschiedenis tevens examens met veel open vragen en een hoge correctielast.

Tabel 3.1.12 De mate waarin docenten de eerste correctie als belastend ervaren per examen per jaar (kolompercentages per examen per jaar optellend tot 100%).

Jaar		NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
2016	Niet belastend	1%	4%	8%	38%	41%	5%	2%	9%	20%	11%	8%	8%
	Enigszins belastend	16%	32%	33%	40%	45%	29%	19%	25%	48%	43%	36%	49%
	Belastend	40%	46%	41%	22%	13%	46%	35%	42%	27%	37%	41%	34%
	Zeer belastend	42%	18%	19%	0%	1%	20%	44%	23%	4%	9%	15%	9%
2011	Niet belastend	4%	7%	10%	25%	31%	10%	4%	10%	15%	15%	17%	15%
	Enigszins belastend	25%	28%	30%	46%	48%	39%	20%	41%	43%	41%	40%	44%
	Belastend	49%	49%	41%	26%	17%	42%	44%	39%	34%	38%	34%	35%
	Zeer belastend	22%	15%	19%	3%	4%	9%	32%	10%	8%	6%	8%	6%

Verschillen tussen jaren per examen

De toetsing van de verschillen tussen 2016 en 2011 geeft bij een aantal vakken significantie te zien ($p < 0,01$). De percentages in tabel 3.1.12 zijn lastig te vergelijken tussen examens. Daarom toont figuur 12 de gemiddelden voor de ervaren belasting door de eerste correctie per jaar per examen, uitgedrukt op een schaal van 1 (niet belastend) tot 4 (zeer belastend). Docenten Nederlands, scheikunde, tehatex en economie ervaren de eerste correctie nu als meer belastend dan toen en docenten Duits, Engels en muziek vinden de eerste correctie nu minder belastend dan destijds.



Figuur 12 Ervaren belasting door de eerste correctie per jaar per examen

Ervaren belasting tweede correctie

In 2016 vindt 8% van de docenten de tweede correctie niet belastend, 35% enigszins belastend, 39% belastend en 19% zeer belastend (zie tabel 3.1.13).

Tabel 3.1.13 *Ervaren belasting van de tweede correctie per jaar (kolompercentages optellend tot 100%)**

	2016	2011
Niet belastend	8	9
Enigszins belastend	35	34
Belastend	39	39
Zeer belastend	19	17

Verschillen tussen jaren

De significantietoetsing van de verschillen tussen jaren laat zien dat docenten de tweede correctie in 2016 als even belastend ervaren als in 2011 ($p = 0,238$).

Verschillen tussen examens

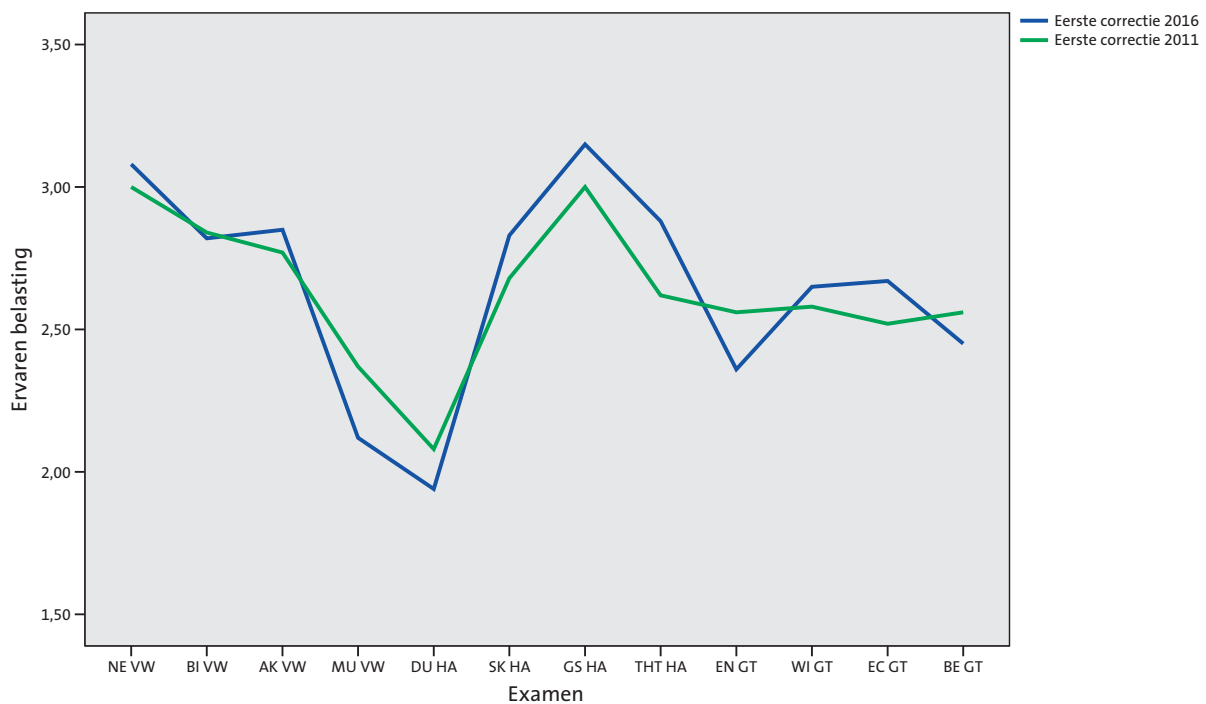
In tabel 3.1.14 is de ervaren belasting van de tweede correctie uitgesplitst naar examens en jaar. De verschillen tussen examens zijn significant, zowel in 2016 als in 2011 ($p < 0,001$). Zo ervaren docenten tehatex de tweede correctie in 2016 veel meer als een belasting dan docenten muziek.

Tabel 3.1.14 *Ervaren belasting van de tweede correctie per examen per jaar (kolompercentages per examen per jaar optellend tot 100%)*

Jaar	Belasting	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	ENG T	WI GT	EC GT	BE GT
2016	Niet belastend	1	2	5	20	29	5	4	2	12	6	7	10
	Enigszins belastend	23	31	27	47	50	30	18	32	49	38	35	45
	Belastend	44	48	46	33	18	43	37	41	29	41	43	35
	Zeer belastend	33	18	22	0	3	23	41	25	9	15	15	10
2011	Niet belastend	3	4	3	15	24	9	3	8	10	10	11	11
	Enigszins belastend	21	28	34	47	49	32	23	40	40	36	40	35
	Belastend	49	47	45	26	22	42	43	34	34	40	36	41
	Zeer belastend	27	21	18	13	5	17	30	18	16	14	13	13

Verschillen tussen jaren per examen

De toetsing van de verschillen tussen 2016 en 2011 geeft alleen bij Engels significantie te zien ($p = 0,002$). Omwille van een eenvoudige interpretatie toont figuur 13 de gemiddelden per jaar per examen op een schaal van 1 (niet belastend) tot 4 (zeer belastend). Docenten Engels vinden de tweede correctie in 2016 minder belastend dan in 2011. Voor de overige elf examens geldt dat docenten de tweede correctie in 2016 als even belastend ervaren als in 2011. Overigens zien we ook bij muziek een relatief groot verschil tussen beide jaren, maar dat is vanwege het kleine aantal waarnemingen niet significant.



Figuur 13 Ervaren belasting door de tweede correctie per jaar per examen

Verschillen tussen de ervaren belasting van de eerste en tweede correctie

De verschillen tussen de ervaren belasting van de eerste versus de tweede correctie in 2016 en 2011 zijn weergegeven in tabel 3.1.15.

Tabel 3.1.15 Ervaren belasting van de eerste en tweede correctie in 2016 en 2011 (kolompercentages optellend tot 100%)

	2016		2011	
	Eerste correctie	Tweede correctie	Eerste correctie	Tweede correctie
Niet belastend	11	8	13	9
Enigszins belastend	35	35	36	34
Belastend	36	39	38	39
Zeer belastend	17	19	13	17

Een multivariate variantieanalyse volgens een herhaalde-metingen ontwerp laat zien dat het hoofdeffect van de correctieprocedure (eerste versus tweede correctie), de beide tweeweg interacties met jaar en examen en de drieweg interactie met jaar en examen significant zijn ($p < 0,01$). De interpretatie is als volgt: de tweede correctie wordt gemiddeld genomen als sterker belastend ervaren dan de eerste correctie, maar dit verschil in beleving is afhankelijk van de specifieke combinatie van jaar en examen.

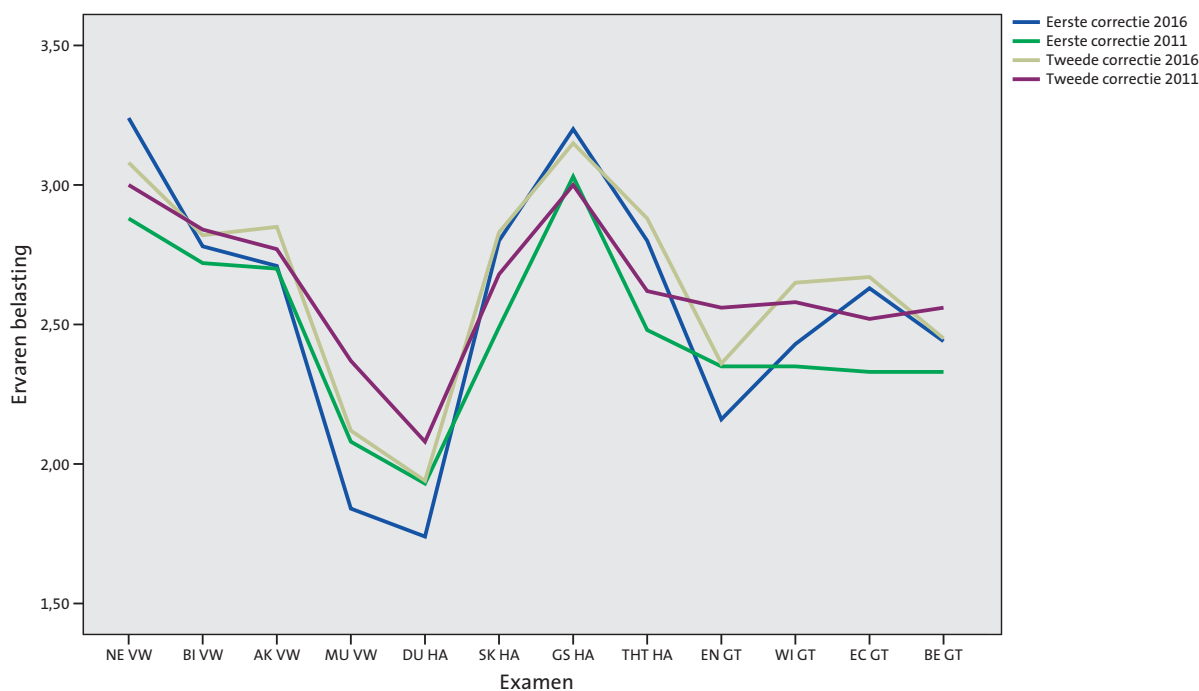
Verschillen tussen de ervaren belasting van de eerste en tweede correctie per examen per jaar

Ter illustratie van de geconstateerde interacties tonen tabel 3.1.16 en figuur 14 de gemiddelden voor de ervaren belasting door de eerste en tweede correctie per jaar per examen. De ervaren belasting is daarbij uitgedrukt op een schaal van 1 (niet belastend) tot 4 (zeer belastend). Zowel in 2016 als in 2011 vonden docenten muziek, Engels en wiskunde de eerste correctie wat minder belastend dan de tweede correctie. Bij docenten Nederlands blijkt het verschil tussen de

ervaren belasting van de eerste en tweede correctie in 2016 echter anders dan in 2011. Net als de andere hiervoor genoemde vakken vonden docenten Nederlands de eerste correctie in 2011 minder belastend dan de tweede correctie, maar in 2016 was het juist de eerste correctie die zij het meest belastend vonden. Ook bij scheikunde en economie constateren we een interactie met het jaar in de zin dat het verschil tussen de ervaren belasting van de eerste en tweede correctie in 2016 kleiner is dan in 2011.

Tabel 3.1.16 Ervaren belasting door de eerste en tweede correctie per jaar per examen

Jaar	Belasting	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BEGT
2016	Eerste correctie	3,24	2,78	2,71	1,84	1,74	2,80	3,20	2,80	2,16	2,43	2,63	2,44
	Tweede correctie	3,08	2,82	2,85	2,12	1,94	2,83	3,15	2,88	2,36	2,65	2,67	2,45
2011	Eerste correctie	2,88	2,72	2,70	2,08	1,93	2,49	3,03	2,48	2,35	2,35	2,33	2,33
	Tweede correctie	3,00	2,84	2,77	2,37	2,08	2,68	3,00	2,62	2,56	2,58	2,52	2,56



Figuur 14 Ervaren belasting door de eerste en tweede correctie per jaar per examen

3.1.6 Toelichtingen van docenten bij de vraag naar de ervaren belasting van de eerste en tweede correctie

Ruim vijfhonderd docenten hebben de moeite genomen hun antwoord op de meerkeuzevraag naar de redenen waarom zij de tweede correctie als (niet) belastend ervaren toe te lichten. Deze vraag kende vier antwoordmogelijkheden: 'Niet belastend', 'Enigszins belastend', 'Belastend' en 'Sterk belastend'. Men kon maximaal twee redenen geven. In totaal 527 respondenten gaven een eerste reden en 516 (daarenboven) een tweede reden. De toelichtingen zijn gecodeerd met behulp van een categorieënsysteem. Voor het uitgebreide verslag van de inhoudsanalyse wordt verwezen naar Loeffen en Eikelboom (in voorbereiding).

De resultaten van de analyse zijn weergegeven in tabel 3.1.17. Deze tabel geeft voor elk van de vier antwoordmogelijkheden van de meerkeuzevraag de verdeling van de redenen per categorie van het categorieënsysteem. Hierbij zijn de eerste en tweede redenen samengenomen. De verdeling van de antwoorden op de meerkeuzevraag in de totale groep tweede correctoren was respectievelijk 8%, 35%, 39% en 19% (zie tabel 3.1.13) en in de groep die een toelichting bij deze vraag gaf respectievelijk 7%, 31%, 41% en 21%.

Tabel 3.1.17 Redenen waarom docenten de tweede correctie al dan niet als belastend ervaren per antwoord-categorie van de vraag naar de ervaren belasting (kolompercentages optellend tot 100%)

	Niet belastend	Enigszins belastend	Belastend	Zeer belastend	Totaal
Correctielast en tijdsdruk					
Tijdrovend, tijdsdruk of tijdgebrek (algemeen)		22	17	17	18
Veel werk/moest veel kandidaten nakijken		3	9	7	6
De tweede correctie moet integraal		0	0	1	0
Tweede correctie moet in kortere tijd dan de eerste correctie		0	0		0
Veel tijd tussen eerste en tweede correctie		1	1		1
Werk stelt hoge eisen					
Vereist concentratie en zorgvuldigheid		3	4	2	3
Vereist vakinhoudelijke deskundigheid en/of beoordelingsdeskundigheid		2	2	2	2
Vereist inlevingsvermogen in (onbekende) leerlingen en collega's		3	2	1	2
Gebrek aan facilitering op school					
Overige werkzaamheden gaan gewoon door	1	12	22	22	17
Geen of beperkte financiële vergoeding		2	2	3	2
Probleem om op school goede werkplek te vinden		1			0
Vragen, correctievoorschrift en aanvullingen van vakverenigingen					
Correctievoorschrift en/of examen biedt (te) veel interpretatieruimte, niet helder, tijdrovend of slecht		2	7	6	5
Aanvullingen vakverenigingen/kringbesprekingen lang, tijdrovend of onduidelijk		0		0	0
Eerste correctie					
Eerste corrector was onzorgvuldig of had niet goed nagekeken		2	7	11	6
Eerste corrector was (te) soepel		1	3	6	3
Eerste corrector heeft het werk niet (goed) geannoteerd		1	2	1	1
Werk van eerste corrector kwam (te) laat binnen		2	0	1	1
Gezamenlijk overleg					
Gezamenlijk overleg kost tijd/duurt lang		7	7	6	6
Gezamenlijk overleg verloopt niet prettig of slechte communicatie met (school) eerste corrector		5	5	6	5
Psychologische oorzaken					
Psychologisch zwaar om na de eerste correctie weer te beginnen aan de tweede correctie		3	3	2	2
Minder gemotiveerd omdat het om onbekende leerlingen gaat		1	1	1	1
Het is een grote verantwoordelijkheid / het is niet prettig een ander te moeten controleren.		1	2	0	1

(vervolg tabel 3.1.17)

	Niet belastend	Enigszins belastend	Belastend	Ze er belastend	Totaal
Redenen waarom tweede correctie NIET BELASTEND is					
Het hoort er gewoon bij	25	5			3
Minder belastend omdat ik de eerste correctie al gedaan heb	21	4	0		3
Eerste corrector had zijn/haar werk goed gedaan	14	5	1		3
Veel meerkeuzevragen	10	2			1
Weinig kandidaten	7	1			1
Interessant om te zien hoe leerlingen of docenten van andere school het examens maken	11	2	0		2
Goede vragen, correctievoorschrift, examens of aanvullingen op het correctievoorschrift	5				0
Gewone lessen vervallen	3	1			0
Overige					
Weet niet/geen mening		1	0	1	1
Off-topic (geen antwoord op de vraag)	4	5	3	3	4
Totaal	100	100	100	100	100

Allereerst beschrijven we de verdeling van de redenen die docenten aanvoeren om aan te geven dat zij de tweede correctie niet belastend vonden. In de totale groep docenten vond 8% de tweede correctie niet belastend en in de groep die een toelichting gaf, was dat 7%.

De belangrijkste redenen zijn:

- Het hoort er gewoon bij (25%);
- Minder belastend omdat je eerste correctie al gedaan hebt (21%);
- De eerste corrector had zijn/haar werk goed gedaan (14%);
- Het is interessant om te zien hoe leerlingen/docenten van andere school het examens maken/corrigeren (11%);
- Veel meerkeuzevragen (10%);
- Weinig kandidaten (7%);
- Goede vragen, correctievoorschrift, examens, aanvullingen op het correctievoorschrift (5%);
- De gewone lessen vervallen (3%).

Vervolgens bespreken we de verdeling van de redenen waarom docenten de tweede correctie als enigszins belastend, belastend of zeer belastend ervaren. In de totale groep docenten vond 92% de tweede correctie enigszins tot zeer belastend en in de groep die een toelichting gaf, was dat 93%. De volgende redenen werden genoemd door meer dan 5% van degenen die de tweede correctie enigszins tot zeer belastend vonden:

- De overige werkzaamheden gaan gewoon door (22%);
- Tijdrovend, tijdsdruk of tijdgebrek (algemeen) (17%);
- De eerste corrector was onzorgvuldig of had niet goed nagekeken (11%);
- Veel werk of veel kandidaten (7%);
- Het gezamenlijk overleg verloopt niet prettig / slechte communicatie met de eerste corrector (6%);
- Het gezamenlijk overleg kost tijd of duurt lang (6%);
- Het correctievoorschrift en/of het examen bieden (te) veel interpretatieruimte, is niet helder, tijdrovend of slecht (6%);
- De eerste corrector was (te) soepel (6%).

3.2 Schoolbeleid en tevredenheid

Hoe faciliteren scholen de eerste en tweede correctie? En hoe tevreden zijn docenten met dit schoolbeleid? Dat zijn de twee vragen die in deze paragraaf centraal zijn.

3.2.1 Het schoolbeleid met betrekking tot de eerste en tweede correctie

Schoolbeleid voor de eerste correctie

De antwoorden op de vraag naar de facilitering van de eerste correctie zijn weergegeven in tabel 3.2.1. In 2016 is 71% van de docenten van mening dat de eerste correctie op hun school niet gefaciliteerd wordt. Van 16% zijn de uren voor de eerste correctie aanwijsbaar opgenomen in de jaartaak (bijv. in de opslagfactor), 7% krijgt een vast bedrag per examenwerk, 2% kan onkosten declareren (bijvoorbeeld telefoonkosten) en 1% ontvangt een kleine attentie (bijv. een boekenbon).

Tabel 3.2.1 De facilitering van de eerste correctie per jaar

	Jaar	N	Gemiddeld percentage
De eerste correctie wordt op mijn school niet gefaciliteerd	2016	3293	71
	2011	3640	75
Ik krijg een vast bedrag per examenwerk	2016	3293	7
	2011	3640	9
Ik kan onkosten declareren (bijv. telefoonkosten)	2016	3293	2
	2011	3640	2
Ik ontvang een kleine attentie (bijv. een boekenbon)	2016	3293	1
	2011	3640	1
De uren voor de eerste correctie zijn aanwijsbaar opgenomen in mijn jaartaak (bijv. in de opslagfactor)	2016	3293	16
	2011	3640	13

Verschillen tussen 2016 en 2011

Met behulp van binaire logistische regressieanalyse is nagegaan in hoeverre de verschillen tussen 2016 en 2011 (zie tabel 3.2.1) statistisch significant zijn. In vergelijking met 2011 zijn in 2016 iets minder docenten van mening dat de eerste correctie op hun school niet gefaciliteerd wordt, zijn de uren voor de eerste correctie wat vaker opgenomen in de jaartaak en krijgen docenten wat minder vaak een vast bedrag per examenwerk. De verschillen voor 'onkosten declareren' en 'kleine attentie' zijn statistisch gezien betekenisloos.

Verschillen tussen examens

Tabel 3.2.2 geeft informatie over hoe docenten van verschillende vakken het schoolbeleid met betrekking tot de eerste correctie in 2016 en 2011 typeren. Een binaire logistische regressieanalyse laat zien dat de verschillen tussen examens in 2016 bij drie van de vijf aspecten van het schoolbeleid significant zijn ($p < 0,010$). Het betreft het ontbreken van faciliteiten voor de eerste correctie, het toewijzen van de uren aan de jaartaak en het toekennen van een vast bedrag per examenwerk. Zo zijn docenten Nederlands en geschiedenis vaker van mening dat de eerste correctie op hun school niet gefaciliteerd wordt dan docenten muziek en economie. Bij docenten economie zijn de uren voor de eerste correctie vaker opgenomen in de jaartaak dan bij docenten

Nederlands en Duits. Docenten muziek en Engels krijgen vaker een vast bedrag per examenwerk dan docenten aardrijkskunde en scheikunde.

Tabel 3.2.2 De facilitering van de eerste correctie per examen per jaar (percentages eerste correctoren)

	Jaar	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BEGT
Geen faciliteiten	2016	78	71	79	62	71	67	77	76	68	71	65	70
	2011	82	80	81	69	69	76	74	70	71	77	72	74
Vast bedrag	2016	6	6	4	13	12	4	6	9	11	6	6	7
	2011	5	9	6	8	13	8	7	13	14	7	9	10
Onkosten	2016	2	4	3	2	0	1	2	4	2	1	1	1
	2011	4	2	3	2	2	2	4	4	2	3	1	1
Attentie	2016	0	0	1	0	1	0	0	2	1	1	1	0
	2011	1	1	0	2	0	0	1	1	1	0	0	1
Jaartaak	2016	11	15	14	17	12	18	12	10	16	18	21	20
	2011	8	11	9	19	15	15	13	12	14	13	18	15

Schoolbeleid voor de tweede correctie

De antwoorden op de vraag naar de facilitering van de tweede correctie zijn samengevat in tabel 3.2.3. In 2016 is ruim de helft (52%) van de docenten van mening dat de tweede correctie op hun school niet gefaciliteerd wordt. Van 6% zijn de uren voor de tweede correctie aanwijsbaar opgenomen in de jaartaak (bijv. in de opslagfactor), 31% krijgt een vast bedrag per examenwerk, 3% kan onkosten declareren (bijv. telefoonkosten) en 1% ontvangt een kleine attentie (bijv. een boekenbon).

Tabel 3.2.3 De wijze waarop scholen de tweede correctie faciliteren per jaar

	Jaar	N	Gemiddeld percentage
De eerste correctie wordt op mijn school niet gefaciliteerd	2016	3247	52
	2011	3460	48
Ik krijg een vast bedrag per examenwerk	2016	3247	31
	2011	3460	39
Ik kan onkosten declareren (bijv. telefoonkosten)	2016	3247	3
	2011	3460	4
Ik ontvang een kleine attentie (bijv. een boekenbon)	2016	3247	1
	2011	3460	2
De uren voor de eerste correctie zijn aanwijsbaar opgenomen in mijn jaartaak (bijv. in de opslagfactor)	2016	3247	6
	2011	3460	5

Verschillen tussen 2016 en 2011

Tabel 3.2.4 geeft informatie over hoe docenten van verschillende vakken het schoolbeleid met betrekking tot de tweede correctie in 2016 en 2011 typeren. De verschillen tussen de jaren zijn bij twee van de vijf aspecten van het schoolbeleid significant ($p < 0,010$). In vergelijking met 2011 zijn docenten in 2016 vaker van mening dat de tweede correctie op hun school niet gefaciliteerd wordt en krijgen zij minder vaak een vast bedrag per examenwerk. De verschillen bij de beleidsmaatregelen 'onkosten declareren', 'kleine attentie' en 'jaartaak' zijn statistisch gezien betekenisloos.

Tabel 3.2.4 De wijze waarop scholen de tweede correctie faciliteren per examen per jaar (percentages docenten per examen per jaar)

	Jaar	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WIG T	EC GT	BE GT
Geen faciliteiten	2016	57	51	55	41	52	48	57	53	48	52	51	48
	2011	55	48	48	55	43	42	53	57	45	48	49	49
Vast bedrag	2016	26	32	32	43	35	29	30	36	33	30	31	36
	2011	34	44	37	35	44	47	32	33	41	37	37	41
Onkosten	2016	5	5	3	0	1	2	4	6	2	2	1	3
	2011	6	3	6	3	3	4	4	7	3	3	4	1
Attentie	2016	2	2	1	4	2	2	1	1	1	1	1	0
	2011	2	2	2	3	2	1	2	0	1	2	1	0
Jaartaak	2016	4	4	3	4	7	9	4	2	8	8	9	8
	2011	3	4	4	5	5	5	4	4	6	5	8	7

Verschillen tussen examens

Tabel 3.2.4 laat zien hoe docenten van verschillende vakken het schoolbeleid met betrekking tot de tweede correctie in 2016 en 2011 typeren. In 2016 zijn de verschillen tussen examens bij drie van de vijf aspecten van de facilitering significant ($p < 0,01$). Het betreft de afwezigheid van de faciliteiten voor de tweede correctie, het toewijzen van uren aan de jaartaak en het toekennen van een vast bedrag per examenwerk. Zo zijn docenten Nederlands en geschiedenis vaker van mening dat er op hun school geen faciliteiten zijn voor de tweede correctie dan docenten muziek, beeldende vakken en Engels. Ook zijn de uren voor de tweede correctie van docenten scheikunde en economie vaker in de jaartaak opgenomen dan van docenten tehatex en aardrijkskunde. Docenten tehatex kunnen vaker onkosten declareren dan docenten muziek.

Verschillen tussen het schoolbeleid voor de eerste en tweede correctie

Scholen voeren met betrekking tot de eerste en tweede correctie een verschillend beleid (zie tabel 3.2.5). Zowel voor 2016 als 2011 geldt dat scholen de tweede correctie sterker faciliteren dan de eerste correctie (2016: 48% versus 29%). Voor de tweede correctie ontvangen docenten vaker een vast bedrag per examenwerk dan voor de eerste correctie (2016: 31% versus 7%). Daar staat tegenover dat de uren voor de tweede correctie minder vaak in de jaartaak zijn opgenomen dan die voor de eerste correctie (2016: 6% versus 16%). Het vergoeden van onkosten (2016: 3% versus 2%) en het verstrekken van een kleine attentie (1% versus 1%) zijn bij de tweede correctie ongeveer even zeldzaam als bij de eerste correctie.

Tabel 3.2.5 *Verschillen tussen het schoolbeleid voor de eerste en tweede correctie (percentages docenten per examen per jaar)*

	Jaar	Eerste correctie	Tweede correctie	Vershil
Geen faciliteiten	2016	71	52	19
	2011	75	48	27
Vast bedrag	2016	7	31	-24
	2011	9	39	-30
Onkosten	2016	2	3	-1
	2011	2	4	-2
Attentie	2016	1	1	0
	2011	1	2	-1
Jaartaak	2016	16	6	10
	2011	13	5	8

Verschillen tussen 2016 en 2011

Vergelijken we de beide jaren (zie tabel 3.2.5) dan lijkt het verschil tussen de facilitering van de eerste en tweede correctie iets kleiner te zijn geworden. In de beleving van de docenten wordt de eerste correctie nu iets beter gefaciliteerd dan toen, maar bij de tweede correctie lijkt er sprake van een kleine verslechtering (het verschil daalt van 27% in 2011 naar 19% in 2016). Ook liggen de percentages docenten die voor de eerste en tweede correctie een vast bedrag per examenwerk ontvangen nu dichter bij elkaar dan toen (het verschil tussen beide percentages daalde van 30% in 2011 naar 24% in 2016). Dit neemt overigens niet weg dat in 2016 minder docenten hiervan gebruik maakten dan in 2011 (van 39% naar 31%).

Overigens is in de vragenlijst alleen gevraagd naar expliciete faciliteringsmaatregelen. Veel scholen kennen de stilzwijgende afspraak dat docenten de correctie uitvoeren in de tijd die vrijkomt doordat de ingeroosterde lessen in examenklassen bij aanvang van het CSE niet meer gegeven worden. In het onderzoek is niet nagegaan in hoeverre de vrijgekomen tijd toereikend is en daadwerkelijk aan de correctie besteed kan worden.

3.2.2 Toelichtingen van docenten bij de vragen naar de facilitering van de eerste en tweede correctie

In totaal 506 docenten maakten gebruik van de mogelijkheid om hun antwoord op de matrix-vraag naar de facilitering van de eerste correctie toe te lichten en 548 docenten deden dat (ook) voor de overeenkomstige vraag naar de facilitering van de tweede correctie. De toelichtingen zijn gecodeerd met behulp van een categorieënsysteem. Voor het uitgebreide verslag van de inhoudsanalyse wordt verwezen naar Loeffen en Eikelboom (in voorbereiding). In totaal 46 respondenten verwezen in hun toelichting bij de vraag naar de facilitering van de tweede correctie naar hun eerder gegeven toelichting bij de vraag naar de facilitering van de eerste correctie (bijvoorbeeld: 'Zie mijn antwoord bij vraag 6'). In deze gevallen is de toelichting vervangen door het eerder gegeven antwoord. De resultaten van de analyse zijn weergegeven in tabel 3.2.6

Tabel 3.2.6 Eerste resultaten van de inhoudsanalyse van de toelichtingen bij de matrixvragen naar de facilitering van de eerste en tweede correctie (kolompercentages optellend tot 100%)

Faciliteringsmaatregel	Eerste correctie	Tweede correctie	Totaal
Geen faciliteiten	9	8	8
Vast bedrag			
Vast 'overall' bedrag zonder specificatie van dat bedrag	2	11	7
Vast 'overall' bedrag met aanduiding van dat bedrag	2	10	6
Vast bedrag per kandidaat zonder specificatie van dat bedrag	0	7	4
Vast bedrag per kandidaat met specificatie van alleen het bedrag per werk	0	2	1
Vast bedrag per kandidaat met specificatie van bedrag per werk en/of het maximum bedrag	3	19	11
Vast bedrag per klas/niveau/schoolsoort/leerlaag zonder specificatie van het bedrag	1	1	1
Vast bedrag per klas/niveau/schoolsoort/leerlaag met specificatie van het bedrag	0	4	2
Onkosten declareren			
Vergoeding reiskosten/telefoonkosten/onkosten	1	1	1
Kleine attentie			
Etentje met collega's	1	1	1
Bedrag voor het sectiepotje/de vakgroep	1	1	1
Jaartaak en opslagfactor			
Uren zitten deels in jaartaak/opslagfactor, maar niet aanwijsbaar	8	3	5
Wel gefaciliteerd, maar niet extra	43	14	29
Hogere opslagfactor/opslagfactor laatste leerjaar/betere lesurenvergoeding	1	0	1
Vast aantal uren/vast aantal uren of minuten per werk	4	1	2
Andere faciliteringsmaatregelen dan genoemd in de matrixvraag zelf			
Niet of minder surveilleren of geen 'verzorgende taken'	2	0	1
Uitgeroosterd, correctie- of nakijkverlof, verlofdagen, een middag geen vergadering	5	1	3
Overig			
Weet niet/geen mening	12	11	11
Overig on-topic (wel antwoord op de vraag, maar niet codeerbaar)	1		1
Off-topic (geen antwoord op de vraag)	3	3	3
Totaal	100	100	100

De open vragen hadden tot doel om de respondenten ertoe te bewegen andere dan de in de vraagstelling beschreven faciliteringsmaatregelen te noemen.

De toelichtingen van de docenten bevatten in verreweg de meeste gevallen een parafraze op of een verbijzondering van een maatregel die reeds in de vraagstelling genoemd was (dit wil zeggen: een vast bedrag, onkosten declareren, een kleine attentie en facilitering middels jaartaak en opslagfactor). Echt nieuwe suggestie werden bijna niet gegeven. Met betrekking tot de facilitering van de eerste correctie verwees 2% van de docenten naar 'niet of minder surveilleren of geen verzorgende taken' en 5% naar 'Uitgeroosterd, correctie- of nakijkverlof,

verlofdagen, een middag geen vergadering'; voor de tweede correctie waren de overeenkomstige percentages 0% en 1%.

Deze gegevens komen overeen met de ervaringen die zijn opgedaan in het kader van het digitale Meldpunt Correctie Centrale Examens waar scholen voorbeelden van goed georganiseerde correctie konden melden. Ook daar kwamen weinig tips ter verbetering van de facilitering van de correctie binnen.

De toelichtingen van de docenten geven ook meer zicht op de verschillen tussen de facilitering van de eerste en tweede correctie. Aan de hand van de gestandaardiseerde residuen van een uitgevoerde Chi-kwadraattoetsing constateren we de volgende overeenkomsten en verschillen:

- Docenten krijgen vaker een vast bedrag per kandidaat voor de tweede correctie dan voor de eerste correctie. Dit geldt voor zes van de zeven typen vaste bedragen die in tabel 3.2.6 genoemd zijn. Overigens gaat het hier in vrijwel alle gevallen om een luttel bedrag.
- De eerste correctie wordt even weinig als de tweede correctie gefaciliteerd door onkosten te vergoeden of een kleine attentie te geven;
- De eerste correctie wordt veel vaker dan de tweede correctie gefaciliteerd via de jaartaak en de opslagfactor; en van de eerste correctie vinden meer docenten dat deze wel gefaciliteerd maar niet extra gefaciliteerd wordt dan van de tweede correctie;
- De niet in de vraagstelling genoemde maatregelen 'Niet of minder surveilleren of geen verzorgende taken' worden bij de eerste correctie wat vaker toegepast dan bij de tweede correctie. Dit geldt ook voor 'uitroosteren, correctieverlof, verlofdagen, nakijkdagen en een middag geen vergadering'.

3.2.3 Tevredenheid over het schoolbeleid

De respondenten hebben aangegeven hoe tevreden zij zijn met de manier waarop hun school de eerste en tweede correctie faciliteert.

Tevredenheid met het schoolbeleid voor de eerste correctie

In 2016 is 43% tevreden noch ontevreden met de manier waarop de school de eerste correctie faciliteert, 1% is zeer tevreden, 16% is tevreden, 25% is ontevreden en 14% is zeer ontevreden.

Tabel 3.2.7 *Tevredenheid met de facilitering van de eerste correctie (bron: eerste corrector; kolompercentages optellend tot 100%)*

	2016	2011
Zeer tevreden	1	1
Tevreden	16	12
Noch tevreden noch ontevreden	43	48
Ontevreden	25	26
Zeer ontevreden	14	13

Verschillen tussen 2016 en 2011

De ordinale logistische regressieanalyse laat zien dat de verschillen tussen 2016 en 2011 niet significant zijn ($p = 0,309$). In 2016 zijn docenten even (on)tevreden over de facilitering van de eerste correctie als in 2011.

Verschillen tussen examens

De verschillen tussen examens zijn in 2016 significant ($p < 0,001$). Zo zijn docenten Nederlands en geschiedenis minder tevreden over het schoolbeleid met betrekking tot de eerste correctie dan docenten Duits en Engels (zie tabel 3.2.8).

Tabel 3.2.8 Tevredenheid met de facilitering van de eerste correctie (bron: eerste corrector; kolompercentages per examen per jaar optellend tot 100%)

Jaar	Tevredenheid	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
2016	Zeer tevreden	1	1	0	4	3	2	0	1	2	2	2	1
	Tevreden	7	11	14	26	17	14	11	10	22	22	18	18
	Noch tevreden noch ontevreden	29	47	41	43	55	44	28	40	46	46	45	49
	Ontevreden	32	30	29	20	16	28	36	33	19	17	26	22
	Zeer ontevreden	31	11	16	7	9	13	24	16	11	13	9	9
2011	Zeer tevreden	0	2	0	3	2	1	0	0	2	1	3	1
	Tevreden	7	10	10	15	17	13	9	9	16	13	16	14
	Noch tevreden noch ontevreden	38	48	42	51	48	50	43	55	49	54	47	55
	Ontevreden	34	25	34	23	23	22	26	27	24	24	25	24
	Zeer ontevreden	22	16	14	8	10	14	22	8	10	8	10	6

Tevredenheid met het schoolbeleid voor de tweede correctie

In 2016 is een derde van de docenten (32%) tevreden noch ontevreden met de wijze waarop de school de tweede correctie faciliteert, 2% is zeer tevreden, 16% is tevreden, 31% is ontevreden en 20% is zeer ontevreden (zie tabel 3.2.9).

Tabel 3.2.9 Tevredenheid met facilitering van de tweede correctie per jaar (bron: eerste corrector; kolompercentages optellend tot 100%)

	2016	2011
Zeer tevreden	2	1
Tevreden	16	15
Noch tevreden noch ontevreden	32	34
Ontevreden	31	31
Zeer ontevreden	20	19

Verschillen tussen 2016 en 2011

De verschillen tussen de (on)tevredenheid over het schoolbeleid in 2016 en 2011 zijn niet significant ($p = 0,856$). In 2016 zijn docenten even (on)tevreden over de manier waarop de schoolleiding de tweede correctie faciliteert als in 2011.

Verschillen tussen examens

De verschillen tussen examens zijn in 2016 significant ($p < 0,001$). Zo zijn docenten Nederlands en geschiedenis minder tevreden met het schoolbeleid met betrekking tot de tweede correctie dan docenten Duits en Engels (zie tabel 3.2.10).

Tabel 3.2.10 *Tevredenheid over de facilitering van de tweede correctie per examen jaar (bron: eerste corrector; kolompercentages per examen per jaar optellend tot 100%)*

Jaar	Tevredenheid	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WIGT	EC GT	BE GT
2016	Zeer tevreden	1	1	0	2	3	1	1	1	4	3	2	0
	Tevreden	8	11	12	24	22	12	11	11	21	21	17	22
	Noch tevreden noch ontevreden	20	32	31	47	36	27	22	29	39	33	35	38
	Ontevreden	38	37	30	16	25	38	37	38	23	28	31	27
	Zeer ontevreden	33	19	26	12	14	22	30	21	13	15	15	13
2011	Zeer tevreden	0	1	1	3	2	1	1	0	2	1	2	0
	Tevreden	9	11	10	8	22	14	9	9	19	16	21	19
	Noch tevreden noch ontevreden	32	36	32	49	35	34	29	43	33	35	30	40
	Ontevreden	32	34	36	28	27	29	32	36	29	31	31	31
	Zeer ontevreden	27	18	21	11	15	23	28	13	18	17	16	10

Verskil tussen de tevredenheid met het schoolbeleid voor de eerste en tweede correctie

Zowel in 2016 als in 2011 zijn docenten minder tevreden met het schoolbeleid voor de tweede correctie dan met dat voor de eerste correctie (zie tabel 3.2.11). In de voor beide jaren afzonderlijk uitgevoerde ordinale logistische regressieanalyse geeft de toetsing van het regressiegewicht voor het verschil tussen de tevredenheid met de eerste en tweede correctie namelijk significantie te zien (beide $p < 0,001$).

Tabel 3.2.11 *Verskil tussen de tevredenheid met het schoolbeleid voor de eerste en tweede correctie per jaar (kolompercentages optellend tot 100%)*

	Eerste correctie		Tweede correctie	
	2016	2011	2016	2011
Zeer tevreden	1	1	2	1
Tevreden	16	12	16	15
Noch tevreden noch ontevreden	43	48	32	34
Ontevreden	25	26	31	31
Zeer ontevreden	14	13	20	19

3.3 Volledigheid van de tweede correctie

Tweede correctoren moeten het werk van de kandidaten integraal nakijken. Dit wil zeggen dat de tweede corrector alle werken nakijkt en per kandidaat het volledige examenwerk met alle vragen. In deze paragraaf gaan we na in hoeverre docenten de tweede correctie volgens de wettelijke voorschriften uitvoeren. De volledigheid van de correctie is vastgesteld door docenten naar hun mening te vragen. Uit onderzoek is bekend dat zelfs kleine verschillen in de vraagstelling van enquêtevragen tot sterk verschillende antwoorden kunnen leiden (o.a. Krosnick, 1999; Kuhlemeier, Hemker & Van den Berg, 2013). Daarnaast blijken er vrijwel altijd respondenten te zijn die de vraag en/of de antwoordmogelijkheden anders interpreteren dan anderen (ook al is de vraag nog zo zorgvuldig geconstrueerd). Daarom hebben wij ervoor

gekozen om de volledigheid van de tweede correctie op verschillende manieren te bevragen. In 2016 hebben de docenten zes vragen beantwoord over de volledigheid van de tweede correctie:

- 1 eerste correctoren is gevraagd of hun examenwerk al dan niet door een tweede corrector is nagekeken (paragraaf 3.3.1);
- 2 eerste correctoren is gevraagd naar de volledigheid van de tweede correctie (paragraaf 3.3.2);
- 3 tweede correctoren is gevraagd naar de volledigheid van hun eigen tweede correctie (paragraaf 3.3.3, 3.3.4 en 3.3.6);
- 4 tweede correctoren is gevraagd in hoeverre zij een niet-integrale werkwijze toepassen waarbij slechts een deel van de kandidaten en/of vragen gecorrigeerd wordt (paragraaf 3.3.5);
- 5 tweede correctoren is gevraagd naar het percentage kandidaten waarvoor zij een of meer examenvragen nakeken (paragraaf 3.3.7);
- 6 tweede correctoren is gevraagd naar het percentage examenvragen waarvoor zij een of meer kandidaten nakeken (paragraaf 3.3.7).

De vragen 1, 2, 3, 5 en 6 zijn ook in het vorige onderzoek gesteld, zodat we de volledigheid anno 2016 kunnen vergelijken met die van 2011.

Meer zekerheid over de volledigheid waarmee de tweede correctie wordt uitgevoerd kan worden verkregen door de antwoorden op de afzonderlijke vragen met elkaar in verband te brengen. Van deze exercitie wordt verslag gedaan in paragraaf 3.3.8.

Aan het einde van deze sectie bespreken we hoe docenten denken over de noodzaak van integrale tweede correctie (paragraaf 3.3.9), de oorzaken van onvolledige tweede correctie (paragraaf 3.3.10 en 3.3.11) en de wenselijkheid van (vermeende) verbeteringsmaatregelen (paragraaf 3.3.12 en 3.3.13).

3.3.1 Wel of geen tweede correctie

Volgens opgave van de eerste corrector is het werk van de kandidaten in 0,5% van de gevallen niet door een tweede corrector nagekeken. In 2011 was dat 0,6%. Zoals blijkt uit de ordinale logistische regressieanalyse is dit verschil statistisch gezien betekenisloos ($p = 0,683$).

Tabel 3.3.1 toont het percentage door een tweede corrector nagekeken examenwerken per examen per jaar. In 2016 zijn de verschillen tussen examens niet significant op 1%-niveau ($p = 0,013$). De tweede correctie wordt bij het ene examen niet minder vaak uitgevoerd dan bij het andere examen.

Tabel 3.3.1 Percentage door tweede corrector nagekeken examenwerken (bron: eerste corrector)

Jaar	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BEGT
2016	100	100	100	100	99	100	100	100	99	99	99	97
2011	100	100	100	100	100	100	100	98	99	99	99	97

3.3.2 Volledigheid van de tweede correctie volgens de eerste corrector

De eerste correctoren is gevraagd naar hun mening over de mate waarin de tweede corrector volledig heeft nagekeken (zie tabel 3.3.2). In 2016 heeft 44% van de tweede correctoren zeer volledig nagekeken, 41% volledig, 10% min of meer volledig, 1% onvolledig, 1% zeer onvolledig en 4% koos voor 'Ik kan niet overzien hoe volledig de tweede corrector heeft nagekeken'.

Tabel 3.3.2 De volledigheid van de tweede correctie (bron: eerste corrector; kolompercentages optellend tot 100%)

Volledigheid	2016	2011
Zeer volledig (d.w.z. alle werken en per kandidaat alle vragen)	44	31
Volledig	41	39
Min of meer volledig	10	18
Onvolledig	1	2
Zeer onvolledig (bijv. enkele vragen van enkele kandidaten)	1	1
Ik kan niet overzien hoe volledig de tweede corrector heeft nagekeken	4	9

Verschillen tussen 2016 en 2011

De significantietoetsing van de verschillen tussen jaren en examens is uitgevoerd nadat de docenten die geen antwoord konden geven buiten beschouwing zijn gelaten. De toetsing van de verschillen tussen jaren geeft significantie te zien ($p < 0,001$). In 2016 wordt de tweede correctie volgens de eerste correctoren vollediger uitgevoerd dan in 2011.

Verschillen tussen examens

Zoals blijkt uit de ordinale logistische regressieanalyse zijn de verschillen tussen examens in 2016 significant ($p < 0,001$). Zo kijken tweede correctoren Nederlands, aardrijkskunde en geschiedenis volgens de eerste correctoren minder volledig na dan onder meer tweede correctoren biologie, scheikunde en Duits (zie tabel 3.3.3).

Tabel 3.3.3 De volledigheid van de tweede correctie per examen per jaar (bron: eerste corrector; kolompercentages per examen per jaar optellend tot 100%)

Jaar	Volledigheid	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
2016	Zeer volledig	44	61	43	54	52	56	30	47	41	47	38	44
	Volledig	42	31	42	43	43	34	50	30	48	45	47	47
	Min of meer volledig	12	6	14	4	5	9	16	18	10	7	13	6
	Onvolledig	2	1	1	0	0	1	2	2	2	1	1	1
	Zeer onvolledig	0	0	0	0	0	0	1	2	1	0	1	3
2011	Zeer volledig	24	42	25	57	47	35	21	28	37	35	34	39
	Volledig	41	41	41	36	42	42	44	41	43	45	44	45
	Min of meer volledig	30	14	28	7	10	21	28	26	17	16	17	15
	Onvolledig	3	3	2	0	1	1	5	2	2	2	2	1
	Zeer onvolledig	2	0	4	0	1	0	3	3	0	2	2	0

3.3.3 Volledigheid van de tweede correctie volgens de tweede corrector

In de vragenlijst zijn de tweede correctoren drie stellingen over de tweede correctie voorgelegd. Elke stelling beschrijft een correctiestrategie. Bij elke vraag kozen de respondenten uit de antwoordmogelijkheden 'Klopt helemaal', 'Klopt', 'Klopt niet' en 'Klopt helemaal niet'.

De eerste stelling beschrijft de situatie waarin de tweede corrector het examenwerk geheel in overeenstemming met de wettelijke voorschriften integraal nakijkt. De eerste stelling was als volgt geformuleerd: 'Ik heb het examenwerk van alle kandidaten volledig nagekeken'.

De verdeling van de antwoorden van de tweede correctoren zijn beschreven in tabel 3.3.4, uitgesplitst naar jaar. Van de tweede correctoren kan 72% zich in 2016 helemaal met deze stelling vereenzelvigen. Bijna een kwart (23%) van de tweede correctoren koos voor de antwoordmogelijkheid 'Klopt', 5% voor 'Klopt niet' en 0% voor 'Klopt helemaal niet'.

Tabel 3.3.4 De volledigheid van de tweede correctie (bron: tweede corrector: kolompercentages optellend tot 100%)

	2016	2011
Klopt helemaal	72	44
Klopt	23	32
Klopt niet	5	20
Klopt helemaal niet	0	4

Verschillen tussen 2016 en 2011

In 2011 is de tweede correctoren nagenoeg dezelfde stelling voorgelegd: 'Ik heb het examenwerk van alle kandidaten helemaal nagekeken (als ware het een eerste correctie)'. We mogen de resultaten van nu dus vergelijken met die van toen. De ordinale logistische regressieanalyse geeft aan dat de tweede correctie volgens de tweede correctoren zelf in 2016 vollediger wordt uitgevoerd dan in 2011 ($p < 0,001$). De verschillen tussen de beide jaren zijn groot. Vond in 2011 nog maar 44% van de tweede correctoren dat de eerste correctie helemaal volledig was uitgevoerd, in 2016 was dit percentage gestegen naar 72%.

Verschillen tussen examens

Tabel 3.3.5 toont de verdeling van de antwoorden per examen per jaar. Volgens de ordinale logistische regressieanalyse zijn de verschillen tussen examens in 2016 significant ($p < 0,001$). Zo kijken tweede correctoren geschiedenis in 2016 naar eigen zeggen minder volledig na dan tweede correctoren Duits.

Tabel 3.3.5 De volledigheid van de tweede correctie per examen per jaar (bron: tweede corrector: kolompercentages per examen per jaar optellend tot 100%)

Jaar	Volledigheid	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WIG T	EC GT	BE GT
2016	Klopt helemaal	67	79	59	86	83	78	48	76	75	79	70	73
	Klopt	26	18	31	12	16	20	35	17	22	18	26	24
	Klopt niet	7	3	10	2	1	2	16	7	2	2	5	3
	Klopt helemaal niet	1	0	0	0	0	1	1	1	0	0	0	0
2011	Klopt helemaal	30	44	31	62	64	43	22	43	56	50	43	56
	Klopt	31	35	31	32	26	33	25	34	31	37	38	29
	Klopt niet	33	20	31	7	9	17	41	20	13	11	17	11
	Klopt helemaal niet	6	2	7	0	1	6	13	3	1	2	2	3

3.3.4 Niet-integrale tweede correctie

De tweede stelling beschrijft een correctiestrategie waarbij alleen een deel van de werken en/of vragen wordt nagekeken. De precieze formulering was: "Ik heb eerst een deel van de werken/vragen nagekeken en ben daarna gestopt". Deze stelling is alleen in 2016 aan docenten

voorgelegd. Het percentage ontbrekende antwoorden was bij deze vraag opvallend hoog (16%). Hierna bespreken we de antwoorden van degenen die de vraag wel beantwoordden. Drie procent van hen herkent zich helemaal in deze niet-integrale correctiestrategie en kiest voor 'klopt helemaal'. De overige respondenten kozen voor 'klopt' (4%), 'klopt niet' (17%) of 'klopt helemaal niet' (77%).

Tabel 3.3.6 Niet-integrale tweede correctie (kolompercentage optellend tot 100%)

	%
Klopt helemaal	3
Klopt	4
Klopt niet	17
Klopt helemaal niet	77

Tabel 3.3.7 toont de verdeling per examen. De verschillen tussen examens zijn significant ($p < 0,001$). Zo kiezen tweede correctoren geschiedenis vaker voor deze vorm van niet-integrale correctie dan tweede correctoren biologie.

Tabel 3.3.7 Niet-integrale tweede correctie per examen (kolompercentages optellend tot 100%)

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Klopt helemaal	2	1	3	0	3	3	2	3	4	3	2	6
Klopt	5	1	5	2	3	2	10	3	6	1	4	6
Klopt niet	18	11	22	22	19	12	27	12	17	11	18	15
Klopt helemaal niet	75	87	70	76	75	83	62	82	72	84	77	73

3.3.5 Gefaseerde volledige tweede correctie

De derde stelling beschrijft een strategie waarbij de tweede corrector eerst een deel van de werken of vragen nakijkt en vervolgens alsnog alle werken/vragen volledig nakijkt. De precieze formulering was: "Ik heb eerst een deel van de werken/vragen nagekeken en daarna heb ik alsnog alle werken/vragen volledig nagekeken (omdat de resultaten van het eerste deel daartoe aanleiding gaven)". Het percentage ontbrekende antwoorden was bij deze vraag opvallend hoog (17%). Hierna bespreken we de antwoorden van degenen die de vraag wel beantwoordden. Getuige hun antwoord 'Klopt helemaal niet' herkent twee derde (66%) van de tweede correctoren zich helemaal niet in deze stelling, een vijfde (20%) antwoordt met 'Klopt niet', 8% met 'Klopt' en 6% met 'Klopt helemaal niet'.

Tabel 3.3.8 Gefaseerde volledige tweede correctie (kolompercentage optellend tot 100%)

	%
Klopt helemaal	6
Klopt	8
Klopt niet	20
Klopt helemaal niet	66

Tabel 3.3.9 toont de verdeling per examen. De verschillen tussen examens zijn significant ($p < 0,001$). Zo beoordelen tweede correctoren geschiedenis en Nederlands vaker eerst een deel om vervolgens alsnog alle werken en vragen na te kijken dan bijvoorbeeld tweede correctoren biologie en wiskunde.

Tabel 3.3.9 Gefaseerde volledige tweede correctie per examen (kolompercentages optellend tot 100%)

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Klopt helemaal	11	2	9	7	2	5	11	12	5	4	6	9
Klopt	8	7	12	4	6	5	15	11	6	5	8	9
Klopt niet	23	13	23	17	20	16	25	22	26	15	20	18
Klopt helemaal niet	58	78	56	72	72	75	49	55	62	76	65	65

3.3.6 Toelichtingen van docenten bij de stellingvragen naar de volledigheid van de tweede correctie

De drie stellingen geven een globale en zeker geen uitputtende beschrijving van de strategieën die docenten bij de tweede correctie toepassen. Aan het einde van de matrixvraag met de drie stellingen is docenten daarom de ruimte geboden om hun reactie op de stellingen toe te lichten. In totaal 280 tweede correctoren hebben van deze mogelijkheid gebruik gemaakt. De inhoud van de toelichtingen is geanalyseerd met behulp van een categorieënsysteem. Voor het uitgebreide verslag van de inhoudsanalyse wordt verwezen naar Loeffen en Eikelboom (in voorbereiding). De resultaten van de analyse zijn weergegeven in tabel 3.3.10. De constructie van deze tabel behoeft enige toelichting. De toelichtingen in deze tabel zijn uitgesplitst naar het antwoord op de eerste stellingvraag naar de volledigheid van de tweede correctie. De vier antwoordmogelijkheden van deze stellingvraag zijn daarbij tot twee teruggebracht door de antwoorden 'Klopt helemaal' en 'Klopt' samen te nemen evenals de antwoorden 'Klopt niet' en 'Klopt helemaal niet'. De eerste kolom van tabel 3.3.10 toont daarmee de verdeling van de toelichtingen van de groep die naar eigen zeggen (helemaal) volledig nakeek (74%), de tweede kolom de verdeling in de groep die (helemaal) niet volledig nakeek (26%) en de derde kolom de verdeling van alle toelichtingen (100%). In de totale groep van docenten die aan het onderzoek deelnam, waren de overeenkomstige percentages 95% tegenover 5% (zie tabel 3.3.2).

In de bespreking van de toelichtingen besteden we eerst aandacht aan de diversiteit van de correctiestrategieën die tweede correctoren zoal hanteren. Dat doen we aan de hand van de percentages in de kolom 'Totaal'. Daarna bespreken we eerst de toelichtingen van degenen die bij de eerste stellingvraag aangeven de tweede correctie (helemaal) volledig uit te voeren. Dat doen we aan de hand van de percentages in de kolom '(Helemaal) volledig'. Tot slot bekijken we de toelichtingen van degenen die bij de eerste stellingvraag aangaven de tweede correctie (helemaal) niet volledig uit te voeren. Dat doen we aan de hand van de percentages in de kolom '(Helemaal) niet volledig'.

De vergelijking van de antwoordpatronen van beide groepen geeft meer inzicht in de manier waarop de docenten de eerste stellingvraag geïnterpreteerd hebben en wat zij precies verstaan onder een volledig uitgevoerde tweede correctie. Vanwege het grote aantal zeer kleine waarden zijn de percentages in tabel 3.3.10 op één decimaal nauwkeurig weergegeven.

Tabel 3.3.10 Verdeling van de toelichtingen bij de drie stellingvragen over de volledigheid van de tweede correctie voor (helemaal) volledig en (helemaal) niet volledig corrigerende tweede correctoren (kolompercentages optellend tot 100%)

	Antwoord op de stellingvraag naar de volledigheid van de tweede correctie		Totaal
	(Helemaal) volledig	(Helemaal) niet volledig	
Volledig nagekeken			
Alle kandidaten en alle vragen volledig nagekeken	62,0	4,5	46,9
Open vragen nagekeken en van de mk-vragen eerst een deel en daarna alsnog alle mk-vragen	1,1		0,8
Kandidaten of vragen volledig nagekeken, maar globaal of minder zorgvuldig			
Eerst een deel van de kandidaten nagekeken en daarna de overige kandidaten globaal of minder zorgvuldig	3,7	3,0	3,5
Eerst een deel van de vragen nagekeken en daarna de overige vragen globaal of minder zorgvuldig	3,7	7,5	4,7
Open vragen en meerkeuzevragen			
Alleen open vragen nagekeken en mk-vragen niet	4,3	4,5	4,3
Open vragen nagekeken en een deel van de mk-vragen	4,3	10,4	5,9
Open vragen nagekeken en mk-vragen globaler of minder zorgvuldig	2,7	1,5	2,4
Open en meerkeuzevragen (on)volledig nagekeken			
Open vragen onvolledig en mk-vragen onvolledig nagekeken	0,5	9,0	2,8
Open vragen niet volledig en mk-vragen wel volledig nagekeken	0,0	0,0	0,0
Eerst of alleen een deel van de kandidaten nagekeken			
Alleen een deel van de kandidaten nagekeken (verder ongespecificeerd)	0,5	4,5	1,6
Eerst een deel van de kandidaten en daarna een deel van de vragen nagekeken	1,1	3,0	1,6
Eerst een deel van de kandidaten en daarna alleen de lastige, opvallende, problematische of discutabele vragen nagekeken	0,5	17,9	5,1
Eerst of alleen een deel van de vragen nagekeken			
Alleen een deel van de vragen nagekeken (verder ongespecificeerd)	1,6	14,9	5,1
Eerst een deel van de vragen en daarna een deel van de overige vragen nagekeken	1,1	3,0	1,6
Eerst een deel van de vragen en daarna een deel van de kandidaten nagekeken		1,5	0,4
Eerst een deel nagekeken (en verder ongespecificeerd)			
Eerst een deel nagekeken (ongespecificeerd; dus onduidelijk of van kandidaten of van vragen)		9,0	2,4
Niet alle kandidaten en niet alle vragen nagekeken			
Zowel een horizontale steekproef (per kandidaat) als een verticale steekproef (per vraag) nagekeken		3,0	0,8
Overig			
Overig on-topic (wel antwoord op de vraag, maar niet codeerbaar)	3,2		2,4
Off-topic (geen antwoord op de vraag)	9,6	3,0	7,9
Totaal	100,0	100,0	100,0

Correctiestrategieën van tweede correctoren

De percentages in de kolom 'Totaal' geven inzicht in de diversiteit van de gehanteerde correctiestrategieën:

- 47% van de 280 docenten heeft de open vraag benut om te benadrukken dat zij de tweede correctie helemaal volledig uitvoeren. Een karakteristieke toelichting is: *'Alles nagekeken als ware het de 1e correctie'*. Onder deze volledig corrigerende tweede correctoren bevinden zich enkelen die gewag maken van een strategie waarbij zij eerst de open vragen nakijken, daarna een deel van de meerkeuzevragen om tot slot alsnog alle meerkeuzevragen te corrigeren.
- Tot de groep docenten die mogelijk volledig nakijkt, rekenen we ook degenen die eerst een deel van de **kandidaten** nakijken om daarna de overige kandidaten wat globaler of minder zorgvuldig te corrigeren (4%) en degenen die eerst een deel van de **vragen** corrigeren en daarna de overige vragen en/of werken globaler en/of minder zorgvuldig nakijken (5%). Een typisch voorbeeld is: *'Eerst per vraag bij alle kandidaten uitgebreid nagekeken t/m vraag 10 toen bleek dat er goed nagekeken was. Daarna iets sneller doorgenomen (per 4 vragen)'*.
- 13% van de tweede correctoren kijkt wel alle open vragen na, maar de meerkeuzevragen niet of niet allemaal. Binnen deze groep zijn drie correctiewijzen te onderscheiden: 4% keek alleen de open vragen na, 6% corrigeerde de open vragen en een deel van de meerkeuzevragen en 3% keek de open vragen na en de meerkeuzevragen globaler of minder zorgvuldig. De laatsten checkten bijvoorbeeld alleen de puntentelling in Wolf in plaats van ook te controleren of de eerste corrector de meerkeuzevragen goed gescoord had (bijvoorbeeld de score 0 of 1 correct in de marge van het examenboekje noteerde). Een typische toelichting is: *'De meerkeuzevragen kijk ik steekproefsgewijs na / de open vragen allemaal'*.
- 4% meldt dat zij de open vragen niet volledig nakijken en doet in diezelfde toelichting ook een uitspraak over de volledigheid waarmee zij de meerkeuzevragen corrigeren. Deze groep is onderverdeeld in 3% die noch de open noch de meerkeuzevragen volledig nakijkt en 1% die de open vragen niet volledig nakijkt maar de meerkeuzevragen wel volledig. Bijvoorbeeld: *'Steekproef meerkeuzevragen. Steekproef open vragen plus problematische vragen'*.
- 2% corrigeert alleen of eerst een deel van de vragen en/of kandidaten zonder te specificeren of het open vragen of meerkeuzevragen betrof en zonder aan te geven of het ging om een deel van de vragen en/of een deel van de kandidaten (bijvoorbeeld: *'Ik heb steekproefsgewijs het examen nagekeken'*).
- 8% corrigeert alleen een deel van de **kandidaten** of keek eerst een deel van kandidaten na en daarna nog meer maar niet alle kandidaten of vragen (zonder aan te geven of het daarbij om open of meerkeuzevragen ging). De laatsten corrigeerden na de eerste selectie bijvoorbeeld alleen wat zij typeerden als lastige vragen, probleemvragen, vragen waarover discussie was (bijvoorbeeld op de internetfora) of vragen waarbij zij discussie met de eerste corrector verwachtten. Een voorbeeld is: *'Eerst een deel van de werken helemaal, toen vragen geselecteerd waarmee ik het oneens was. Die vragen bij alle leerlingen nagekeken. Dus 20% van de werken helemaal, rest op geselecteerde vragen'*.
- 7% corrigeert alleen een deel van de **vragen** of keek eerst een deel van vragen na en daarna nog meer maar niet alle vragen of alle kandidaten (zonder te specificeren of het daarbij om open of meerkeuzevragen ging). Een karakteristieke toelichting is: *'Enkele vragen helemaal bij iedereen nagekeken, zeer nauwkeurig. Enkele vragen steekproefsgewijs nagekeken. Merendeel 'normaal' nagekeken'*.
- 1% corrigeerde eerst zowel een deel van de kandidaten (horizontaal) als een deel van de vragen (verticaal) zonder precies aan te geven wat zij daarna deden. Bijvoorbeeld: *'Heb aantal nagekeken, beste/slechtste + aantal discutabele/lastige vragen'*.
- 2% van de toelichtingen bevatte een antwoord op de vraag, maar kon niet met het gebruikte categorieënsysteem geïnclassificeerd worden. Een voorbeeld: *'Ik maak gebruik van de notities in de kantlijn van de 1e correctie. Daarnaast ken je de antwoorden beter, omdat je een secure 1e correctie hebt gedaan'*.
- 8% van de antwoorden was off-topic in de zin dat het geen uitspraak over de volledigheid van de tweede correctie bevatte.

Tabel 3.3.10 geeft ook inzicht in de verschillen tussen de toelichtingen van degenen die bij de eerste stellingvraag aangaven de tweede correctie (helemaal) volledig uit te voeren en degenen die dat (helemaal) niet volledig deden. Door de toelichtingen per groep te bekijken, kunnen we de antwoorden op de eerste stellingvraag beter interpreteren en komen we meer te weten over wat docenten precies verstaan onder een volledig uitgevoerde tweede correctie. Hierna bespreken we eerst de toelichtingen van degenen die bij de eerste stellingvraag aangeven de tweede correctie (helemaal) volledig uit te voeren. Dat doen we aan de hand van de percentages in de kolom '(Helemaal) volledig'. Vervolgens bekijken we de toelichtingen van degenen die bij de eerste stellingvraag aangaven de tweede correctie (helemaal) niet volledig uit te voeren. Dat doen we aan de hand van de percentages in de kolom '(Helemaal) niet volledig'.

Correctiestrategieën van tweede correctoren die naar eigen zeggen (helemaal) volledig nakijken

Omdat afronding in dit geval tot onnauwkeurige uitkomsten zou leiden, presenteren we hierna de percentages op één of twee decimalen nauwkeurig. Van degenen die naar eigen zeggen de tweede correctie (helemaal) volledig uitvoeren, blijkt 63,1% (= 62,0% + 1,1%) dat volgens hun toelichtingen ook te doen. In hun toelichting bevestigen zij hun eerder gegeven antwoord dat zij alle kandidaten en alle vragen volledig nakijken.

Daarnaast is er een groep van 3,74% die eerst een deel van de kandidaten nakeek en daarna de overige kandidaten globaal of minder zorgvuldig corrigeerde. Dat zij alle kandidaten beoordeelden, betekent niet automatisch dat zij van deze kandidaten ook alle vragen hebben nagekeken. Bovendien kan globaal of minder zorgvuldig ook betekenen dat er kandidaten of vragen zijn overgeslagen. Omdat zij bij de stellingvraag te kennen gaven de tweede correctie (helemaal) volledig uit te voeren krijgen zij het voordeel van de twijfel en rekenen we hen toch tot de volledig nakijkende tweede correctoren. Hetzelfde doen we met de andere groep van 3,74% die eerst een deel van de vragen nakeek en daarna de overige vragen globaal of minder zorgvuldig corrigeerde. Hiermee komt de groep die waarschijnlijk helemaal volledig corrigeerde op 70,6% (= 63,1% + 3,74% + 3,74%).

De overige 29,4% valt uiteen in een groep die de tweede correctie volgens hun toelichtingen niet volledig uitvoert (16,6%) en een groep die een niet-codeerbare of anderszins onbruikbare toelichting gaf (12,8%).

De 16,6% die in eerste instantie naar eigen zeggen (helemaal) volledig corrigeert maar dat gezien hun toelichtingen toch niet lijkt te doen kan worden onderverdeeld in de volgende vier groepen:

- 11,23% keek de open vragen na en de meerkeuzevragen niet (4,28%), gedeeltelijk (4,28%) of globaal of minder zorgvuldig (2,67%);
- vrijwel niemand keek noch de open vragen noch de meerkeuzevragen volledig na (0,53);
- 2,14% keek niet alle kandidaten na;
- 2,67% keek niet alle vragen na.

Al met al blijkt (na afronding) 71% van degenen die volledig zeggen na te kijken dat volgens hun toelichtingen inderdaad te doen. Daarnaast is er een groep van 17% die de tweede correctie (helemaal) volledig uitvoert maar dat gezien hun toelichtingen toch niet gedaan heeft.

Deze groep valt uiteen in twee groepen: degenen die alleen (een deel van) de meerkeuzevragen niet nakijken (12%) en degenen die of niet alle kandidaten en/of niet alle vragen nakijken (5%).

Correctiestrategieën van tweede correctoren die naar eigen zeggen (helemaal) niet volledig nakijken

Vervolgens bekijken we de toelichtingen van degenen die bij de eerste stellingvraag aangaven de tweede correctie (helemaal) niet volledig uit te voeren. Uit de inhoudsanalyse van deze toelichtingen komt naar voren dat 4% toch alle vragen en alle kandidaten nakeek en 3% een

niet-buikbare toelichting gaf. De overige 93% voerde de tweede correctie volgens hun toelichting en in overeenstemming met hun antwoord op de eerste stellingvraag inderdaad niet volledig uit. Deze groep van 93% onvolledig nakijkende tweede correctoren kan worden onderverdeeld in de volgende groepen:

- 10% corrigeerde eerst een deel van de kandidaten of eerst een deel van de vragen om vervolgens de overige kandidaten of vragen globaal of minder zorgvuldig na te kijken. Omdat deze respondenten bij de eerste stellingvraag al aangaven de tweede correctie niet volledig uit te voeren rekenen we deze 10% tot de niet-volledig nakijkende tweede correctoren.
- 16% keek de open vragen na en de meerkeuzevragen niet (4%), gedeeltelijk (10%) of globaal of minder zorgvuldig (2%).
- 9% keek noch de open vragen noch de meerkeuzevragen volledig na.
- 57% keek niet alle kandidaten en/of niet alle vragen na. Deze groep valt uiteen in een groep die niet alle kandidaten nakeek (26%), een groep die niet alle vragen nakeek (19%), een groep waarvan we alleen weten dat zij alleen een gedeelte nakeken (9%) en een groep die aangaf niet alle vragen en niet alle kandidaten na te kijken (3%).

Al met al blijkt 93% van de tweede correctoren die volgens eigen opgave niet volledig nakijken dat gezien hun toelichtingen inderdaad niet te doen. Daarnaast is er een groep van 4% die toch volledig corrigeert alhoewel zij bij de eerste stelling te kennen gaven niet volledig na te kijken (en een groep van 3% waarvan de toelichtingen voor de inhoudsanalyse onbruikbaar waren).

Samenvatting

Van de tweede correctoren die bij de eerste stellingvraag aangaven **(helemaal) volledig** na te kijken, heeft gezien hun toelichting 71% daadwerkelijk volledig nagekeken, keek 12% de open vragen na maar de meerkeuzevragen niet of slechts gedeeltelijk, keek 5% niet alle kandidaten en/of niet alle vragen na, gaf 3% een toelichting die niet gecodeerd kon worden en gaf 10% een toelichting die niet met de vraagstelling in verband stond. Het gegeven dat 17% toch een niet-volledige correctiewijze blijkt toe te passen, geeft aanleiding tot twijfel aan de validiteit van de antwoorden op de eerste stellingvraag naar de volledigheid van de tweede correctie. Kennelijk vindt een substantieel deel van de docenten dat zij (helemaal) volledig nakijken terwijl zij dat in werkelijkheid (helemaal) niet doen.

Van de tweede correctoren die bij de eerste stellingvraag aangaven **(helemaal) niet volledig** na te kijken, heeft 92% gezien hun toelichting daadwerkelijk onvolledig nagekeken, 4% keek toch alle kandidaten en alle vragen na en 3% gaf een antwoord dat niet aan de vraagstelling gerelateerd was.

Als we de hiervoor besproken resultaten uitdrukken als een percentage van alle docenten die een toelichting gaven, heeft 56% van de tweede correctoren alle kandidaten en alle vragen gecorrigeerd, 13% keek de open vragen na maar sloeg (een deel van) de meerkeuzevragen over, 21% corrigeerde niet alle kandidaten en/of niet alle vragen, 2% gaf een toelichting die niet gecodeerd kon worden en 8% van de toelichtingen was off-topic.

3.3.7 Percentage gecorrigeerde kandidaten en vragen

Volgens de wettelijke voorschriften moeten tweede correctoren de antwoorden van alle kandidaten op alle vragen corrigeren. Behalve met de drie stellingen is nog een indicatie verkregen van de mate waarin de tweede correctie wordt uitgevoerd. De tweede correctoren is gevraagd van welk deel van de kandidaten zij een of meer vragen hebben gecorrigeerd en van welk deel van de vragen zij het werk van een of meer kandidaten hebben nagekeken

(zie tabel 3.3.11). In de rapportage hierna zijn alleen de gegevens gebruikt van de tweede correctoren die ten minste één kandidaat en ten minste één examenvraag nakeken.

Tabel 3.3.11 Percentage gecorrigeerde kandidaten en examenwerken per jaar (bron: tweede corrector)

	Jaar	N	Min	Max	Gem	Se-Gem	Stddev	P10	P25	P75	P90
Van welk deel van de kandidaten heeft u een of meer vragen nagekeken (als tweede corrector)?	2016	3213	2	100	97,88	0,20	11,37	100,00	100,00	100,00	100,00
	2011	3330	1	100	94,84	0,30	17,45	90,00	100,00	100,00	100,00
Welk deel van de vragen uit het examen heeft u nagekeken (voor een of meer kandidaten)?	2016	3206	10	100	95,98	0,23	13,04	90,00	100,00	100,00	100,00
	2011	3312	5	100	90,84	0,35	19,98	60,00	100,00	100,00	100,00

In 2016 heeft de gemiddelde tweede corrector uit het onderzoek naar eigen zeggen van 98% van de kandidaten een of meer vragen nagekeken en van 96% van de vragen een of meer kandidaten gecorrigeerd.

Opgemerkt wordt dat het hier niet gaat om percentages integrale correctie. Het nakijkpercentage van 98% van de examenvragen betekent bijvoorbeeld niet noodzakelijkerwijs dat de tweede corrector deze vragen voor alle examenkandidaten heeft nagekeken.

De verschillen tussen docenten zijn klein. Dat blijkt uit de percentiepunten P25, P75 en P90 (zie tabel 3.3.11) die alle 100 bedragen.

Verschillen tussen vragen, jaren en examens

Met behulp van multivariate variantieanalyse is nagegaan in hoeverre de verschillen tussen de antwoorden op de beide vragen (percentage nagekeken kandidaten versus percentage nagekeken vragen), de beide jaren en de twaalf examens significant zijn. De toetsing van de binnen-subjecten effecten brengt een significant hoofdeffect van vraag aan het licht ($p < 0,001$) evenals significante tweeweg interacties tussen vraag en jaar ($p < 0,001$) en tussen vraag en examen ($p = 0,002$), maar de drieweg interactie tussen vraag, jaar en examen is niet significant ($p = 0,025$). De toetsing van de tussen-subjecten effecten laat zien dat de hoofdeffecten van jaar en examen beide significant zijn evenals de interactie tussen beide ($p < 0,001$). Deze uitkomsten betekenen dat de nakijkpercentages variëren al naar gelang de vraag (percentage kandidaten versus percentage vragen), het jaar en het examen.

Verschillen tussen 2016 en 2011

Gemiddeld is er sprake van een toename van de percentages nagekeken kandidaten en vragen (ze tabel 3.3.11):

- In 2016 werden van 98% van de kandidaten een of meer vragen nagekeken tegen 95% in 2011.
- In 2016 werden van 96% van de vragen een of meer kandidaten gecorrigeerd tegen 91% in 2011.

De percentages tweede correctoren die 100% van de kandidaten of 100% van de vragen nagekeken, zijn uiteraard iets lager:

- In 2016 keek 95% van de tweede correctoren volgens eigen opgave alle kandidaten na tegenover 87% in 2011;
- In 2016 keek 87% van de tweede correctoren alle examenvragen na tegenover 77% in 2011.

Tegelijkertijd kunnen we constateren dat de verschillen tussen docenten kleiner geworden zijn. Dat wordt duidelijk als we de spreiding van de antwoorden in 2016 vergelijken met die in 2011 (zie tabel 3.3.11). De standaarddeviaties voor de percentages nagekeken kandidaten en examenvragen zijn in 2016 beduidend kleiner dan in 2011.

Verschillen tussen examens

Tabel 3.3.12 geeft inzicht in de verschillen tussen examens. Te zien is onder meer dat tweede correctoren geschiedenis in vergelijking met collega's van andere vakken weinig kandidaten en weinig examenvragen corrigeerden.

Tabel 3.3.12 Gemiddeld percentage gecorrigeerde kandidaten en examenwerken per examen per jaar (bron: tweede corrector)

	Jaar	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Percentage nagekeken kandidaten	2016	99	99	97	100	97	99	94	97	98	99	98	97
	2011	95	98	93	98	98	93	84	94	97	96	96	98
Percentage nagekeken examenvragen	2016	95	97	95	98	97	98	90	95	96	99	95	97
	2011	85	91	89	97	95	94	78	92	94	96	92	94

Verschillen tussen jaren en examens

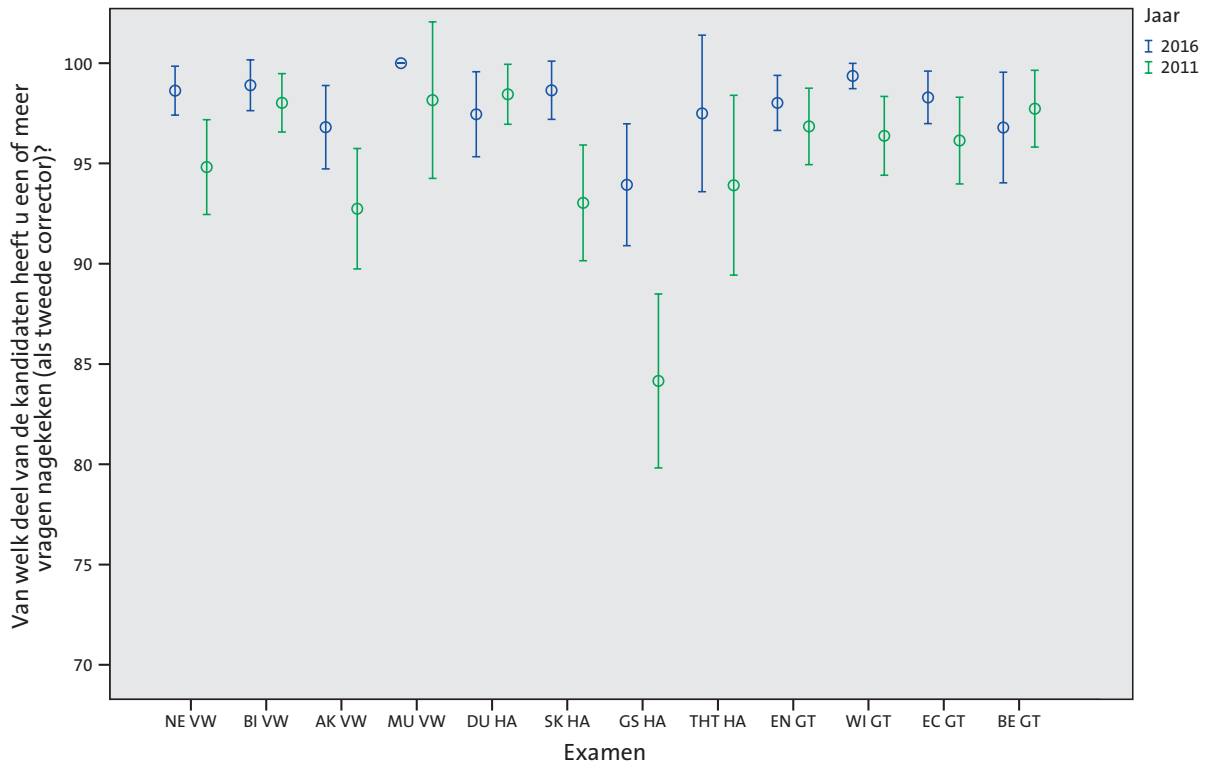
Hiervoor zagen we dat het verschil tussen de nakijkpercentages in 2016 en 2011 niet onafhankelijk is van de vraag (percentage kandidaten versus percentage vragen) en het examen. De geconstateerde tweeweg interacties tussen vraag en examen, tussen vraag en jaar en tussen jaar en examen zijn zichtbaar in tabel 3.3.12. In figuur 15 en 16 zijn de gemiddelde percentages nagekeken kandidaten en vragen per examen per jaar gevisualiseerd. Zo hebben tweede correctoren Nederlands, geschiedenis, wiskunde en scheikunde in 2016 van meer kandidaten een of meer vragen nagekeken dan in 2011. Het duidelijkst is dit te zien bij de examens geschiedenis en Nederlands.

Bij Nederlands, geschiedenis, aardrijkskunde, biologie, scheikunde en wiskunde is het percentage examenvragen dat voor een of meer kandidaten is nagekeken duidelijk toegenomen.

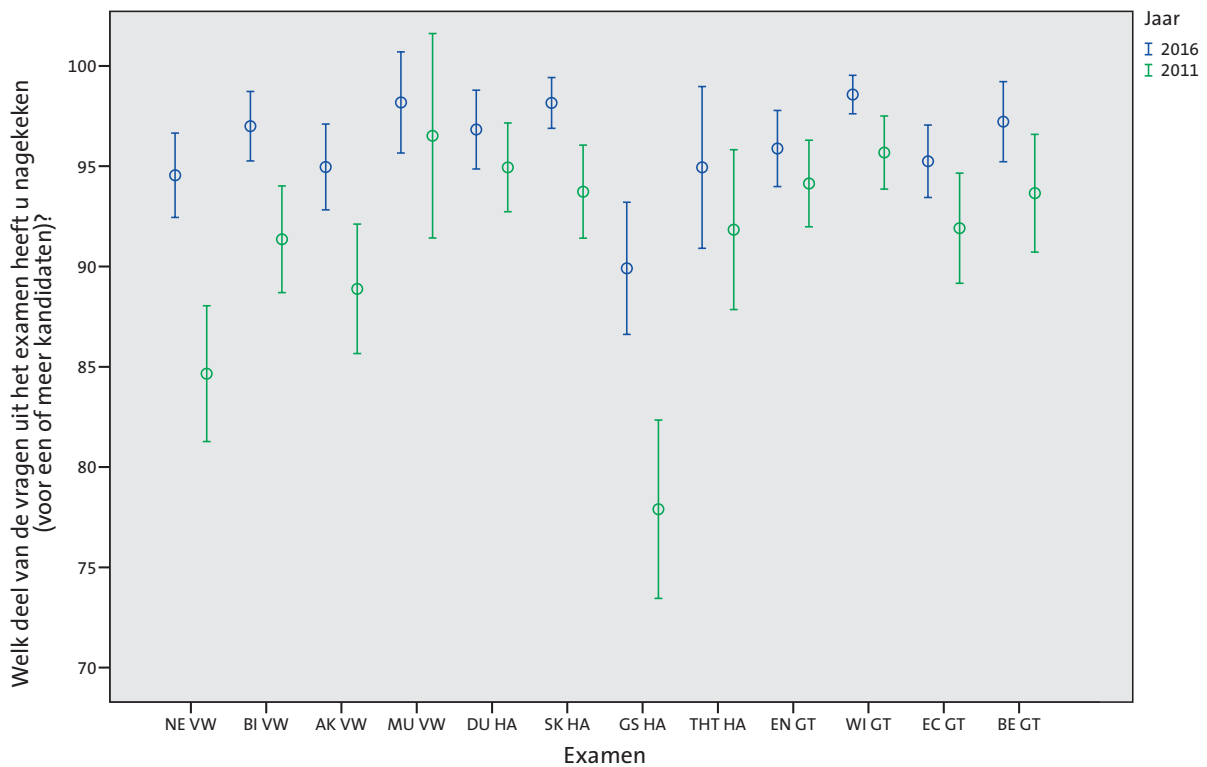
Deze toename is het grootst bij Nederlands en geschiedenis, twee arbeidsintensieve examens met antwoordmodellen die veel ruimte laten voor verschillende interpretaties (Kuhlemeier, Gitsels, Boom & Van de Kerkhof, 2011).

Slotopmerking

Wellicht ten overvloede merken we nogmaals op dat de hier gerapporteerde percentages geen integrale nakijkpercentages zijn. Een deel van de docenten zal niet alle vragen van de nagekeken examenwerken hebben gecorrigeerd en/of niet alle examenwerken van de nagekeken vragen hebben gecorrigeerd.



Figuur 15 Percentage gecorrigeerde kandidaten per examen per jaar (bron: tweede corrector)



Figuur 16 Percentage gecorrigeerde vragen per examen per jaar (bron: tweede corrector)

3.3.8 Percentage integraal nakijkende tweede correctoren

In deze paragraaf proberen we meer precieze informatie te verkrijgen over het percentage docenten dat daadwerkelijk alle examenwerken en alle examenvragen heeft nagekeken. Bijna een kwart (23%) heeft op de eerste stellingvraag niet geantwoord met 'klopt helemaal' maar met 'klopt'. Er zijn drie verklaringen te geven waarom de docenten 'klopt' hebben ingevuld. De eerste en meest voor de hand liggende verklaring is dat de docent de tweede correctie niet helemaal volledig heeft uitgevoerd. De tweede verklaring gaat er van uit dat de docent niet zeker genoeg is om voor het uiterste te kiezen. In de perceptie van de docent zijn er heel veel vragen en kandidaten gecorrigeerd en misschien is er wel een overgeslagen. Ook de derde verklaring gaat ervan uit dat niet alles is nagekeken maar dat de tweede correctie naar het oordeel van de docent wel **voldoende** volledig is uitgevoerd (bijvoorbeeld omdat de eerste corrector zorgvuldig en objectief heeft nagekeken). De gedachte die de docent daarbij gehad kan hebben is 'deze stelling klopt grotendeels'. Daarbij kan bedacht worden dat het antwoord 'klopt niet' voor zijn of haar gevoel verder van de waarheid af ligt. Om bovenstaande redenen geven we hier een conservatieve en een optimistische raming van het aantal docenten dat volledig corrigeert. In de conservatieve of strenge schatting zijn alleen degenen meegenomen die bij de eerste stellingvraag 'klopt helemaal' hebben ingevuld. De optimistische of soepele schatting neemt ook de docenten mee die op de eerste stellingvraag 'klopt' hebben geantwoord.

We rapporteren de uitkomsten van drie verschillende berekeningswijzen:

- een conservatieve raming op basis van alleen de gegevens van 2016;
- een optimistische raming op basis van alleen de gegevens van 2016;
- een optimistische raming op basis van alleen de vragen die zowel in 2016 als in 2011 zijn gesteld.

Conservatieve raming volgens het strenge vierdelige criterium

In het onderzoek van 2016 zijn vijf indicaties verkregen van de volledigheid waarmee de tweede correctie is uitgevoerd:

- 1 72% van de tweede correctoren heeft naar eigen zeggen al het examenwerk helemaal volledig nagekeken. Op de eerste stelling 'Ik heb het examenwerk van alle kandidaten helemaal nagekeken' reageerden zij met 'klopt helemaal' (en niet met 'klopt', 'klopt niet' of 'klopt helemaal niet').
- 2 7% van de tweede correctoren heeft volgens eigen opgave eerst een deel van de werken/vragen nagekeken en is daarna gestopt. Op de tweede stelling 'Ik heb eerst een deel van de werken/vragen nagekeken en ben daarna gestopt' reageerden zij met 'Klopt' of 'Klopt helemaal'.
- 3 95% van de tweede correctoren heeft naar eigen zeggen voor alle kandidaten voor een of meer vragen gecorrigeerd.
- 4 87% van de tweede correctoren heeft volgens eigen opgave alle examenvragen voor een of meer kandidaten nagekeken.
- 5 volgens de eerste correctoren heeft 44% van de tweede correctoren zeer volledig nagekeken (dat wil zeggen alle werken en per kandidaat alle vragen).

Meer precieze informatie over het percentage docenten dat daadwerkelijk alle examenwerken en alle examenvragen heeft nagekeken is verkregen door de antwoorden op vier vragen over de volledigheid van de tweede correctie met elkaar in verband te brengen. Allereerst zijn de reacties op de (eerste) stelling over integrale correctie in verband gebracht met de reacties op de (tweede) stelling over niet-integrale correctie. De uitkomsten van deze analyse geeft aanleiding de antwoorden op de eerste stelling over de volledigheid van de tweede correctie met enige voorzichtigheid te interpreteren. Van degenen die op de stelling over integrale correctie

reageerden met 'Klopt helemaal' bleek namelijk 4% de tweede vraag over niet-integrale correctie toch met 'Klopt' of 'Klopt helemaal' te hebben beantwoord. En van degenen die deze eerste vraag met 'Klopt' beantwoordden, was het overeenkomstige percentage 5%. Kennelijk vinden deze tweede correctoren dat zij (helemaal) integraal hebben nagekeken ook al hebben zij niet alle werken en/of alle vragen gecorrigeerd. Voor een overzicht van de (deels begrijpelijke) motieven die hieraan ten grondslag liggen, wordt verwezen naar paragraaf 3.3.6 en naar de publicatie van Loeffen en Eikelboom (in voorbereiding).

Een meer precieze schatting van het percentage integraal nakijkende tweede correctoren is verkregen door uit te rekenen hoeveel docenten zich volledig herkennen in de eerste stelling en zich tegelijkertijd (helemaal) niet herkennen in de tweede stelling die een niet-integrale correctiewijze beschrijft. Van alle tweede correctoren voldoet ruim de helft (69%) aan dit dubbele criterium. Van hen mag met wat meer zekerheid worden aangenomen dat zij de tweede correctie volledig overeenkomstig de wettelijke voorschriften uitvoeren.

Nog meer zekerheid kan worden verkregen door de antwoorden op de vragen naar de percentages gecorrigeerde kandidaten en examenvragen in het criterium te verdisconteren. Van alle tweede correctoren voldoet 66% aan dit uit vier onderdelen samenstelde criterium. Zij voeren de tweede correctie naar eigen zeggen helemaal volledig uit, passen geen niet-integrale correctiewijze toe en corrigeren alle kandidaten en alle vragen. Van de 72% die op de eerste stellingvraag 'klopt helemaal' antwoordde, zijn er dus nog 6% die op de andere vragen over de volledigheid van de correctie niet consequent antwoorden dat ze 'helemaal volledig' corrigeren. Van deze 66% mag met een grote mate van zekerheid worden aangenomen dat zij de tweede correctie volledig overeenkomstig de wettelijke voorschriften uitvoeren.

Verschillen tussen examens in 2016

De verschillen tussen examens zijn significant ($p < 0,001$). Tabel 3.3.13 toont het percentage integraal nakijkende tweede correctoren per examen (berekend op basis van het versoepelde vierdelige criterium). De verschillen tussen de examens zijn groot. Zo kijkt 76% van de tweede correctoren biologie volledig na tegen 44% van de tweede correctoren geschiedenis.

Tabel 3.3.13 Percentage integraal nakijkende tweede correctoren per examen (berekend op basis van het vierdelige criterium)

NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WIGT	EC GT	BEGT
64	76	52	80	75	73	44	69	64	75	61	62

Optimistische raming volgens het soepele vierdelige criterium

In de berekening is de tweede correctie van degenen die de eerste stelling met 'Klopt helemaal' beantwoordden als integraal beschouwd (ook al bleek 4% van hen naar eigen zeggen niet-integraal te beoordelen). Een optimistischere raming kan worden verkregen door ook degenen die de eerste stelling met 'Klopt' beantwoordden als integraal te beschouwen (ook al past 5% van hen naar eigen zeggen een niet-integrale werkwijze toe).

Van de 95% die op de eerste stellingvraag 'klopt helemaal' of 'klopt' antwoordden, blijkt 14% niet consequent te antwoorden dat ze volledig corrigeren op de andere drie vragen over de volledigheid. Volgens de optimistische raming met het soepele vierdelige criterium voert 81% van de docenten de tweede correctie volledig uit.

Het percentage docenten dat volledig nakijkt op basis van de eerste stellingvraag daalt na het meenemen van alle vragen over de volledigheid. Deze daling is 6% (van 72 naar 66) onder het strenge criterium en 14% (van 95 naar 81) onder het soepele criterium. Dit laat zien dat de

groep docenten die op de eerste stellingvraag ‘klopt’ antwoordde (23%) in grotere mate op de vervolgvragen alsnog aangeeft niet helemaal volledig na te kijken.

Wanneer we het soepele criterium hanteren, heeft volgens de eerste correctoren 85% van de tweede correctoren (zeer) volledig nagekeken.

Verschillen tussen examens in 2016

De verschillen tussen examens zijn opnieuw significant ($p < 0,001$). Tabel 3.3.14 toont het percentage volledig nakijkende tweede correctoren per examen (berekend op basis van het gerelaxeerde vierdelige criterium). De verschillen tussen de examens zijn wederom groot. Zo kijkt ruim driekwart van de tweede correctoren biologie volledig na tegen de helft van de tweede correctoren geschiedenis.

Tabel 3.3.14 Percentage integraal nakijkende tweede correctoren per examen (berekend op basis van het versoepelde vierdelige criterium)

NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
64	77	61	57	69	78	50	57	62	75	62	57

Alternatieve raming op basis van de gemeenschappelijke vragen van 2011 en 2016

Drie vragen over de volledigheid van de tweede correctie zijn zowel in 2016 als 2011 aan de tweede correctoren gesteld, zodat we de uitkomsten kunnen vergelijken. Het betreft de vraag naar de volledigheid van de tweede correctie (eerste stelling), het percentage gecorrigeerde kandidaten en het percentage nagekeken vragen. Berekend is hoeveel tweede correctoren de tweede correctie naar eigen zeggen helemaal volledig uitvoeren, alle kandidaten corrigeren (voor een of meer vragen) en alle vragen nakijken (voor een of meer kandidaten). Het verschil tussen 2016 en 2011 blijkt significant ($p < 0,001$). Van alle tweede correctoren voldoet in 2016 69% aan dit driedelige criterium tegen 42% in 2011. Wellicht ten overvloede merken we op dat het percentage integraal corrigerende docenten hier waarschijnlijk overschat wordt. In het criterium zijn de antwoorden op de stelling over niet-integrale tweede correctie immers niet meegenomen (omdat deze stelling alleen in 2016 aan respondenten is voorgelegd).

Verschillen tussen jaren en examens

De verschillen tussen examens zijn significant, zowel in 2016 als 2011 ($p < 0,001$). Tabel 3.3.15 toont het percentage integraal nakijkende tweede correctoren per jaar per examen (berekend op basis van het driedelige criterium). De tabel laat zien dat de verschillen tussen jaren en tussen examens groot zijn. In 2016 kijken tweede correctoren biologie en wiskunde bijvoorbeeld veel vollediger na dan hun collega's geschiedenis, aardrijkskunde en tehatex. Het verschil tussen 2016 en 2011 verschilt sterk van examens tot examens. Bij Nederlands en geschiedenis is het percentage integraal nakijkende tweede correctoren sinds 2011 meer dan verdubbeld terwijl de toename bij Engels en Duits naar verhouding bescheiden is.

Tabel 3.3.15 Percentage integraal nakijkende tweede correctoren per jaar per examen (berekend op basis van het driedelige criterium)

Jaar	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
2016	64	77	55	83	77	76	44	69	71	78	64	71
2011	29	42	29	57	61	41	20	41	52	47	40	53

Samenvattend

Afhankelijk van het gehanteerde criterium, lopen de schattingen van het percentage docenten dat de tweede correctie nog niet (helemaal) volledig uitvoert ver uiteen. Ze bedragen 5% (op basis van alleen de stellingvraag naar de volledigheid van de tweede correctie), 31% (op basis van de eerste en tweede stellingvraag tezamen), 34% (op basis van het strenge vierdelige criterium) en 19% (op basis van het soepele vierdelige criterium).

Deze percentages zijn berekend voor de groep docenten die de desbetreffende vragen beantwoord hebben. In het geval van het vierdelige criterium gaat het om 82%. De ontbrekende 18% is vooral toe te schrijven aan het grote aantal docenten dat de tweede stellingvraag 'Ik heb eerst een deel van de werken/vragen nagekeken en ben daarna gestopt' overgeslagen heeft (16%). Een eerste verklaring veronderstelt dat docenten die de eerste stellingvraag over de volledigheid van de tweede correctie positief beantwoorden de vervolgvraag over onvolledige tweede correctie als redundant beschouwen en om die reden onbeantwoord laten. Een tweede verklaring is dat het een gevoelige kwestie betreft die sommige docenten liever onbeantwoord laten. Een alternatieve schatting is verkregen door de totale groep van 3540 respondenten als percentagebasis te nemen. De percentages docenten die de tweede correctie nog niet (helemaal) volledig uitvoeren, liggen dan doorgaans aanzienlijk hoger. Afhankelijk van het gehanteerde criterium gaat het om 5% (op basis van de eerste stellingvraag), 42% (op basis van de beide stellingvragen tezamen), 46% (op basis van het strenge vierdelige criterium) en 34% (op basis van het soepele vierdelige criterium). Deze getallen tonen dus de percentages docenten die op deze vragenlijst hebben aangegeven volledig te corrigeren. Omdat we niet zeker weten waarom de docenten de vragen hebben overgeslagen en hoe ze de tweede correctie precies hebben uitgevoerd, is deze berekeningswijze niet als leidend genomen.

3.3.9 Mening over de noodzaak van integrale tweede correctie

In de vragenlijst van 2016 zijn de respondenten twee stellingen voorgelegd over de noodzaak van integrale tweede correctie:

- De tweede correctie moet integraal worden uitgevoerd (dat wil zeggen: alle examenwerken en per kandidaat alle vragen).
- Als de tweede corrector een deel heeft nagekeken en tot de conclusie komt dat de eerste corrector goed heeft nagekeken, is het niet nodig om alles na te kijken.

Eerste stelling

De overgrote meerderheid van de docenten - 82% - is van mening dat de tweede correctie integraal moet worden uitgevoerd (zie tabel 3.3.16). Bijna de helft (47%) is het er helemaal mee eens, meer dan een derde (35%) is het ermee eens, 4% heeft geen mening of weet het niet, 11% is het er niet mee eens en 3% is het er helemaal niet mee eens.

Tabel 3.3.16 De noodzaak van integrale tweede correctie (kolompercentage optellend tot 100%)

	%
Helemaal mee oneens	3
Mee oneens	11
Geen mening/Weet niet	4
Mee eens	35
Helemaal mee eens	47

De mening over de noodzaak van tweede correctie verschilt van vak tot vak ($p < 0,001$). Zo zijn docenten geschiedenis, tehatex en aardrijkskunde daar minder van overtuigd dan docenten muziek, biologie en wiskunde.

Tabel 3.3.17 De noodzaak van integrale tweede correctie per examen (kolompercentages optellend tot 100%)

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Helemaal mee oneens	4	2	2	3	4	3	4	4	3	2	3	3
Mee oneens	17	8	17	5	6	10	26	11	8	6	11	7
Geen mening/Weet niet	2	1	9	0	2	3	7	10	5	3	5	8
Mee eens	34	36	39	23	29	36	40	39	34	30	36	40
Helemaal mee eens	43	53	33	68	58	48	23	37	52	59	46	42

Tweede stelling

De tweede stelling beschrijft een vorm van niet-integrale correctie waarbij niet alles wordt nagekeken als blijkt dat de eerste corrector zijn of haar werk goed gedaan heeft (zie tabel 3.3.18). Bijna driekwart van de docenten (71%) staat (sterk) afwijzend tegenover deze vorm van niet-integrale correctie. Van de docenten is 30% het er helemaal mee oneens, 41% mee oneens, 8% heeft geen mening of weet het antwoord niet, 16% kiest voor mee eens en 4% voor helemaal mee eens.

Tabel 3.3.18 De wenselijkheid van niet-integrale tweede correctie (kolompercentages optellend tot 100%)

	%
Helemaal mee oneens	30
Mee oneens	41
Geen mening/Weet niet	8
Mee eens	16
Helemaal mee eens	4

De verschillen tussen vakken zijn significant ($p < 0,001$). Zo vinden docenten geschiedenis, aardrijkskunde en tehatex vaker dat je niet alles hoeft te corrigeren als de eerste corrector goed werk heeft afgeleverd dan docenten muziek, biologie en wiskunde (zie tabel 3.3.19).

Tabel 3.3.19 De wenselijkheid van niet-integrale tweede correctie per examen (kolompercentages optellend tot 100%)

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Helemaal mee oneens	26	35	19	50	44	25	11	20	38	37	29	29
Mee oneens	40	47	36	34	42	45	32	37	43	43	43	42
Geen mening/Weet niet	7	8	8	5	6	8	11	9	9	8	7	13
Mee eens	21	6	31	9	6	18	35	30	9	10	16	13
Helemaal mee eens	6	4	6	2	3	4	10	4	2	3	4	4

3.3.10 Oorzaken van niet-integrale tweede correctie

De docenten is in 2016 gevraagd naar hun mening over de vermoedelijke oorzaken van onvolledige tweede correctie. De exacte vraagstelling luidde:

“Uit onderzoek blijkt dat de tweede correctie niet altijd integraal wordt uitgevoerd (d.w.z. niet alle kandidaten en/of niet alle examenvragen). Waarom denkt u dat een deel van de docenten de tweede correctie niet volledig uitvoert?”

De docenten konden daarbij kiezen uit een of meer van de oorzaken die zijn opgesomd in tabel 3.3.20. Tijdgebrek blijkt voor 82% de belangrijkste oorzaak, gevolgd door chronische overbelasting (48%) en onvoldoende financiële compensatie (38%). Ruim een kwart (28%) verwijst naar antwoordmodellen waarvan de interpretatie veel tijd kost. Een vijfde tot een kwart denkt dat een deel van de docenten de tweede correctie onvolledig uitvoert omdat een steekproef van kandidaten voldoende garantie biedt (25%) of omdat een steekproef van examenvragen voldoende garantie biedt (20%). Dat de noodzaak van integrale tweede correctie onvoldoende is aangetoond wordt door 7% van de respondenten gekozen.

Tabel 3.3.20 Oorzaken van onvolledige tweede correctie

	N	Percentage
Tijdgebrek (bijvoorbeeld omdat de eerste correctie al zoveel tijd vergt)	3400	82
Onvoldoende financiële compensatie (kleine attentie of onkosten declareren volstaat niet)	3400	38
Chronische overbelasting (aan het einde van een lang schooljaar snakt menigeen naar de grote vakantie)	3400	48
De noodzaak van integrale tweede correctie is onvoldoende aangetoond	3400	7
Een steekproef van kandidaten biedt voldoende garantie	3400	25
Een steekproef van examenvragen biedt voldoende garantie	3400	21
De interpretatie van de antwoordmodellen kost veel tijd	3400	28

Behoudens een uitzondering - de “oorzaak” dat de noodzaak van integrale tweede correctie onvoldoende is aangetoond - zijn de verschillen tussen de examens significant ($p < 0,001$). Tabel 3.3.21 laat bijvoorbeeld zien dat tijdgebrek voor docenten scheikunde, geschiedenis en Nederlands een belangrijkere oorzaak van niet-integrale tweede correctie is dan voor docenten wiskunde. Dat een steekproef van kandidaten of vragen voldoende garantie biedt, noemen docenten geschiedenis en aardrijkskunde veel vaker dan docenten muziek en beeldende vakken. De tijdrovende interpretatie van de antwoordmodellen wordt door docenten tehatex, beeldende vakken, geschiedenis, Nederlands en biologie vaker gemeld als een oorzaak voor de onvolledigheid van de correctie dan door docenten Engels, Duits en wiskunde.

Tabel 3.3.21 Oorzaken van onvolledige tweede correctie per examen (percentages per oorzaak per examen)

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WIG T	EC GT	BEG T
Tijdgebrek	87	85	86	72	81	87	92	79	80	73	79	73
Financiële compensatie	44	37	44	33	41	43	39	33	32	36	41	27
Overbelasting	59	53	53	42	47	49	57	52	44	37	43	42
Integraal niet nodig	10	6	8	4	5	6	8	8	6	8	9	6
Steekproef kandidaten voldoende garantie	22	21	37	4	19	24	37	18	22	27	26	15
Steekproef vragen voldoende garantie	22	20	30	5	12	18	30	23	16	21	25	18
Antwoordmodellen tijdrovend	38	42	32	25	20	20	38	46	18	17	26	45

3.3.11 Toelichtingen van docenten bij de vraag naar de oorzaken van onvolledige tweede correctie

In totaal 668 respondenten hebben de moeite genomen hun antwoord op de matrixvraag naar de mogelijke oorzaken van onvolledige tweede correctie toe te lichten. De inhoud van de toelichtingen is geanalyseerd met behulp van een categorieënsysteem. Voor het uitgebreide verslag van de inhoudsanalyse wordt verwezen naar Loeffen en Eikelboom (in voorbereiding). De resultaten van de analyse zijn weergegeven in tabel 3.3.22.

Tabel 3.3.22 Verdeling van de toelichtingen van docenten bij de vraag naar de oorzaken van onvolledige tweede correctie (kolompercentages optellend tot 100%)

Toelichting	%
Twijfel aan noodzaak of effectiviteit van volledige tweede correctie	
Twijfel aan effectiviteit of noodzaak van (volledige) tweede corrector/correctie	21
Redenen die samenhangen met het examen	
Vraagstelling/(open) vragen onduidelijk, slecht of meervoudig te interpreteren	3
Correctievoorschrift of antwoordmodel is onvolledig, onduidelijk, meervoudig te interpreteren en/of tijdrovend	2
De tweede correctie kost teveel tijd / is teveel werk (niet nader gespecificeerd)	6
Aanvullingen op correctiemodel tijdrovend, lang en/of tegenstrijdig met correctiemodel	0
Facilitering door de school	
Geen of gebrekkige financiële compensatie	1
Geen of beperkte vrijstelling/uitroosting/taakverlichting	5
Gezamenlijk overleg	
Tegenzin of angst voor overleg/confrontatie met eerste corrector	4
(Persoons)kenmerken van de docenten	
Gebrek aan motivatie van de tweede corrector	23
De betrokkenheid met onbekende leerlingen is minder groot	3
Gebrekkige professionaliteit/competentie/attitude van de tweede corrector (als beoordelaar)	2

(vervolg tabel 3.3.22)

Toelichting	%
Oneens met de vraag	
De tweede correctie wordt wel degelijk altijd volledig uitgevoerd.	7
Overig	
Weet niet/geen mening	17
Overig on-topic (wel antwoord op de vraag, maar niet codeerbaar)	0
Off-topic (geen antwoord op de vraag)	6
Totaal	100

In hun toelichtingen hebben docenten hun mening gegeven over de vermoedelijke oorzaken van onvolledige tweede correctie. Als belangrijkste oorzaak noemen zij een gebrek aan motivatie bij de tweede corrector (23%). Enkele karakteristieke toelichtingen luiden: 'Geen zin', 'Laksheid', 'Luiheid', 'Opzien tegen al het werk', 'Onvoldoende arbeidsethos' en 'Niet plichtsgetrouw'.

In hun toelichtingen verwijst een op de vijf docenten (20%) naar twijfel bij tweede correctoren aan de effectiviteit of noodzaak van de tweede correctie (als oorzaak van onvolledige tweede correctie). Enkele typische uitspraken zijn: 'De eerste corrector komt ermee weg', 'Het maakt weinig uit op het totale cijfer', 'De tweede correctie is niet nodig', 'De eerste correctoren kijken al goed en zorgvuldig na', 'Een steekproef is voldoende', 'De meerkeuzevragen hoeven niet, vergelijking met Wolf is voldoende' en 'Alleen lastige vragen is voldoende'.

Iets meer dan een op de tien docenten (11%) verwijst naar het examen als oorzaak van onvolledige tweede correctie. Bijvoorbeeld: 'Meer meerkeuzevragen', 'Meer vrijheid als professional om antwoorden van leerlingen te mogen interpreteren', 'De interpretatie van de antwoorden van de leerlingen kost veel tijd' en 'Regiobesprekingen schijnen soms een ander licht op de correctie'.

Een gebrekkige facilitering door de school wordt door 5% van de docenten genoemd. Bijvoorbeeld: 'Geen financiële compensatie' en 'Ander schoolwerk gaat gewoon door'.

Vier procent denkt dat de tweede correctie niet altijd volledig wordt uitgevoerd omdat tweede correctoren opzien tegen de confrontatie met de eerste corrector. Bijvoorbeeld: 'Heeft geen zin: collega met 'grote mond' wint altijd', 'Angst voor confrontatie met eerste correctie', 'Angst voor oeverloze discussie met andere woorden laag rendement' en 'Dan hoef je ook niet de discussie aan te gaan met de eerste corrector als je verschillen constateert (Is namelijk regelmatig een onprettig gesprek)'

Tot slot merken we nog op dat 19% van de docenten te kennen gaf de vraag niet te kunnen beantwoorden of geen mening te hebben. Twee karakteristieke toelichtingen zijn: 'Geen idee, dat kan ik niet voor een ander bepalen' en 'Vraag het de docenten die niet-integraal nakijken of die wel belasting ervaren'.

Volledigheid tweede correctie en tijdsdruk

Hiervoor noemden docenten tijdgebrek de belangrijkste oorzaak waardoor een deel van de docenten de tweede correctie de correctie niet altijd volledig uitvoert. Aanvullend is de tweede correctoren gevraagd in hoeverre zij het eens zijn met de stelling 'Door tijdsdruk heb ik de tweede correctie minder zorgvuldig kunnen uitvoeren dan wenselijk is'. Een derde (33%) is het er helemaal niet mee eens en 39% is het er niet mee eens, 15% kiest voor 'noch mee eens, noch mee oneens', 9% voor 'mee eens' en 5% voor 'helemaal mee eens'. Slechts één docent gaf aan dit jaar geen tweede corrector te zijn geweest.

Tabel 3.3.23 Tijdsdruk als oorzaak van een minder zorgvuldige tweede correctie (bron: tweede correctoren; kolompercentages optellend tot 100%)

	%
Helemaal niet mee eens	33
Niet mee eens	39
Noch mee eens, noch mee oneens	15
Mee eens	9
Helemaal mee eens	5
Ik was dit jaar geen tweede corrector	0

De verschillen tussen examens zijn significant ($p < 0,001$) en weergegeven in tabel 3.3.24. Zo is tijdgebrek voor docenten geschiedenis een belangrijker oorzaak van onzorgvuldige tweede correctie dan voor docenten Duits, Engels en wiskunde.

Tabel 3.3.24 Tijdsdruk als oorzaak van een minder zorgvuldige tweede correctie per examen (bron: tweede correctoren; kolompercentages optellend tot 100%)

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BEGT
Helemaal niet mee eens	26	32	24	55	55	30	12	21	40	43	32	28
Niet mee eens	39	45	38	29	31	36	33	37	40	39	39	50
Noch mee eens, noch mee oneens	19	14	17	16	9	19	20	21	12	11	15	13
Mee eens	10	7	13	0	4	11	19	16	5	6	10	6
Helemaal mee eens	6	2	9	0	1	5	16	3	2	1	4	2

3.3.12 Maatregelen ter bevordering van integrale tweede correctie

De docenten is gevraagd naar de noodzaak van potentiële maatregelen ter stimulering van integrale tweede correctie. De precieze vraagstelling luidde: “Welke maatregelen zijn naar uw mening nodig om te bevorderen dat alle docenten de tweede correctie integraal (kunnen) uitvoeren?” De respondenten konden daarbij kiezen uit een of meer van de maatregelen die zijn opgesomd in tabel 3.3.25. Het opnemen van de uren in de jaartaak (61%) is de meest gekozen maatregel, gevolgd door taakverlichting tijdens de examenperiode (59%) en een betere financiële compensatie (46%). Ruim een derde (36%) vindt dat er sterk sturende antwoordmodellen ontwikkeld moeten worden met bijvoorbeeld meer voorbeelden van volledig juiste, gedeeltelijk juiste en foute antwoorden. Relatief weinig docenten onderschrijven de noodzaak van een betere controle op de uitvoering van de tweede correctie (19%) of een betere voorlichting over het belang van tweede correctie (10%).

Tabel 3.3.25 De wenselijkheid van maatregelen ter bevordering van integrale tweede correctie (percentage docenten)

	N	%
Geen maatregelen nodig	3472	7
Taakverlichting tijdens de examenperiode (bijvoorbeeld door vrijstelling van surveilleren of uitroosteren van lesuren)	3472	59
De uren voor de tweede correctie aanwijsbaar opnemen in de jaartaak	3472	61
Financiële compensatie voor de gewerkte uren (anders dan een kleine attentie of het declareren van telefoonkosten)	3472	46
Betere voorlichting over het belang van integrale tweede correctie	3472	10
Een betere controle op de uitvoering van de tweede correctie	3472	19
Ontwikkeling van sterker sturende antwoordmodellen met bijv. meer voorbeelden van volledig juiste, gedeeltelijk juiste en foute antwoorden (meer dicht getimmerd)	3472	36

Behoudens één uitzondering zijn de verschillen tussen de examens bij alle genoemde maatregelen significant ($p < 0,01$). De uitzondering doet zich voor bij het alternatief 'geen maatregelen nodig'. Zo hebben docenten Nederlands en geschiedenis meer behoefte aan taakverlichting tijdens de examenperiode dan docenten economie en Engels (zie tabel 3.3.26). Ook vinden docenten Nederlands, geschiedenis en scheikunde het opnemen van uren in de jaartaak wenselijker dan docenten Engels, economie en wiskunde. Financiële compensatie wordt vaker genoemd door docenten Nederlands en aardrijkskunde dan door docenten Engels, tehatex, muziek en beeldend. Sterker sturende antwoordmodellen blijken meer een wens van docenten Nederlands, economie en geschiedenis dan van docenten muziek, wiskunde en scheikunde.

Tabel 3.3.26 Maatregelen ter bevordering van integrale tweede correctie per examen (percentage docenten)

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Geen maatregelen nodig	4	7	8	9	7	6	6	5	7	11	8	9
Taakverlichting	79	62	67	53	55	65	78	66	50	43	50	53
Jaartaak	71	62	60	53	61	68	65	59	56	54	57	61
Financiële compensatie	55	43	51	29	46	48	45	41	45	44	46	38
Betere voorlichting	8	9	7	9	12	10	5	10	14	15	9	11
Betere controle	15	20	16	29	23	14	12	15	22	23	24	18
Sterker sturende antwoordmodellen	44	43	41	17	32	27	43	31	35	27	44	39

3.3.13 Toelichtingen van docenten bij de vraag naar maatregelen ter bevordering van volledige tweede correctie

In totaal 534 docenten hebben de moeite genomen hun antwoord op de matrixvraag naar maatregelen ter bevordering van volledige tweede correctie toe te lichten. De toelichtingen zijn gecodeerd met behulp van een categorieënsysteem. Voor het uitgebreide verslag van de inhoudsanalyse wordt verwezen naar Loeffen en Eikelboom (in voorbereiding). De resultaten van de analyse zijn weergegeven in tabel 3.3.27:

Tabel 3.3.27 Verdeling van de toelichtingen van docenten bij de vraag naar maatregelen ter bevordering van volledige tweede correctie (kolompercentages optellend tot 100%)

Toelichting	%
Oplossingen gerelateerd aan het examen	
Vraagstelling duidelijker, eenduidiger, minder subjectief en/of minder (open) vragen	13
Correctiemodel/antwoordmodel duidelijker, vollediger, minder tijdrovend en/of minder dicht getimmerd	10
Examen beter, korter, minder tijdrovend, objectiever, minder nuancegevoelig en/of minder teksten	5
Aanvullingen op correctiemodel (zowel meer als minder), meer status geven, eerder aanleveren, aangeven welke toegestaan zijn	4
Handmatig invoeren minder belastend dan via Wolf	0
Digitalisering van examens	2
Oplossingen gerelateerd aan de correctieprocedure	
Eerlijkere toewijzing van examens (door DUO of school)	1
Geen contact tussen eerste en tweede corrector, middelen en daarna bij onenigheid derde corrector	2
Geheel onafhankelijke eerste en tweede correctie of tweede correctie door onafhankelijke instantie	5
Omkering correctievolgorde	8
Annoteren niet meer toestaan	1
Afspraken over aanlevering examenwerk aan tweede corrector vastleggen	2
Betere feedback op de uitvoering van de tweede correctie	0
Moeilijk na te kijken examens eerder inroosteren / niet aan het einde	1
Correctietermijn verlengen	5
Fysiek contact tussen eerste en tweede corrector	1
Oplossingen gerelateerd aan de facilitering door de school	
Financiële compensatie	2
Vrijstelling/vrij-roostering/taakverlichting	5
Meer of betere controle op de tweede correctie (met consequenties bij slecht werk)	4
Opnemen uren voor tweede correctie in jaartaak	1
Scholing van docenten in het nakijken van de examens	1
Oplossing (voorlopig) niet nodig	
Geen maatregelen nodig	12
Eerst onderzoeken of integrale tweede correctie noodzakelijk is	1
Overige oplossingen	
Dwang	1
Cultuurverandering (beter arbeidsethos)	3
Oplossing hangt af van het vak	3
Overig	
Weet ik niet / geen mening	2
Overig on-topic (wel antwoord op de vraag, maar niet codeerbaar)	3
Off-topic (geen antwoord op de vraag)	2
Totaal	100

De toelichtingen van de docenten bevatten een schat aan suggesties ter bevordering van de volledigheid van de tweede correctie die nadere bestudering meer dan waard zijn:

- Ruim een derde van de docenten (34%) doet suggesties ter verbetering van het examen als oplossing voor het gegeven dat niet alle docenten de tweede correctoren volledig uitvoeren. Enkele karakteristieke uitspraken zijn: 'Examenvragen maken die correct zijn en niet subjectief', 'Minder open vragen', 'Er moet meer recht gedaan worden aan andere goede antwoorden van leerlingen', 'Het geschiedenisexamen kent onnodig veel vragen', 'Landelijke examenbesprekingen meer op elkaar afstemmen. Nu spreken ze elkaar soms tegen' en 'Samenvoegen van het officiële (Cito!) CvTE-antwoordmodel en het verslag in Levende Talen'.
- Ruim een kwart (26%) ziet de oplossing voor de onvolledige tweede correctie in een verbetering van de huidige examenprocedure. Het meest genoemd zijn: omkering van de correctievolgorde (8%), geheel onafhankelijke eerste en tweede correctie of tweede correctie door onafhankelijke instantie (5%) en verlenging van de correctietermijn (5%). Enkele typerende toelichtingen bij deze vraag zijn: 'Het staatsexamen model. Twee onafhankelijke correctoren kijken de boel na. Bij meer dan 5% verschil kijkt een derde onafhankelijke corrector de boel na', 'Omdraaien correctievolgorde', 'Omkering 1e en 2e correctie: mits fraude wordt voorkomen (degene die de uiteindelijke cijfers invoert, moet niet de eigen docent zijn)' en 'Zet alle examenkandidaten van een regio bij elkaar en geef het examenwerk mee aan de verschillende docenten'.
- Ruim één op de tien (13%) noemt een betere facilitering door de school als oplossing voor de onvolledige tweede correctie. Bijvoorbeeld: 'En dan een echte financiële compensatie. Eigenlijk belachelijk dat we voor 'niets' zoveel werk moeten verrichten', 'Taakverlichting bij de vakken waarbij nakijken veel tijd kost (bijv. Nederlands)' en 'Ik heb mijn directie op de hoogte gebracht van 1: De verplichting 70 eindexamen leerlingen te moeten nakijken. 2: 20 uren te surveilleren. 3: 6 uren te moeten geven'.
- Een op de tien docenten vindt dat een oplossing (voorlopig) niet nodig is of dat er eerst onderzoek gedaan moet worden om te bepalen of volledige correctie echt noodzakelijk is. Enkele toelichtingen: 'Integrale tweede correctie lijkt me niet nodig als je als tweede corrector ziet dat de eerste correctie zorgvuldig en consequent volgens het correctievoorschrift is nagekeken'.
- Zeven procent van de docenten verwijst naar andere oplossingen dan die hiervoor besproken zijn of merkt op dat de oplossing afhankelijk is van het vak. Bijvoorbeeld: 'Dwang (sorry, maar mensen nemen de vrijheid die ze krijgen)', 'Beroep doen op arbeidsethos' en 'Dit geldt echter voor bepaalde vakken als Nederlands en geschiedenis. Met heel veel uitgebreide antwoorden. Niet voor moderne vreemde talen'.
- Tot slot heeft 7% een niet-codeerbaar of anderszins onbruikbaar antwoord gegeven. Bijvoorbeeld: 'Eerlijk gezegd zou ik niet weten wat zou kunnen helpen, bovenstaande suggesties zijn eigenlijk niet ideaal omdat ook met beloning niet gegarandeerd is dat collega's de tweede correctie serieus nemen'.

3.4 Kwaliteit van de eerste en tweede correctie

Hiervoor stond de volledigheid van de tweede correctie centraal. Examinatoren moeten het werk niet alleen volledig nakijken, maar de kandidaten ook beoordelen volgens de door de overheid verstrekte normen. In deze paragraaf presenteren we de uitkomsten met betrekking tot de zorgvuldigheid en objectiviteit van de beoordeling. Daarnaast rapporteren we over hoe docenten denken over de wenselijkheid van potentiële maatregelen ter verbetering van de correctie.

3.4.1 Toepassing van het correctievoorschrift

De tweede correctoren is gevraagd in hoeverre de eerste corrector het correctievoorschrift correct heeft toegepast (zie tabel 3.4.1). In 2016 past - aldus de tweede correctoren - 8% van de eerste correctoren het correctievoorschrift volledig correct toe, ruim de helft (54%) doet dat correct, 27% min of meer correct, 9% niet correct en 2% helemaal niet correct.

Tabel 3.4.1 De toepassing van het correctievoorschrift door de eerste corrector per jaar (bron: tweede corrector; kolompercentages optellend tot 100%)

	2016	2011
Volledig correct	8	5
Correct	54	53
Min of meer correct	27	31
Niet correct	9	9
Helemaal niet correct	2	3

Verschillen tussen 2016 en 2011

In 2016 vinden de tweede correctoren dat de eerste correctoren het correctievoorschrift in 2016 wat beter hebben toegepast dan in 2011 ($p < 0,001$). Inspectie van de percentages in tabel 3.4.1 laat zien dat het hier om een zeer klein verschil gaat.

Verschillen tussen examens

De naar examen en jaar uitgesplitste gegevens zijn weergegeven in tabel 3.4.2. In 2016 is het correctievoorschrift Duits bij de eerste correctie beter toegepast dan het correctievoorschrift geschiedenis, zo zeggen de tweede correctoren van deze twee examens.

Tabel 3.4.2 De toepassing van het correctievoorschrift door de eerste corrector per examen per jaar (bron: tweede corrector; kolompercentages per examen per jaar optellend tot 100%)

Jaar		NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
2016	Volledig correct	9	7	4	6	19	8	5	8	12	7	6	3
	Correct	49	52	55	63	55	53	52	46	49	59	55	59
	Min of meer correct	28	28	30	27	22	25	29	29	30	24	29	24
	Niet correct	12	10	9	4	3	12	12	11	8	8	7	10
	Helemaal niet correct	2	2	2	0	0	3	3	7	2	2	3	4
2011	Volledig correct	6	5	3	18	7	6	3	3	4	4	5	5
	Correct	49	50	52	44	59	61	55	46	46	56	55	50
	Min of meer correct	30	35	33	35	27	25	31	31	33	31	30	31
	Niet correct	10	8	10	3	6	5	8	17	13	6	8	10
	Helemaal niet correct	5	2	2	0	1	2	4	3	3	2	2	5

3.4.2 Zorgvuldigheid van de correctie

In de vragenlijst is eerste correctoren gevraagd naar de zorgvuldigheid waarmee de tweede corrector heeft nagekeken. En de tweede correctoren is gevraagd naar de zorgvuldigheid van de eerste correctoren. De vraag voor de eerste corrector luidde: "Hoe zorgvuldig of slordig heeft de

tweede corrector het werk van uw kandidaten naar uw mening nagekeken (bijv. verschillende beoordelingen van gelijke antwoorden of uitwerkingen; er zijn goede en/of foute antwoorden over het hoofd gezien)?” De vraag aan de tweede correctoren was precies zo geformuleerd met dien verstande dat het onderwerp de eerste correctie was.

Zorgvuldigheid van de tweede correctie (volgens de eerste correctoren)

In 2016 heeft - aldus de eerste correctoren - ruim een derde (36%) van de tweede correctoren het werk zeer zorgvuldig nagekeken, de helft deed dat zorgvuldig (51%), 7% min of meer zorgvuldig, 1% slordig en 0% zeer slordig (zie tabel 3.4.3). Vier procent gaf te kennen de vraag naar de zorgvuldigheid van de tweede correctie niet te kunnen beantwoorden.

Tabel 3.4.3 De zorgvuldigheid van de tweede correctie per jaar (bron: eerste corrector; kolompercentages optellend tot 100%)

	2016	2011
Zeer zorgvuldig	36	24
Zorgvuldig	51	54
Min of meer zorgvuldig	7	13
Slordig	1	1
Zeer slordig	0	1
Ik kan niet overzien hoe zorgvuldig de tweede corrector heeft nagekeken	4	8

Verschillen tussen 2016 en 2011

In de analyse van de verschillen tussen jaren en examens zijn de docenten die deze vraag niet konden beantwoorden buiten beschouwing gelaten. De toetsing van de verschillen tussen jaren gaf significantie te zien ($p < 0,001$). In 2016 vinden naar verhouding meer eerste correctoren dat de tweede correctie zorgvuldig is uitgevoerd dan in 2011.

Verschillen tussen examens in 2016

De verschillen tussen examens zijn significant ($p < 0,001$). Zo blijken tweede correctoren Duits en biologie in 2016 volgens de eerste correctoren iets zorgvuldiger na te kijken dan hun collega's geschiedenis (zie tabel 3.4.4).

Tabel 3.4.4 De zorgvuldigheid van de tweede correctie per examen per jaar (bron: eerste corrector; kolompercentages per examen per jaar optellend tot 100%)

Jaar		NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BEGT
2016	Zeer zorgvuldig	43	48	38	39	47	43	26	36	36	35	32	31
	Zorgvuldig	49	48	50	56	48	49	62	48	56	56	58	61
	Min of meer zorgvuldig	6	4	11	6	4	8	10	12	7	8	9	7
	Slordig	2	1	1	0	1	0	1	3	1	1	1	1
	Zeer slordig	0	0	0	0	0	0	1	1	0	0	1	1
2011	Zeer zorgvuldig	25	31	20	44	36	28	16	18	23	24	28	28
	Zorgvuldig	56	60	59	46	54	58	63	60	60	62	56	57
	Min of meer zorgvuldig	17	7	20	10	9	13	21	20	15	11	13	14
	Slordig	2	1	0	0	1	0	1	2	2	2	2	1
	Zeer slordig	1	1	1	0	1	0	0	1	1	0	2	0

Zorgvuldigheid van de eerste correctie (volgens de tweede correctoren)

De vraag naar de zorgvuldigheid van de eerste correctie is gesteld aan de tweede correctoren. In 2016 heeft 14% van de eerste correctoren het werk zeer zorgvuldig nagekeken, ruim de helft deed dat zorgvuldig (53%), 21% min of meer zorgvuldig, 9% slordig en 3% zeer slordig.

Tabel 3.4.5 De zorgvuldigheid van de eerste correctie per examen (bron: tweede corrector; kolompercentages optellend tot 100%)

	2016	2011
Zeer zorgvuldig	14	10
Zorgvuldig	53	57
Min of meer zorgvuldig	21	21
Slordig	9	8
Zeer slordig	3	3

Verschillen tussen 2016 en 2011

In de analyse van de verschillen tussen jaren en examens zijn de docenten die deze vraag niet konden beantwoorden buiten beschouwing gelaten. De verschillen tussen beide jaren blijken niet significant ($p = 0,104$). In 2016 hebben eerste correctoren volgens de tweede correctoren even zorgvuldig nagekeken als in 2011.

Verschillen tussen examens in 2016

De toetsing van de verschillen tussen examens leverde significantie op ($p < 0,001$). Zo keken eerste correctoren Duits in 2016 volgens de tweede correctoren zorgvuldiger na dan hun collega's Nederlands en Engels.

Tabel 3.4.6 De zorgvuldigheid van de eerste correctie per examen per jaar (bron: tweede corrector; kolompercentages per examen per jaar optellend tot 100%)

Jaar		NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
2016	Zeer zorgvuldig	14	17	15	26	25	15	11	16	13	12	10	8
	Zorgvuldig	47	52	51	50	51	49	56	45	52	59	57	59
	Min of meer zorgvuldig	22	22	25	18	19	20	17	22	24	17	23	23
	Slordig	13	8	7	6	5	13	13	14	8	8	7	7
	Zeer slordig	5	1	2	0	1	4	3	2	2	3	3	3
2011	Zeer zorgvuldig	9	12	7	23	15	12	9	6	6	11	9	10
	Zorgvuldig	53	57	59	48	62	59	60	48	53	61	61	53
	Min of meer zorgvuldig	23	22	25	24	17	21	18	33	25	18	20	21
	Slordig	11	6	5	5	6	4	8	11	11	8	7	12
	Zeer slordig	5	3	3	0	0	3	4	2	4	2	3	4

Verschillen tussen de zorgvuldigheid van de eerste en tweede correctie

Een multivariate variantieanalyse met de correctie (eerste versus tweede) als binnen-subject-factor en jaar en examen als tussen-subjectenfactoren laat zien dat het hoofdeffect van correctie significant is ($p < 0,001$), evenals de tweeweg interactie met jaar ($p < 0,001$) en met examen ($p < 0,001$) maar dat de drieweg interactie tussen correctie, jaar en examen dat niet is ($p = 0,536$). De interpretatie van het hoofdeffect van correctie is dat tweede correctoren

kritischer zijn over de zorgvuldigheid van de eerste correctie dan eerste correctoren over de zorgvuldigheid van de tweede correctie. Zo corrigeerde in 2016 volgens de tweede correctoren slechts 14% van de eerste correctoren zeer zorgvuldig, terwijl volgens de eerste correctoren 37% van de tweede correctoren zeer zorgvuldig naaktek (zie tabel 3.4.7).

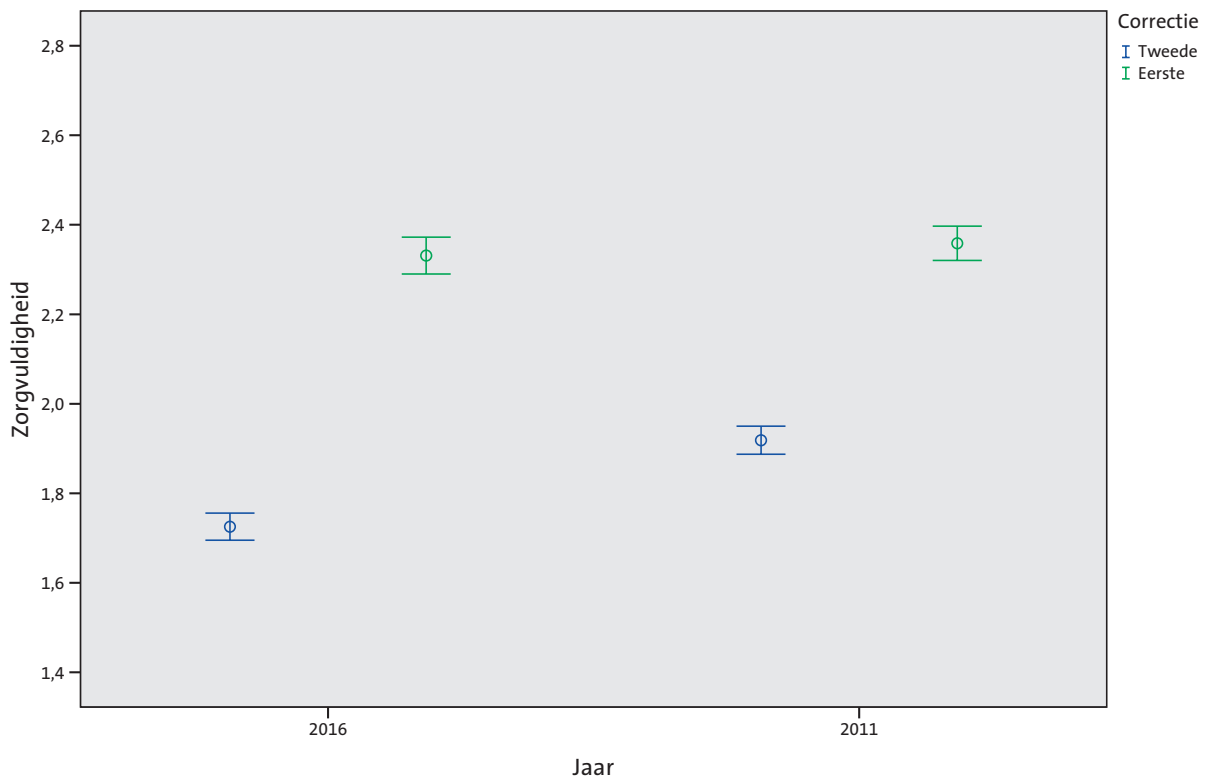
Tabel 3.4.7 Zorgvuldigheid van de tweede correctie (bron: eerste corrector) versus de eerste correctie (bron: tweede corrector; kolompercentages per jaar per correctieprocedure optellend tot 100%)

Jaar		Zorgvuldigheid van de tweede correctie volgens de eerste correctoren*	Zorgvuldigheid van de eerste correctie volgens de tweede correctoren
2016	Zeergvuldig	37	14
	Zorgvuldig	54	54
	Min of meer zorgvuldig	8	20
	Slordig	1	9
	Zeerglordig	0	3
2011	Zeergvuldig	26	10
	Zorgvuldig	58	58
	Min of meer zorgvuldig	14	21
	Slordig	1	8
	Zeerglordig	1	3

* Deze percentages wijken enigszins af van die in de tabellen 3.4.3 en 3.4.5 omdat de antwoorden van de eerste correctoren die de vraag naar de zorgvuldigheid van de tweede correctie niet konden beantwoorden in de analyse buiten beschouwing zijn gelaten.

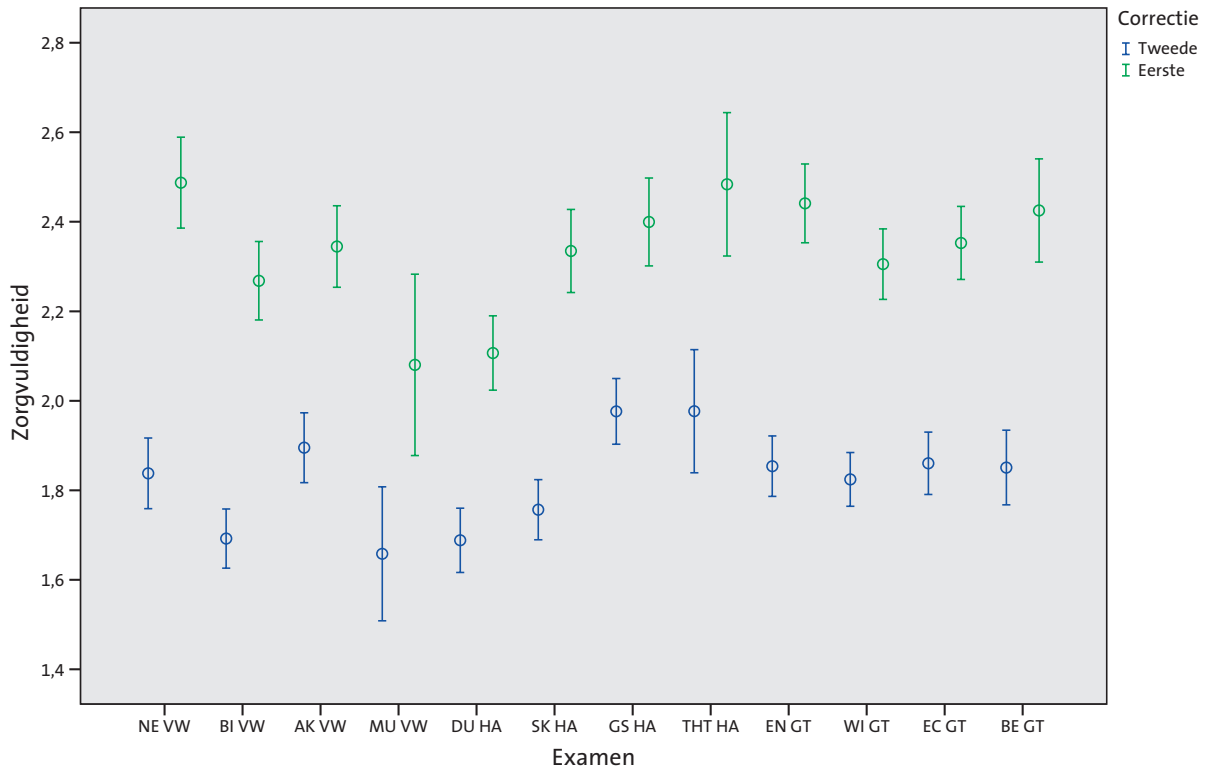
Verschillen tussen 2016 en 2011

De significante tweeweg interactie van correctie met jaar en examen betekent dat het verschil tussen de zorgvuldigheid van de eerste en tweede correctie niet onafhankelijk is van het jaar en het examen. Figuur 17 toont de gemiddelden voor de eerste en tweede correctie per jaar zoals uitgedrukt op een schaal van 1 (zeergvuldig) tot 5 (zeerglordig). De figuur bevestigt dat tweede correctoren kritischer zijn over de zorgvuldigheid van de eerste correctie dan eerste correctoren over de zorgvuldigheid van de tweede correctie. Te zien is verder dat de meningen van de eerste en tweede correctoren over de zorgvuldigheid van de ander in 2016 wat verder uiteenlopen dan in 2011.



Figuur 17 Verschil in zorgvuldigheid tussen de tweede correctie (bron: eerste corrector) en de eerste correctie (bron: tweede corrector) per jaar

Figuur 18 maakt duidelijk dat het verschil tussen de zorgvuldigheid van de eerste en tweede correctie voor het ene examen groter is dan voor het andere examen. Zo liggen de meningen van de tweede en eerste correctoren over de zorgvuldigheid van de eerste respectievelijk tweede correctie bij Nederlands verder uit elkaar dan bij Duits.



Figuur 18 *Verskil in zorgvuldigheid tussen de tweede correctie (bron: eerste corrector) en de eerste correctie (bron: tweede corrector) per examen*

3.4.3 Soepelheid van de correctie

Behalve volledig en zorgvuldig moet de correctie ook objectief zijn. Eerste en tweede correctoren mogen kandidaten bijvoorbeeld niet bevoordelen of benadelen door een te soepele of te strenge beoordeling. In de vragenlijst zijn eerste correctoren bevraagd naar de soepelheid van de tweede correctoren en de tweede correctoren zijn bevraagd naar de soepelheid van de eerste correctoren.

Soepelheid van de tweede correctie (volgens de eerste corrector)

In 2016 heeft - volgens de eerste correctoren - 7% van de tweede correctoren te streng beoordeeld, 21% beoordeelde streng, 67% noch te soepel noch te streng, 2% soepel en 0% te soepel. Drie procent van eerste correctoren geeft te kennen de vraag naar de soepelheid van de tweede correctie niet te kunnen beantwoorden.

Tabel 3.4.8 De soepelheid van de tweede correctie per examen (bron: eerste correctoren; kolompercentages optellend tot 100%)

	2016	2011
Te soepel (d.w.z. kandidaten zijn bevoordeeld)	0	0
Soepel	2	3
Noch te soepel noch te streng	67	65
Streng	21	21
Te streng (d.w.z. kandidaten zijn benadeeld)	7	5
Ik kan niet overzien hoe soepel of streng de tweede corrector heeft nagekeken	3	4

Verschillen tussen 2016 en 2011

In de analyse van de verschillen tussen jaren en examens zijn de antwoorden van de docenten die deze vraag niet konden beantwoorden buiten beschouwing gelaten. De verschillen tussen beide jaren zijn niet significant ($p = 0,078$). In 2016 vonden eerste correctoren de tweede correctie even soepel of streng als in 2011.

Verschillen tussen examens in 2016

De toetsing van de verschillen tussen examens gaf significantie te zien ($p < 0,001$). Zo kijken tweede correctoren biologie in 2016 volgens de eerste correctoren strenger na dan docenten wiskunde (zie tabel 3.4.9).

Tabel 3.4.9 De soepelheid van de tweede correctie per examen per jaar (bron: eerste correctoren; kolompercentages per examen per jaar optellend tot 100%)

Jaar	Soepelheid	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WIGT	EC GT	BEGT
2016	Te soepel	1	0	0	0	0	0	0	0	0	0	0	1
	Soepel	1	2	2	0	2	3	3	4	1	3	2	5
	Noch te soepel noch te streng	68	59	69	65	74	71	73	51	71	78	66	64
	Streng	24	30	23	28	18	20	17	33	21	15	23	22
	Te streng	6	10	6	7	6	6	8	12	7	3	9	8
2011	Te soepel	1	0	0	0	0	1	0	0	1	0	0	0
	Soepel	2	2	3	6	4	2	3	4	3	6	4	5
	Noch te soepel noch te streng	62	67	74	68	67	77	76	72	61	69	65	66
	Streng	29	25	18	21	23	16	15	18	26	20	24	23
	Te streng	5	6	5	5	6	4	5	6	9	5	7	6

Soepelheid van de eerste correctie (volgens de tweede corrector)

De vraag naar de soepelheid van de correctie is ook gesteld aan de tweede correctoren. Nu was de eerste correctie het onderwerp. Volgens de tweede correctoren heeft 9% van de eerste correctoren in 2016 te soepel beoordeeld in de zin dat kandidaten bevoordeeld zijn, 17% beoordeelde soepel, bijna twee derde (65%) was noch te soepel noch te streng, 7% was streng en 1% was te streng in de zin dat kandidaten benadeeld zijn.

Tabel 3.4.10 De soepelheid van de eerste correctie (bron: tweede corrector; kolompercentages optellend tot 100%)

	2016	2011
Te soepel (d.w.z. kandidaten zijn bevoordeeld)	9	10
Soepel	17	16
Noch te soepel noch te streng	65	66
Streng	7	8
Te streng (d.w.z. kandidaten zijn benadeeld)	1	1

Verschillen tussen 2016 en 2011

De verschillen tussen beide jaren zijn niet significant ($p = 0,960$). In 2016 vonden tweede correctoren de eerste correctie even soepel of streng als in 2011.

Verschillen tussen examens in 2016

De toetsing van de verschillen tussen examens gaf significantie te zien ($p < 0,001$). Zo kijken in 2016 eerste correctoren Duits volgens de tweede correctoren iets strenger na dan docenten tehatex en Nederlands.

Tabel 3.4.11 De soepelheid van de eerste correctie per examen per jaar (bron: tweede corrector; kolompercentages per examen per jaar optellend tot 100%)

Jaar	Soepelheid	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WIGT	EC GT	BEGT
2016	Te soepel	11	13	10	4	3	10	10	18	10	7	7	11
	Soepel	19	18	19	14	12	15	22	21	23	13	19	18
	Noch te soepel noch te streng	61	60	65	68	76	68	61	54	59	72	63	62
	Streng	7	8	6	14	8	6	7	7	7	7	9	8
	Te streng	1	1	0	0	0	1	1	0	2	1	0	2
2011	Te soepel	12	10	8	5	6	7	12	19	13	6	9	14
	Soepel	20	19	21	24	17	11	10	17	22	14	13	17
	Noch te soepel noch te streng	60	63	65	65	69	74	69	58	56	71	67	58
	Streng	8	7	6	6	8	7	9	6	8	8	9	10
	Te streng	1	0	0	0	1	1	0	0	2	1	1	1

Verschillen tussen de soepelheid van de eerste en tweede correctie

Hiervoor constateerden wij dat tweede correctoren minder positief denken over de zorgvuldigheid van de eerste correctie dan eerste correctoren over de zorgvuldigheid van de tweede correctie. Met betrekking tot de soepelheid van de beoordeling vinden we een soortgelijk patroon. De multivariate variantieanalyse laat zien dat het hoofdeffect van correctie en de tweeweg interactie met examen significant zijn ($p < 0,001$), maar de tweeweg interactie tussen correctie en jaar en de drieweg interactie tussen correctie, jaar en examen zijn dat niet (p respectievelijk 0,066 en 0,026). Tweede correctoren vinden de eerste correctoren vaker (te) soepel dan dat eerste correctoren de tweede corrector (te) soepel vinden. Zo is in 2016 van de tweede correctoren 9% van mening dat de eerste corrector te soepel is geweest, terwijl slechts 0% van de eerste correctoren vindt dat de tweede correctie te soepel was. In datzelfde jaar vinden 21% en 7% van de eerste correctoren de tweede correctie streng respectievelijk te streng,

terwijl de overeenkomstige percentages voor de mening van de tweede correctoren over de strengheid van de eerste corrector slechts 7% en 1% zijn.

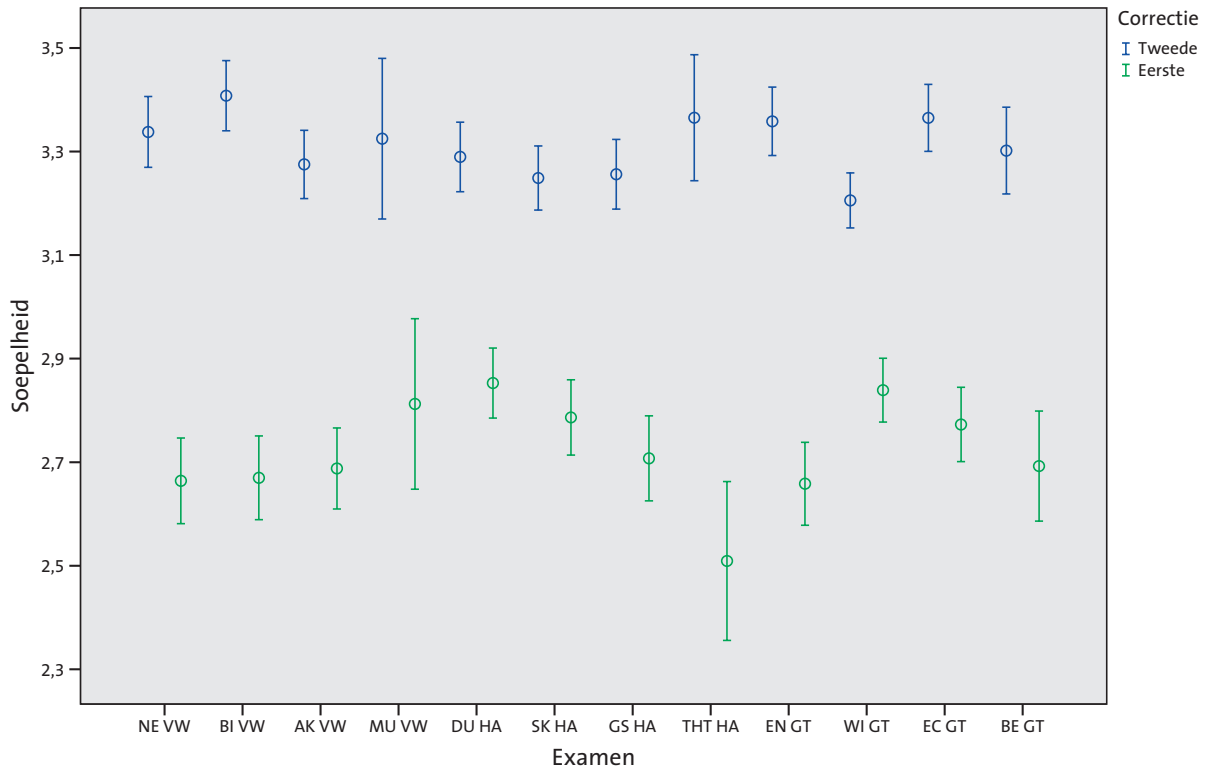
Tabel 3.4.12 Soepelheid van de eerste versus de tweede correctie (kolompercentages per correctieprocedure per jaar optellend tot 100%)

Jaar		Soepelheid van de tweede correctie volgens de eerste correctoren*	Soepelheid van de eerste correctie volgens de tweede correctoren
2016	Te soepel	0	9
	Soepel	2	17
	Noch te soepel noch te streng	70	65
	Streng	21	7
	Te streng	7	1
2011	Te soepel	0	10
	Soepel	3	16
	Noch te soepel noch te streng	69	66
	Streng	22	8
	Te streng	6	1

* Deze percentages wijken enigszins af van die in tabel 3.4.8 omdat de antwoorden van degenen die deze vraag niet konden beantwoorden in de analyse buiten beschouwing zijn gelaten.

Verschillen tussen jaren en examens

De significante interactie tussen correctie en examens betekent dat het verschil tussen de soepelheid van de eerste en tweede correctie niet onafhankelijk is van het examen. Figuur 19 toont de gemiddelden voor de eerste en tweede correctie per jaar op een schaal van 1 (zeer soepel) tot 5 (zeer streng). Uit de figuur komt naar voren dat het verschil voor het ene examen groter is dan voor het andere examen. Zo is het verschil bij wiskunde en Duits veel kleiner dan bij Nederlands, biologie of tehatex.



Figuur 19 Verschil in soepelheid tussen de tweede correctie (bron: eerste corrector) en eerste correctie (bron: tweede corrector) per examen

3.4.4 Verbeteringsmaatregelen

In 2016 zijn de docenten vier stellingen over potentiële maatregelen ter verbetering van de correctie voorgelegd. De eerste twee stellingen betreffen majeure wijzigingen in de huidige correctiesystematiek:

- Het is wenselijk dat de correctievolgorde wordt omgekeerd (d.w.z.: eerst de docent van de andere school en daarna de 'eigen' docent);
- Het is wenselijk dat de eerste en tweede correctie anoniem, door twee onafhankelijke correctoren worden uitgevoerd, dus niet meer door de 'eigen' docent.

De derde en vierde stelling gaan over de adviezen voor het corrigeren die de vakinhoudelijke verenigingen tijdens de correctieperiode publiceren:

- Het is wenselijk dat de vakinhoudelijke verenigingen adviezen voor het corrigeren ter beschikking stellen;
- Het is wenselijk dat deze adviezen uiterlijk twee werkdagen na de landelijke bespreking gepubliceerd worden.

In de toelichting bij de derde en vierde stelling is de docenten erop gewezen dat het hier NIET gaat om de officiële, bindende aanvullingen die het College voor Toetsing en Examens (CVTE) publiceert naar aanleiding van onvolkomenheden in het correctievoorschrift.

Er zijn signalen dat tweede correctoren het door de eerste corrector nagekeken werk soms pas zeer laat in de examencampagne ontvangen. Het vroegtijdig maken van een afspraak over de aanlevering van het werk kan voorkomen dat de tweede corrector in tijdnood komt. Vandaar dat de tweede correctoren is gevraagd of zij met de eerste corrector een afspraak maakten over de datum van aanlevering en hoeveel dagen voor het einde van de correctieperiode zij de examenwerken ontvingen.

Tot slot is nog een vraag over het gezamenlijk overleg tussen eerste en tweede corrector gesteld. Uit het vorige onderzoek (Kuhlemeier & Kremers, 2013) is bekend dat er grote verschillen zijn in de duur van het overleg waarin eerste en tweede gezamenlijk de uiteindelijke scores bepalen. Dit roept de vraag op in hoeverre eerste en tweede correctoren erin slagen om alle geconstateerde scoreverschillen te bespreken. In 2016 is de tweede correctoren daarom gevraagd of zij alle verschillen (plussen en minnen) met de eerste corrector besproken hebben.

Omkering van de correctievolgorde

Docenten zijn in overgrote meerderheid tegen omkering van de correctievolgorde: 66% is het er helemaal mee oneens, 14% mee oneens, 10% heeft geen mening of weet het niet, 6% is het ermee eens en 4% is het er helemaal mee eens.

Tabel 3.4.13 *Mening over de wenselijkheid van omkering van de correctievolgorde (kolompercentages optellend tot 100%)*

	%
Helemaal mee oneens	66
Mee oneens	14
Geen mening/Weet niet	10
Mee eens	6
Helemaal mee eens	4

Docenten van het ene vak staan minder negatief tegenover omkering van de correctievolgorde dan die van het andere vak ($p < 0,001$). Zo staan docenten Engels en economie daar minder afwijzend tegenover dan docenten scheikunde, geschiedenis en Nederlands.

Tabel 3.4.14 *Mening over de wenselijkheid van omkering van de correctievolgorde per examen (kolompercentage per examen optellend tot 100%)*

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WIGT	EC GT	BE GT
Helemaal mee oneens	75	63	68	71	64	79	72	65	58	66	64	51
Mee oneens	9	13	13	9	17	10	13	14	19	15	13	21
Geen mening/Weet niet	9	12	8	14	12	4	8	12	12	10	11	15
Mee eens	3	8	6	3	5	5	4	4	7	5	6	8
Helemaal mee eens	3	5	5	3	2	2	3	4	4	5	6	4

Onafhankelijke eerste en tweede correctie

De antwoorden op de vraag naar de wenselijkheid van onafhankelijke eerste en tweede correctie zijn weergegeven in tabel 3.4.15. Deze verbeteringsmaatregel kent weinig voorstanders: 58% is het er helemaal mee oneens, 19% mee oneens en 10% heeft geen mening of weet het niet; 7% is het ermee eens en 5% is het er helemaal mee eens.

Tabel 3.4.15 *Mening over onafhankelijke eerste en tweede correctie (kolompercentages optellend tot 100%)*

	%
Helemaal mee oneens	58
Mee oneens	19
Geen mening/Weet niet	10
Mee eens	7
Helemaal mee eens	5

De verschillen tussen examens zijn significant ($p < 0,001$). Zo staan docenten muziek sterker afwijzend tegenover onafhankelijke eerste en tweede correctie dan hun collega's van andere vakken.

Tabel 3.4.16 *Mening over onafhankelijke eerste en tweede correctie per examen (kolompercentages per examen optellend tot 100%)*

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	ENG T	WIG T	EC GT	BEG T
Helemaal mee oneens	53	58	56	79	57	63	55	60	55	64	58	51
Mee oneens	19	19	18	9	18	18	16	20	22	18	23	25
Geen mening/Weet niet	12	11	10	10	13	9	10	11	11	9	8	14
Mee eens	9	7	9	0	7	6	12	4	8	5	8	6
Helemaal mee eens	7	6	7	2	4	5	7	4	5	4	3	4

Adviezen van de vakinhoudelijke verenigingen voor het corrigeren

De adviezen voor het corrigeren van de vakinhoudelijke verenigingen kunnen op brede ondersteuning van docenten rekenen (zie tabel 3.4.17). Bijna de helft (48%) is het er helemaal mee eens, 37% mee eens, 8% heeft geen mening of weet het niet, 4% is het er mee oneens en 3% helemaal mee oneens.

Tabel 3.4.17 *Mening over de wenselijkheid van adviezen voor het corrigeren van de vakinhoudelijke verenigingen (kolompercentages optellend tot 100%)*

	%
Helemaal mee oneens	3
Mee oneens	4
Geen mening/Weet niet	8
Mee eens	37
Helemaal mee eens	48

Er zijn verschillen tussen examens in de mate waarin docenten de adviezen van de vakinhoudelijke verenigingen op prijs stellen ($p < 0,001$). Zo zijn docenten aardrijkskunde, geschiedenis, scheikunde, Nederlands en biologie het er nog sterker mee eens dan docenten Engels, Duits en wiskunde.

Tabel 3.4.18 *Mening over de wenselijkheid van adviezen voor het corrigeren van de vakinhoudelijke verenigingen per examen (kolompercentages per examen optellend tot 100%)*

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Helemaal mee oneens	4	1	1	5	7	2	0	3	6	5	4	2
Mee oneens	2	4	1	2	9	1	1	7	8	5	3	8
Geen mening/Weet niet	2	2	1	10	10	3	1	16	21	12	8	12
Mee eens	24	34	24	37	41	27	20	44	44	48	43	51
Helemaal mee eens	67	60	73	46	33	66	77	30	22	30	41	27

Tijdige verstrekking van de adviezen voor de correctie

De overgrote meerderheid van de docenten zou het op prijs stellen als de adviezen voor de correctie uiterlijk twee werkdagen na de landelijke bespreking gepubliceerd worden. Bijna de helft (46%) kiest voor helemaal mee eens, 37% voor mee eens, 10% voor heeft geen mening of weet het niet, 4% voor mee oneens en 3% voor helemaal mee oneens.

Tabel 3.4.19 *Mening over de wenselijkheid van tijdige verstrekking van de adviezen voor de correctie (kolompercentages optellend tot 100%)*

	%
Helemaal mee oneens	3
Mee oneens	4
Geen mening/Weet niet	10
Mee eens	37
Helemaal mee eens	46

De verschillen tussen examens zijn significant ($p < 0,001$). Zo stellen docenten Nederlands, aardrijkskunde, biologie, scheikunde en geschiedenis tijdige verstrekking van de adviezen nog sterker op prijs dan docenten Duits, Engels en wiskunde.

Tabel 3.4.20 *Mening over de wenselijkheid van tijdige verstrekking van de adviezen voor de correctie per examen (kolompercentages per examen optellend tot 100%)*

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Helemaal mee oneens	3	1	1	2	6	1	0	3	5	4	4	1
Mee oneens	2	3	1	4	7	1	3	2	6	5	3	7
Geen mening/Weet niet	2	6	1	9	14	5	3	18	26	14	9	14
Mee eens	28	33	27	35	39	32	27	46	41	44	42	50
Helemaal mee eens	65	59	69	51	34	61	65	31	21	33	42	27

Afspraak tussen eerste en tweede corrector over de aanlevering van het werk

Van de tweede correctoren heeft 8% met de eerste corrector een afspraak over de aanlevering van het examenwerk gemaakt. De verschillen tussen examens zijn significant ($p = 0,004$). Zo maken tweede correctoren muziek en tehatex vaker een afspraak met de eerste corrector dan hun collega's aardrijkskunde en scheikunde.

Tabel 3.4.21 *Het maken van een afspraak met de eerste corrector over het aanleveren van het examenwerk (bron: tweede corrector; kolompercentages optellend tot 100%)*

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Nee	93	91	95	84	90	95	95	89	89	94	91	90
Ja	7	9	5	16	10	5	5	11	11	6	9	10

Moment van aanlevering van het examenwerk

De tweede correctoren ontvangen het examenwerk gemiddeld zeven dagen voor het einde van de correctieperiode (zie tabel 3.4.22). De verschillen tussen docenten zijn groot ($p < 0,001$). De percentielscores geven het percentage respondenten met een even hoog of lager aantal dagen. De 10% die de envelop met het examenwerk het laatst ontving (P10) heeft het werk drie dagen of nog later voor het einde van de correctieperiode ontvangen. De 10% die het werk het vroegst ontving (P90) kreeg dat twaalf dagen of nog eerder voor het einde van de correctieperiode.

Tabel 3.4.22 *Verdeling van het aantal dagen voor het einde van de correctieperiode dat de tweede corrector het examenwerk van de eerste corrector ontving (bron: tweede corrector)*

N	Min	Max	Gem	Se-Gem	Stddev	P10	P25	P75	P90
2865	0	23	7,17	0,07	3,68	3,00	5,00	10,00	

Bij het ene examen ontvangt de tweede corrector het werk wat eerder dan bij het andere examen ($p < 0,001$). Zo konden docenten Nederlands al negen dagen voor het einde van de correctieperiode met het nakijken beginnen en docenten aardrijkskunde pas zes dagen ervoor.

Tabel 3.4.23 *Gemiddeld aantal dagen voor het einde van de correctieperiode dat de tweede corrector het examenwerk van de eerste corrector ontving per examen (bron: tweede corrector)*

NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
9,03	6,50	5,91	5,87	6,62	7,48	7,75	6,36	6,67	7,66	6,54	8,12

Volledigheid van het bespreken van de verschillen tijdens het gezamenlijk overleg

In 2016 is de tweede correctoren daarom gevraagd of zij ALLE verschillen (plussen en minnen) met de eerste corrector besproken hebben. Veertien procent zegt dat niet gedaan te hebben. De verschillen tussen examens zijn significant ($p < 0,001$) en weergegeven in tabel 3.4.24. Het percentage docenten dat alle scoreverschillen bespreekt, is bijvoorbeeld bij geschiedenis en tehatex veel lager dan bij Nederlands, biologie, economie, Engels, Duits en muziek. De twee laatstgenoemde examens spannen de kroon aangezien alle docenten aangeven alle scoreverschillen te hebben besproken.

Tabel 3.4.24 Antwoorden op de vraag of alle scoreverschillen besproken zijn (kolompercentages optellend tot 100%)

	NE VW	BI VW	AK VW	MU VW	DU HA	SK HA	GS HA	THT HA	EN GT	WI GT	EC GT	BE GT
Nee	13	13	22	0	0	18	33	30	2	15	9	10
Ja	87	87	78	100	100	82	67	70	98	85	91	90

Werkwijze als niet alle scoreverschillen besproken worden

De 450 tweede correctoren die niet alle scoreverschillen met de eerste corrector bespraken, is gevraagd hoe zij daarbij te werk gegaan zijn. De respondenten konden daarbij kiezen uit een of meer van de werkwijzen die zijn opgesomd in tabel 3.4.25. Veertig procent van de tweede correctoren kijkt naar de totaalscores en bespreekt de verschillen alleen als zijn of haar totaalscore opvallend afwijkt van die van de eerste corrector. Een derde (33%) kiest de vragen met de grootste scoreverschillen en bespreekt van deze vragen alle scoreverschillen. Ruim een vijfde (22%) kiest de kandidaten met de grootste scoreverschillen en bespreekt dan alle verschillen. Slechts 2% vertrouwt erop dat de eerste corrector de scores van de niet besproken kandidaten en/of vragen overeenkomstig (de wel besproken gevallen) zal aanpassen. Hierbij merken we op dat de respondenten meer antwoorden konden aankruisen en de ene werkwijze de andere dus niet uitsluit. Vanwege het kleine aantal respondenten is analyse van de verschillen tussen examens hier niet zinvol.

Tabel 3.4.25 Werkwijze als niet alle scoreverschillen met de eerste corrector besproken zijn

	N	%
Ik kies de vragen met de grootste scoreverschillen en bij deze vragen bespreek ik alle verschillen (plus en min)	450	33
Ik kies de kandidaten met de grootste scoreverschillen en bij deze kandidaten bespreek ik alle verschillen (plus en min)	450	22
Ik kijk naar de totaalscores en bespreek verschillen alleen als mijn totaalscore opvallend afwijkt van de eerste corrector	450	40
Als ik niet alle kandidaten en/of vragen bespreek, vertrouw ik erop dat de eerste corrector de scores van de overige kandidaten/vragen overeenkomstig aanpast	450	2

Toelichtingen van docenten bij de vraag of zij alle scoreverschillen met de eerste corrector besproken hebben

In totaal 252 respondenten hebben gebruik gemaakt van de mogelijkheid om hun antwoord nader toe te lichten. De inhoud van de toelichtingen is geanalyseerd met behulp van een categorieënsysteem. Voor het uitgebreide verslag van de inhoudsanalyse wordt verwezen naar Loeffen en Eikelboom (in voorbereiding). De resultaten van de analyse zijn weergegeven in tabel 3.4.26. Van de respondenten die een toelichting gaven had 81% daarvoor aangegeven dat zij alle scoreverschillen met de eerste corrector hadden besproken en 19% had dat niet; in de responsgroep van alle docenten waren de overeenkomstige percentages 86% en 14%.

Tabel 3.4.26 *Bespreekprocedures tijdens het gezamenlijk overleg voor degenen die niet en wel alle scoreverschillen bespreken (kolompercentages optellend tot 100%)*

	Alle scoreverschillen besproken?		Totaal
	Nee	Ja	
Bespreking scoreverschillen per vraag			
Bespreekt alle vragen met scoreverschil	1	5	1
Bespreekt alleen vragen met scoreverschil groter dan 1 à 2 punten	5		4
Bespreekt de scoreverschillen per vraag totdat de eerste corrector geen uitleg meer nodig heeft	4		3
Bespreekt alleen vragen waarbij de eerste corrector het niet met zijn/haar score eens is	2	2	2
Bespreekt alleen de vragen waar inhoudelijke bezwaren zijn	7	5	7
Bespreking scoreverschillen per kandidaat			
Bespreekt alle kandidaten met scoreverschil	30	11	26
Bespreekt alleen kandidaten met scoreverschil groter dan 1 à 2 punten	4		3
Bespreekt voor een deel van de kandidaten alleen scoreverschillen per kandidaat groter dan 1 à 2 punten	1		0
Bespreekt alleen kandidaten waarbij de eerste corrector het niet met zijn/haar score eens is	2		2
Bespreekt alleen de vragen die voordeel opleveren voor de totaalscore van een kandidaat	6	2	5
Bespreking scoreverschillen zonder dat duidelijk is of dat per vraag of per kandidaat gebeurt			
Bespreekt alle verschijscores	6	55	15
Bespreekt alleen scoreverschillen groter dan 1 à 2 punten	3	2	3
Bespreekt alle scoreverschillen behalve spelfouten	1		1
Er waren te veel scoreverschillen om ze allemaal te kunnen bespreken	4	2	3
Alleen extreme scoreverschillen besproken	2		2
Niet alles besproken vanwege slordige correctie en/of grote weerstand bij eerste corrector	7		6
Bespreekt scoreverschillen waarbij de eerste corrector het niet met zijn/haar score eens is	9	2	8
Overig			
Overig on-topic	2	2	2
Off-topic	5	11	6
Totaal	100	100	100

De bespreekprocedures van de tweede correctoren die aangaven niet alle scoreverschillen met de eerste corrector te bespreken zijn weergegeven in de kolom 'Nee' (van tabel 3.4.26). Van hen hanteert 19% bespreking per vraag, 42% bespreking per kandidaat en van 32% is niet duidelijk of zij dat per vraag of per kandidaat doen (terwijl 7% van de toelichtingen on-topic was maar niet geclassificeerd kon worden en 5% off-topic was).

De bespreekprocedures van de tweede correctoren die aangaven wel alle scoreverschillen met de eerste corrector te bespreken zijn weergegeven in de kolom 'Ja' (van tabel 3.4.26). Van hen doet dat 12% per vraag, 13% per kandidaat en 61% bespreekt alle scoreverschillen zonder te specificeren of dat per vraag of per kandidaat gebeurt (terwijl 2% van de toelichtingen on-topic was maar niet geclassificeerd kon worden en 11% off-topic was).

Wellicht ten overvloede melden we nog dat de percentages in de kolom Totaal gebaseerd zijn op alle docenten die een toelichting gaven.

4 Samenvatting, discussie en aanbevelingen

4 Samenvatting, discussie en aanbevelingen

Dit rapport bevat het verslag van het tweede onderzoek naar de praktijk van de eerste en tweede correctie van de centraal schriftelijke examens in het voortgezet onderwijs. De inventarisatie is een vervolg op het onderzoek van 2011 (Kuhlemeier & Kremers, 2011).

Het onderzoek heeft tot doel een antwoord te geven op de volgende onderzoeksvragen:

- Hoeveel tijd kosten de eerste en tweede correctie? En in hoeverre ervaren docenten de correctie als belastend?
- Hoe wordt de eerste en tweede correctie op school gefaciliteerd? En in hoeverre zijn docenten tevreden over dit schoolbeleid?
- In hoeverre voeren docenten de tweede correctie volledig uit? Waarom wordt de tweede correctie niet altijd volledig uitgevoerd? En hoe denken docenten over mogelijke oplossingen?
- Wat vinden docenten van de voorgestelde omkering van de correctievolgorde en andere (vermeende) verbeteringen?
- Hoe denken eerste en tweede correctoren over de volledigheid, zorgvuldigheid en objectiviteit van elkaars beoordelingen? En welke maatregelen vinden docenten wenselijk om de kwaliteit van de correctie waar nodig te verbeteren?
- In hoeverre is de correctiepraktijk sinds 2011 veranderd? Omdat een groot deel van de enquêtevragen ook in 2011 gesteld is, kunnen we de gegevens van 2016 vergelijken met die van 2011.

Net als in 2011 zijn in 2016 twaalf examenvakken in het onderzoek betrokken, te weten Nederlands vwo, Biologie vwo, Aardrijkskunde vwo, Muziek vwo, Duits havo, Scheikunde havo, Geschiedenis havo, Tehatex havo, Engels gt, Wiskunde gt, Economie gt en Beeldende vakken gt. De vakken zijn gelijkelijk verdeeld over de vier typen vakken (talen, bèta, gamma en kunst) en de drie opleidingstypen (vwo, havo en gt).

Er is gekozen voor dezelfde twaalf examenvakken als in 2011. Dit is gedaan om de resultaten anno 2016 optimaal te kunnen vergelijken met die van toen. Van de 6148 verzonden vragenlijsten uit ons onderzoek naar de correctiepraktijk kwam 58% ingevuld retour. De rapportage is gebaseerd op de antwoorden van 3540 docenten.

Drie vakinhoudelijke verenigingen hebben de vragenlijst op eigen initiatief aan hun leden voorgelegd. Het betreft de Nederlandse Vereniging van Wiskundeleraren (NVvW), de Nederlandse Vereniging voor Onderwijs in de Natuurwetenschappen (NVON) en Levende Talen. De vragen zijn ook gesteld aan docenten van andere examens dan de twaalf uit ons onderzoek. De resultaten van dit flankerend onderzoek zijn opgenomen in bijlage 1 van dit onderzoeksrapport. In de hierna volgende samenvatting en discussie zal hier enkele keren naar verwezen worden.

Het vervolg van dit hoofdstuk bevat eerst een vrij uitvoerige beschrijving van de onderzoeksresultaten waarbij we zo dicht mogelijk bij de cijfermatige gegevens zijn gebleven (paragraaf 4.1). Daarna vatten we de belangrijkste resultaten en conclusies samen, doen we een eerste poging tot interpretatie en geven we alvast enkele aanbevelingen (paragraaf 4.2).

In dit laatste hoofdstuk wordt slechts zeer beperkt aandacht besteed aan de verschillen tussen de resultaten voor de twaalf examens. Bij vrijwel alle aspecten van de correctiepraktijk

constateerden wij grote verschillen tussen examens qua tijdbesteding, ervaren belasting, schoolbeleid, tevredenheid over het schoolbeleid, de volledigheid van de tweede correctie, de zorgvuldigheid en objectiviteit van de correctie en de wenselijkheid van (vermeende) verbeteringsmaatregelen. Voor het ene examen bleek het verschil tussen de correctiepraktijk in 2016 en 2011 veel groter of juist kleiner dan voor het andere examen. In het kader van een beknopte samenvatting is het niet mogelijk al deze verschillen tussen examens te beschrijven. De geïnteresseerde lezer wordt daartoe verwezen naar de hoofdtekst en naar de resultaten van het flankerend onderzoek van de drie vakinhoudelijke verenigingen in bijlage 1.

4.1 Samenvatting

Tijdbesteding en belasting

Correctielast en tijdbesteding aan de eerste correctie

In 2016 heeft de eerste corrector het werk van 39 examenkandidaten nagekeken en daar in totaal 21 klokuren aan besteed in ruim 4 dagen. De gemiddelde tijdbesteding per kandidaat was 38 minuten.

De correctielast en de tijdbesteding aan de eerste correctie blijkt sterk van docent tot docent te verschillen. De 10% met de minste correctielast keek het werk van hooguit 15 kandidaten na en besteedde daar tot 5 klokuren aan in hooguit 2 dagen. De 10% met de zwaarste correctielast keek het werk van minimaal 70 kandidaten na en besteedde daar minimaal 40 uur aan in 8 dagen of meer.

Correctielast en tijdbesteding aan de tweede correctie

In 2016 heeft de gemiddelde tweede corrector het werk van 40 examenkandidaten nagekeken, daar in totaal 13 klokuren aan besteed in ruim 3 dagen. De gemiddelde tijdsbesteding aan de tweede correctie per kandidaat is 23 minuten.

De ene docent besteedt veel meer tijd aan de tweede correctie dan de andere. De 10% met de minste correctielast heeft het werk van hooguit 17 kandidaten nagekeken en besteedde daar tot 4 klokuren aan in hooguit 1 dag. De 10% met de zwaarste correctielast keek het werk van minimaal 70 kandidaten na en besteedde daar minimaal 25 uur aan in minimaal 5 dagen.

Correctielast en tijdbesteding aan de eerste en tweede correctie tezamen

In 2016 besteedden docenten evenveel tijd aan de eerste en tweede correctie tezamen als in 2011. De gemiddelde correctietijd per kandidaat was eveneens gelijk. Wel hebben docenten de correctie over een wat langere periode gespreid: 7,7 dagen in 2016 versus 7,1 dagen in 2011.

Verskil tussen de tijdbesteding aan de eerste en tweede correctie

In 2016 kostte de eerste correctie veel meer tijd dan de tweede correctie: gemiddeld 21 versus 13 uur. In 2011 ging het om 22 versus 12 uur. Het flankerend onderzoek van de drie vakinhoudelijke verenigingen geeft een vergelijkbaar beeld te zien (zie bijlage 1).

Vergelijking van de tijdbesteding aan de eerste en tweede correctie in 2016 versus 2011

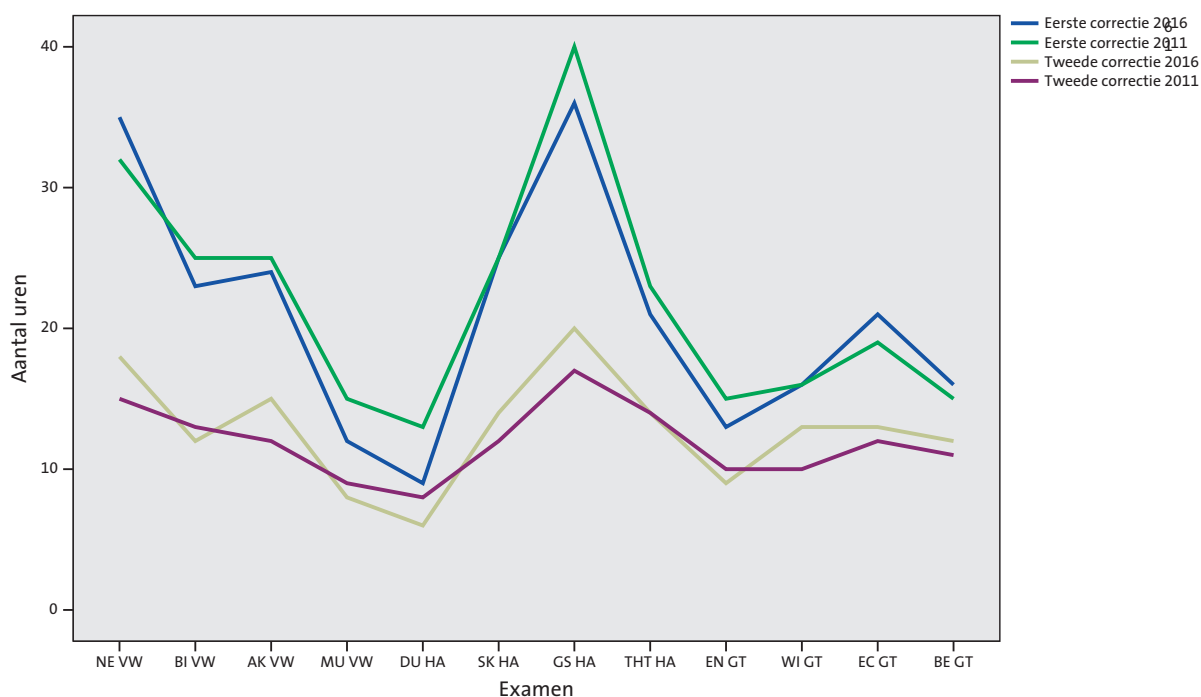
Aan de eerste correctie werd in 2016 gemiddeld bijna één uur minder besteed dan in 2011 (van 22 naar 21 uur). Deze afname was echter statistisch niet significant zodat we moeten concluderen dat de tijdbesteding aan de eerste correctie niet veranderd is. Ook de gemiddelde correctietijd per kandidaat was in beide jaren vergelijkbaar. Wel was het aantal dagen waarin de docenten in 2016 de eerste correctie uitvoerden iets groter: 4,5 dagen in 2016 versus 4,3 dagen in 2011.

De tweede correctie vergde in 2016 iets meer dan één uur meer tijd dan in 2011 (van 12 naar 13 uur). Aangezien dit verschil statistisch niet significant was, moeten we ervan uitgaan dat de tijdbesteding aan de tweede correctie gelijk gebleven is. De gemiddelde correctietijd per kandidaat was wat hoger (23 minuten in 2016 versus 21 minuten in 2011). Verder was het aantal dagen waarin de docenten in 2016 de tweede correctie uitvoerden iets groter: gemiddeld 3,2 dagen in 2016 versus 2,8 dagen in 2011.

Al met al is de tijdbesteding aan de eerste en tweede correctie sinds 2011 niet aantoonbaar veranderd.

Verschillen tussen examenvakken in tijdbesteding

Het nakijken van het ene examen kost meer tijd dan het andere terwijl ook het verschil tussen de tijdbesteding in 2016 en 2011 van examen tot examen verschilt. Dat wordt duidelijk uit figuur 4.1 waarin voor elk examen het totaal aantal bestede uren aan de eerste en tweede correctie in 2016 en 2011 visueel is weergegeven. De figuur laat zien dat de verschillen tussen de twaalf vakken groot zijn. Zo besteden docenten geschiedenis en Nederlands veel meer tijd aan de eerste en tweede correctie dan docenten Duits, Engels en muziek. Uit het flankerend onderzoek van de drie vakinhoudelijke verenigingen (zie bijlage 1) blijkt overigens dat er ook binnen hetzelfde vak grote verschillen kunnen bestaan tussen de diverse examens voor de opleidingstypen vmbo, havo en vwo.

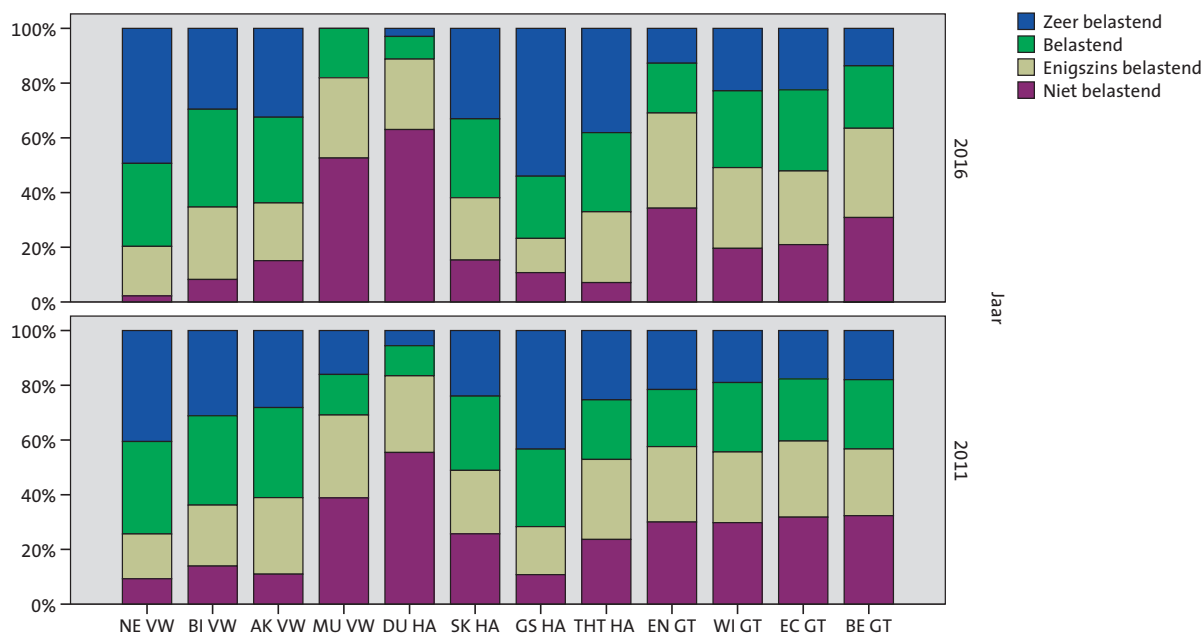


Figuur 4.1 Tijdbesteding in uren aan de eerste en tweede correctie per examen per jaar

Ervaren belasting van de eerste en tweede correctie

In 2016 vond 54% van de docenten de **eerste** correctie (zeer) belastend, 35% enigszins belastend en 11% niet belastend. De drie vakinhoudelijke verenigingen vonden vergelijkbare percentages (zie bijlage 1). In 2016 vonden docenten de eerste correctie een grotere belasting dan in 2011. Zo vond 17% de eerste correctie in 2016 zeer belastend tegenover 13% in 2011.

In 2016 vindt 57% van de docenten de **tweede** correctie (zeer) belastend, 35% enigszins belastend en 8% niet belastend. De ervaren belasting van de tweede correctie in 2016 verschilt niet van die in 2011. Dat de verschillen tussen de examenvakken groot zijn, wordt duidelijk uit figuur 4.2. Zo ervaren docenten geschiedenis en Nederlands de tweede correctie als een veel grotere belasting dan docenten Duits, Engels, muziek en beeldende vakken. De figuur bevestigt ook dat de verschillen tussen de ervaren belasting in 2016 en 2011 ook binnen de afzonderlijke vakken klein zijn.



Figuur 4.2 Ervaren belasting door de tweede correctie per examen in 2016 (boven) en 2011 (onder)

Docenten vinden de tweede correctie over het algemeen sterker belastend dan de eerste correctie.

Facilitering door de school en tevredenheid

Schoolbeleid ten aanzien van de eerste correctie

In 2016 wordt de eerste correctie volgens 71% van de docenten op hun school niet gefaciliteerd. Van 16% zijn de uren voor de eerste correctie aanwijsbaar opgenomen in de jaartaak (bijvoorbeeld in de opslagfactor), 7% krijgt een vast bedrag per examenwerk, 2% kan onkosten declareren (bijvoorbeeld telefoonkosten) en 1% ontvangt een kleine attentie (bijv. een boekenbon). Omdat respondenten bij deze vraag meer alternatieven konden aankruisen, tellen de percentages niet op tot 100%.

In vergelijking met 2011 vinden wat minder docenten dat de eerste correctie op hun school niet gefaciliteerd wordt (van 75% naar 71%). Ook zijn de uren voor de eerste correctie wat vaker opgenomen in de jaartaak (van 13% naar 16%). Van de andere kant krijgen docenten wat minder vaak een vast bedrag per examenwerk (van 9% naar 7%). Eerste correctoren declareren in 2016 even vaak onkosten en ontvangen even vaak een kleine attentie als in 2011.

In 2016 is 40% van de docenten (zeer) ontevreden met de manier waarop de school de eerste correctie faciliteert, 17% is (zeer) tevreden en 43% is tevreden noch ontevreden. In 2016 zijn docenten even (on)tevreden over de manier waarop de schoolleiding de eerste correctie faciliteert als in 2011.

Schoolbeleid ten aanzien van de tweede correctie

In 2016 vindt ruim de helft (52%) van de docenten dat de tweede correctie op hun school niet gefaciliteerd wordt. Van 6% zijn de uren voor de tweede correctie aanwijsbaar opgenomen in de jaartaak (bijv. in de opslagfactor), 31% krijgt een vast bedrag per examenwerk, 3% kan onkosten declareren (bijv. telefoonkosten) en 1% ontvangt een kleine attentie (bijv. een boekenbon). Omdat respondenten bij deze vraag meer alternatieven konden aankruisen, tellen de percentages niet op tot 100%.

In vergelijking met 2011 zijn docenten in 2016 vaker van mening dat de tweede correctie op hun school niet gefaciliteerd wordt (van 48% naar 52%). Ook krijgen zij minder vaak een vast bedrag per examenwerk (van 39% naar 31%). De verschillen bij de beleidsmaatregelen 'onkosten declareren', 'kleine attentie' en 'jaartaak' zijn verwaarloosbaar klein.

In 2016 is 51% van de docenten (zeer) ontevreden over de wijze waarop de school de tweede correctie faciliteert, 18% is (zeer) tevreden en 32% is tevreden noch ontevreden.

Verschillen tussen het schoolbeleid voor de eerste en tweede correctie

Scholen hebben vaker een beleid voor de tweede correctie dan voor de eerste correctie (in 2016: 48% versus 29%). Voor de tweede correctie ontvangen docenten vaker een vast bedrag per examenwerk dan voor de eerste correctie (31% versus 7%). Daar staat tegenover dat de uren voor de tweede correctie minder vaak in de jaartaak zijn opgenomen (2016: 6% versus 16%). Het vergoeden van onkosten (2016: 3% versus 2%) en het verstrekken van een kleine attentie (1% versus 1%) zijn bij de tweede correctie ongeveer even zeldzaam als bij de eerste correctie.

Verschillen tussen de tevredenheid met het schoolbeleid voor de eerste en tweede correctie

Zowel in 2016 als in 2011 zijn docenten minder tevreden met het schoolbeleid voor de tweede correctie dan met dat voor de eerste correctie.

De gegevens over de facilitering van de eerste en tweede correctie op school en de tevredenheid met het schoolbeleid komen overeen met die van de drie vakinhoudelijke verenigingen (zie bijlage 1).

De volledigheid van de tweede correctie

Volgens de wettelijke voorschriften moeten tweede correctoren de antwoorden van alle kandidaten op alle vragen corrigeren.

Wel of geen tweede correctie

Volgens opgave van de eerste corrector is het werk van de kandidaten in 0,5% van de gevallen niet door een tweede corrector nagekeken. In 2011 was dat 0,6%. In 2016 wordt de tweede correctie dus even weinig niet uitgevoerd als in 2011.

Volledigheid van de tweede correctie volgens de eerste corrector

De eerste correctoren is gevraagd naar hun mening over de volledigheid van de tweede correctie. In 2016 heeft 85% van de tweede correctoren (zeer) volledig nagekeken, 10% min of meer volledig, 2% (zeer) onvolledig en 4% kon niet overzien hoe volledig de tweede corrector had nagekeken.

In 2016 is de tweede correctie volgens de eerste correctoren vollediger uitgevoerd dan in 2011. Zo vond in 2011 nog slechts 31% de tweede correctie zeer volledig tegenover 44% in 2016.

Volledigheid van de tweede correctie volgens de tweede corrector

In de vragenlijst zijn de docenten verschillende vragen over de volledigheid van de tweede correctie gesteld. Twee ervan waren geformuleerd in de vorm van stellingen.

De eerste stellingvraag beschrijft de situatie waarin de tweede corrector het examenwerk geheel overeenkomstig de wettelijke voorschriften volledig nakijkt. De formulering was: 'Ik heb het examenwerk van alle kandidaten volledig nagekeken'. In 2016 kan 95% van de tweede correctoren zich (helemaal) met deze stelling vereenzelvigen. Zij kozen voor 'Klopt helemaal' of 'Klopt'. De overige tweede correctoren kozen voor 'Klopt niet' (5%) of 'Klopt helemaal niet' (0%).

De docenten is de ruimte geboden om hun antwoorden op de vragen naar de volledigheid van de tweede correctie toe te lichten (zie ook paragraaf 3.3.6). Allereerst bespreken we de toelichtingen van degenen die bij de eerste stellingvraag aangaven **(helemaal) volledig** na te kijken. In deze groep heeft blijkens de toelichtingen 71% daadwerkelijk volledig nagekeken, 11% keek de open vragen na maar de meerkeuzevragen niet of slechts gedeeltelijk, 5% keek niet alle kandidaten en/of niet alle vragen na (en 3% gaf een toelichting die niet gecodeerd kon worden en 10% gaf een toelichting die niet met de vraagstelling in verband stond). Het gegeven dat 16% van hen toch een niet-volledige correctiewijze toepast, geeft aanleiding om de antwoorden op de stellingvraag met enige voorzichtigheid te interpreteren. Kennelijk vindt een substantieel deel van de docenten dat zij (helemaal) volledig nakijken terwijl zij dat in werkelijkheid niet doen.

De toelichtingen van de degenen die bij de eerste stellingvraag aangaven **niet volledig** na te kijken, zijn als volgt verdeeld. Volgens hun toelichtingen heeft 92% daadwerkelijk onvolledig nagekeken, 4% keek toch alle kandidaten en alle vragen na en 3% gaf een antwoord dat niet aan de vraagstelling gerelateerd was. Hierbij merken we op dat deze percentages (na afronding op een geheel getal) optellen tot 99% in plaats van tot 100%.

In 2016 is de tweede correctie - aldus de tweede correctoren zelf - vollediger uitgevoerd dan in 2011. De verschillen tussen 2016 en 2011 zijn groot. Vond bijvoorbeeld in 2011 nog maar 44% van de tweede correctoren dat de tweede correctie helemaal volledig was uitgevoerd, in 2016 was dit percentage gestegen naar 72%.

De tweede stellingvraag beschrijft een niet-volledige correctiestrategie waarbij alleen een deel van de werken en/of vragen wordt nagekeken. De precieze formulering was: "Ik heb eerst een deel van de werken/vragen nagekeken en ben daarna gestopt". Deze stelling is alleen in 2016 aan de tweede correctoren voorgelegd. Van de tweede correctoren vindt 93% dat deze stelling (helemaal) niet klopt en 7% vindt dat deze wel (helemaal) klopt. De laatsten kijken naar eigen zeggen niet volledig na.

Een opvallend grote groep tweede correctoren heeft de tweede stellingvraag overgeslagen (16%). Een eerste verklaring veronderstelt dat docenten die de eerste stellingvraag over de volledigheid van de tweede correctie positief beantwoordden de vervolgvraag over onvolledige tweede correctie als redundant beschouwden en om die reden onbeantwoord lieten. Een tweede verklaring is dat het een gevoelige vraag betreft die sommige docenten liever onbeantwoord laten.

Percentages nagekeken kandidaten en vragen

De volledigheid van de tweede correctie is een uitermate complex onderwerp. Docenten denken verschillend over wat in de praktijk onder volledigheid moet worden volstaan. Vandaar dat er behalve de twee stellingen nog meer vragen over dit onderwerp gesteld zijn. De tweede correctoren is ook gevraagd van welk deel van de kandidaten zij een of meer vragen hebben

gecorrigeerd en van welk deel van de vragen zij het werk van een of meer kandidaten hebben nagekeken. In 2016 heeft de gemiddelde tweede corrector naar eigen zeggen van 98% van de kandidaten een of meer vragen gecorrigeerd en van 96% van de vragen een of meer kandidaten nagekeken. Opgemerkt wordt dat het hier niet gaat om percentages volledige correctie. Het nakijkpercentage van 96% van de examenvragen betekent bijvoorbeeld niet noodzakelijkerwijs dat de tweede corrector deze vragen voor alle examenkandidaten heeft nagekeken.

De percentages nagekeken kandidaten en vragen zijn iets gestegen. In 2016 keken de tweede correctoren van 98% van de kandidaten een of meer vragen na tegenover 95% in 2011. In 2016 werden van 96% van de vragen een of meer kandidaten gecorrigeerd tegenover 91% in 2011.

De percentages tweede correctoren die 100% van de kandidaten of 100% van de vragen nakeken, liggen uiteraard iets lager. In 2016 keek 95% van de tweede correctoren volgens eigen opgave voor alle kandidaten een of meer vragen na tegenover 87% in 2011. In 2016 keek 87% van de tweede correctoren voor alle examenvragen een of meer kandidaten na tegenover 77% in 2011.

De gegevens over de volledigheid van de eerste en tweede correctie voor 2016 zijn vergelijkbaar met die van de drie vakinhoudelijke verenigingen (zie bijlage 1).

Mening over de noodzaak van volledige tweede correctie

In de vragenlijst zijn de respondenten twee stellingen voorgelegd over de noodzaak van integrale tweede correctie.

De eerste stelling luidde: 'De tweede correctie moet integraal worden uitgevoerd (d.w.z.: alle examenwerken en per kandidaat alle vragen)'. De overgrote meerderheid van de docenten blijkt overtuigd van de noodzaak van integrale tweede correctie: 82% is het er (helemaal) mee eens, 14% is het er (helemaal) niet mee eens en 4% heeft geen mening of weet het niet.

De tweede stelling beschrijft een vorm van niet-volledige tweede correctie: 'Als de tweede corrector een deel heeft nagekeken en tot de conclusie komt dat de eerste corrector goed heeft nagekeken, is het niet nodig om alles na te kijken'. Bijna driekwart van de docenten staat (sterk) afwijzend tegenover deze vorm van niet-volledige correctie: 71% was het er (helemaal) mee oneens, 21% (helemaal) mee eens en 8% heeft geen mening of weet het antwoord niet.

Oorzaken van niet-volledige tweede correctie

De docenten is in 2016 gevraagd naar de oorzaken van onvolledige tweede correctie. Tijdgebrek wordt door 82% als oorzaak genoemd, gevolgd door chronische overbelasting (48%) en onvoldoende financiële compensatie (38%). Ruim een kwart (28%) verwijst naar antwoordmodellen waarvan de interpretatie te veel tijd kost (om alles volledig te kunnen beoordelen). Een vijfde tot een kwart denkt dat een deel van de docenten de tweede correctie onvolledig uitvoert omdat een steekproef van kandidaten voldoende garantie biedt (25%) of omdat een steekproef van examenvragen voldoende garantie biedt (20%). De reden dat de noodzaak van volledige tweede correctie onvoldoende is aangetoond wordt door 7% van de respondenten gekozen. Omdat respondenten bij deze vraag meer alternatieven konden aankruisen, tellen de percentages niet op tot 100%.

Volledigheid tweede correctie en tijdsdruk

Hiervoor zagen we dat tijdgebrek een belangrijke oorzaak is waardoor docenten de tweede correctie niet altijd volledig (kunnen) uitvoeren. Aanvullend is de tweede correctoren gevraagd in hoeverre zij het eens zijn met de stelling 'Door tijdsdruk heb ik de tweede correctie minder

zorgvuldig kunnen uitvoeren dan wenselijk is'. Bijna driekwart (72%) is het er (helemaal) niet mee eens, 12% is het er (helemaal) mee eens en 15% kiest voor 'noch mee eens, noch mee oneens'.

Volledigheid van de bespreking van de verschillen tijdens het gezamenlijk overleg

Uit het vorige onderzoek naar de correctiepraktijk (Kuhlemeier & Kremers, 2013) is bekend dat er grote verschillen zijn in de duur van het overleg waarin de eerste en tweede corrector gezamenlijk de uiteindelijke scores bepalen. Dit roept de vraag op in hoeverre eerste en tweede correctoren erin slagen om ALLE geconstateerde scoreverschillen (plussen en minnen) te bespreken. In 2016 heeft 86% van de tweede correctoren alle scoreverschillen met de eerste corrector tijdens het gezamenlijk overleg besproken.

De relatief kleine groep van 450 docenten - 14% van de 3540 respondenten - die niet alle scoreverschillen met de eerste corrector besprak, is gevraagd hoe zij daarbij te werk gingen. Als we de antwoorden omrekenen naar percentages van alle docenten uit het onderzoek, dan ontstaat het volgende beeld. Van alle 3540 docenten kijkt 6% naar de totaalscores en bespreekt de verschillen alleen als zijn of haar totaalscore opvallend afwijkt van die van de eerste corrector, 5% kiest de vragen met de grootste scoreverschillen en bespreekt van deze vragen alle scoreverschillen, 3% kiest de kandidaten met de grootste scoreverschillen en bespreekt dan alle verschillen en 0% (na afronding op een geheel getal) vertrouwt erop dat de eerste corrector de scores van de niet besproken kandidaten en/of vragen overeenkomstig (de wel besproken gevallen) zal aanpassen. Hierbij merken we op dat de respondenten meer antwoorden konden aankruisen en de ene werkwijze de andere dus niet uitsluit.

Maatregelen ter bevordering van volledige tweede correctie

De docenten is gevraagd naar de wenselijkheid van potentiële maatregelen ter stimulering van volledige tweede correctie. Het opnemen van de uren in de jaartaak (61%) is de meest genoemde maatregel, gevolgd door taakverlichting tijdens de examenperiode (59%) en een betere financiële compensatie (46%). Ruim een derde (36%) vindt dat er sterker sturende antwoordmodellen ontwikkeld moeten worden met bijvoorbeeld meer voorbeelden van volledig juiste, gedeeltelijk juiste en foute antwoorden. Bij docenten Nederlands, economie en geschiedenis leeft deze wens overigens sterker dan bij docenten scheikunde, wiskunde en muziek. Relatief weinig docenten verwijzen naar een betere controle op de uitvoering van de tweede correctie (19%) of een betere voorlichting over het belang van tweede correctie (10%). Omdat de respondenten meer maatregelen konden aankruisen, tellen de percentages niet op tot 100%.

Toelichtingen van docenten bij de vraag naar verbeteringsmaatregelen

De toelichtingen van de docenten bevatten een schat aan suggesties ter bevordering van de volledigheid van de tweede correctie die nadere bestudering meer dan waard zijn. Een derde van de docenten doet suggesties ter verbetering van het examen. Enkele van de vele suggesties zijn: maak betere en minder subjectieve vragen, neem minder open vragen in het examen op, houd in het antwoordmodel meer rekening met andere goede antwoorden van kandidaten en voeg het officiële antwoordmodel van CvTE samen met de aanvullende adviezen voor de correctie van de vakinhoudelijke verenigingen. Bijna een kwart (24%) ziet de oplossing voor de onvolledige tweede correctie in een verbetering van de bestaande correctieprocedure. Het meest genoemd zijn omkering van de correctievolgorde en geheel onafhankelijke eerste en tweede correctie. Dertien procent noemt een betere facilitering door de school als oplossing voor de onvolledige tweede correctie. Daarentegen vindt 13% dat een oplossing (voorlopig) niet nodig is of dat er eerst onderzoek gedaan moet worden om te bepalen of volledige tweede correctie echt noodzakelijk of effectief is.

Kwaliteit van de correctie

Examinatoren moeten het werk van de kandidaten niet alleen volledig nakijken. Van even groot belang is dat zij de antwoorden van de kandidaten zorgvuldig en objectief nakijken. Onderzocht is wat eerste en tweede correctoren van elkaars correctiewerk vinden.

Zorgvuldigheid van de correctie

In de vragenlijst is de eerste corrector gevraagd naar de zorgvuldigheid waarmee de tweede corrector heeft nagekeken. Tegelijkertijd is de tweede corrector gevraagd naar de zorgvuldigheid van de eerste corrector.

Zorgvuldigheid van de tweede correctie (volgens de eerste correctoren)

Volgens de eerste correctoren heeft 87% van de tweede correctoren (zeer) zorgvuldig nagekeken, 7% min of meer zorgvuldig, 1% (zeer) slordig en 4% kon niet overzien hoe zorgvuldig de tweede corrector heeft nagekeken.

In 2016 vinden naar verhouding meer eerste correctoren dat de tweede correctie zorgvuldig is uitgevoerd dan in 2011. Vond in 2011 nog slechts een kwart (24%) van de eerste correctoren de tweede correctie zeer zorgvuldig, in 2016 was dit percentage gestegen naar ruim een derde (36%).

Zorgvuldigheid van de eerste correctie (volgens de tweede correctoren)

In 2016 heeft 67% van de eerste correctoren het werk (zeer) zorgvuldig nagekeken, 21% min of meer zorgvuldig en 12% (zeer) slordig. In 2016 hebben eerste correctoren volgens de tweede correctoren even zorgvuldigheid nagekeken als in 2011.

Verskil tussen de zorgvuldigheid van de eerste en tweede correctie

Tweede correctoren zijn kritischer over de zorgvuldigheid van de eerste correctie dan eerste correctoren over de zorgvuldigheid van de tweede correctie. Zo heeft in 2016 volgens de tweede correctoren 37% van de eerste correctoren zeer zorgvuldig nagekeken, terwijl slechts 14% van de tweede correctoren volgens de eerste correctoren zeer zorgvuldig gecorrigeerd heeft.

Objectiviteit van de correctie

Behalve volledig en zorgvuldig moet de correctie objectief zijn. Eerste en tweede correctoren mogen kandidaten bijvoorbeeld niet bevoornden of benadelen door een te soepele of te strenge beoordeling. De eerste correctoren is gevraagd naar de soepelheid van de tweede correctoren en de tweede correctoren is gevraagd naar de soepelheid van de eerste correctoren. Daarbij konden zij kiezen uit de antwoordmogelijkheden 'Te soepel (d.w.z. kandidaten zijn bevoornden)', 'Soepel', 'Noch te soepel noch te streng', 'Streng' en 'Te streng (d.w.z. kandidaten zijn benadeeld)'.
Te soepel (d.w.z. kandidaten zijn bevoornden)
Soepel
Noch te soepel noch te streng
Streng
Te streng (d.w.z. kandidaten zijn benadeeld)

Soepelheid van de tweede correctie (volgens de eerste corrector)

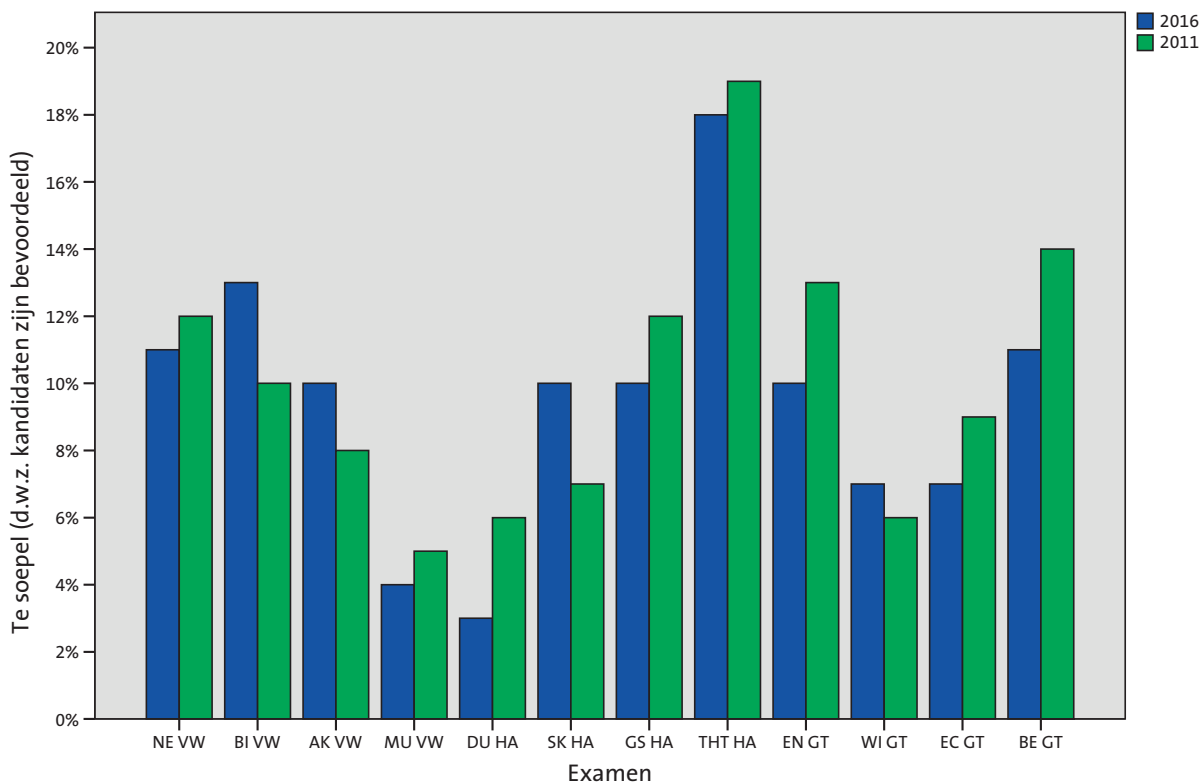
In 2016 heeft - volgens de eerste correctoren - 28% van de tweede correctoren (te) streng beoornden, 2% (te) soepel, 67% noch te soepel noch te streng en 3% kon niet overzien hoe soepel of streng de tweede corrector heeft nagekeken. In 2016 vonden eerste correctoren de tweede corrector even soepel of streng als in 2011.

Soepelheid van de eerste correctie (volgens de tweede corrector)

Volgens de tweede correctoren heeft 9% van de eerste correctoren in 2016 te soepel beoornden in de zin dat kandidaten bevoornden zijn, 17% beoorndde soepel, bijna twee derde (65%) was noch te soepel noch te streng, 7% was streng en 1% was te streng in de zin dat kandidaten

benadeeld zijn. Het percentage van 9% te soepel beoordelende eerste correctoren is ook gevonden in een recente enquête van de Algemene Onderwijsbond (AOB, 2015).

In 2016 vonden tweede correctoren de eerste correctie even soepel of streng als in 2011. In figuur 4.3 is voor elk examen en jaar het percentage eerste correctoren weergegeven dat volgens de tweede correctoren te soepel nakijkt in de zin dat kandidaten bevoordeeld zijn. De figuur maakt duidelijk dat de bevoordeling van kandidaten door de 'eigen' docent van examens tot examens verschilt (bij tehatex, beeldende vakken, geschiedenis en Nederlands speelt dit bijvoorbeeld meer dan bij muziek en Duits). De figuur laat ook zien dat de verschillen tussen de beide jaren duidelijk kleiner zijn dan die tussen examens.



Figuur 4.3 Percentage eerste correctoren dat te soepel beoordeelde in de zin dat kandidaten bevoordeeld zijn (bron: tweede corrector)

Verschil tussen de soepelheid van de eerste en tweede correctie

Hiervoor constateerden wij dat tweede correctoren minder positief denken over de zorgvuldigheid van de eerste correctie dan eerste correctoren over de zorgvuldigheid van de tweede correctie. Met betrekking tot de soepelheid van de beoordeling vinden we een vergelijkbaar patroon. Tweede correctoren vinden de eerste correctoren vaker (te) soepel dan dat eerste correctoren de tweede corrector (te) soepel vinden. In 2016 is bijvoorbeeld 9% van de tweede correctoren van mening dat de eerste corrector te soepel was, terwijl 0% van de eerste correctoren vindt dat de tweede corrector te soepel was. In datzelfde jaar vinden 21% en 7% van de eerste correctoren de tweede correctie streng respectievelijk te streng, terwijl de overeenkomstige percentages voor de mening van de tweede corrector over de strengheid van de eerste correctie slechts 7% en 1% zijn.

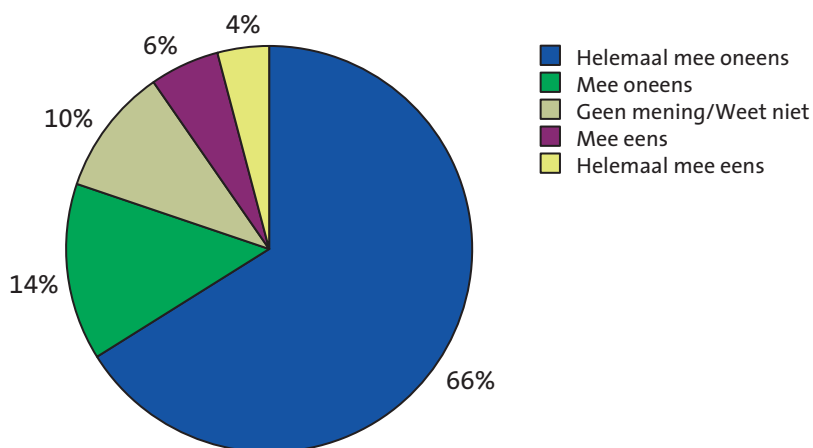
De gegevens over de objectiviteit van de correctie komen overeen met die uit het flankerend onderzoek van de drie vakinhoudelijke verenigingen (zie bijlage 1).

Maatregelen ter verbetering van de correctie

In 2016 zijn de docenten verschillende stellingen over de wenselijkheid van (vermeende) verbeteringsmaatregelen voorgelegd. Twee stellingen betreffen majeure wijzigingen in de huidige correctiesystematiek:

- Het is wenselijk dat de correctievolgorde wordt omgekeerd (d.w.z.: eerst de docent van de andere school en daarna de 'eigen' docent);
- Het is wenselijk dat de eerste en tweede correctie anoniem, door twee onafhankelijke correctoren worden uitgevoerd, dus niet meer door de 'eigen' docent.

Docenten zijn in overgrote meerderheid tegen omkering van de correctievolgorde: 80% is het er (helemaal) mee oneens, 10% is het er (helemaal) mee eens en 10% heeft geen mening of weet het niet (zie figuur 4.4).



Figuur 4.4 Mening van docenten over de wenselijkheid van omkering van de correctievolgorde

Onafhankelijke eerste en tweede correctie kent weinig voorstanders: 77% is het er (helemaal) mee oneens, 12% is het er (helemaal) mee eens en 10% heeft geen mening of weet het niet.

De massale weerstand tegen omkering is ook gevonden in een recente enquête van de AOb (2015).

Het gebruik van de adviezen van de vakinhoudelijke verenigingen voor het corrigeren

Twee andere stellingen gaan over de aanvullende adviezen voor het corrigeren die de vakinhoudelijke verenigingen tijdens de correctieperiode publiceren:

- Het is wenselijk dat de vakinhoudelijke verenigingen adviezen voor het corrigeren ter beschikking stellen;
- Het is wenselijk dat deze adviezen uiterlijk twee werkdagen na de landelijke bespreking gepubliceerd worden.

De aanvullende adviezen voor het corrigeren van de vakinhoudelijke verenigingen kunnen op brede ondersteuning van docenten rekenen: 84% is het er (helemaal) mee eens, 7% (helemaal) mee oneens en 8% heeft geen mening of weet het niet.

De overgrote meerderheid van de docenten zou het op prijs stellen als de adviezen voor de correctie uiterlijk twee werkdagen na de landelijke bespreking gepubliceerd worden: 83% kiest voor (helemaal) mee eens, 7% voor (helemaal) mee oneens en 10% voor heeft geen mening of weet het niet.

Het maken van afspraken over de aanlevering van het examenwerk

Er zijn signalen dat tweede correctoren het door de eerste corrector nagekeken werk soms zeer laat ontvangen. Het maken van een afspraak met de eerste corrector over de aanlevering van het examenwerk kan voorkomen dat de tweede corrector in tijdnood komt. De tweede correctoren is gevraagd of zij met de eerste corrector een afspraak maakten over de datum van aanlevering en hoeveel dagen voor het einde van de correctieperiode zij de examenwerken ontvingen.

Van de tweede correctoren heeft 8% met de eerste corrector een afspraak over de aanlevering van het examenwerk gemaakt. De tweede correctoren ontvangen het examenwerk gemiddeld zeven dagen voor het einde van de correctieperiode. De verschillen tussen docenten zijn groot. De 10% die de envelop met het examenwerk het laatst ontving, heeft het werk drie dagen of nog dichter tegen het einde van de correctieperiode ontvangen. De 10% die het werk het vroegst ontving, kreeg dat twaalf dagen of nog eerder voor het einde van de correctieperiode.

4.2 Discussie en aanbevelingen

De voorgaande paragraaf bevatte een cijfermatige beschrijving van de belangrijkste onderzoeksresultaten. In deze paragraaf doen we een eerste poging tot interpretatie en geven we alvast enkele aanbevelingen. Daarbij besteden we speciale aandacht aan enkele ogenschijnlijke tegenstrijdigheden in de onderzoeksresultaten die het lastig maken een eenduidig antwoord te geven op de vraag of de tweede correctie nu vollediger wordt uitgevoerd dan in 2011.

Tijdbesteding en belasting

Sinds het vorige onderzoek naar de correctiepraktijk is er onder andere in de media veel aandacht geschonken aan het belang van een volledige tweede correctie. Vanuit de veronderstelling dat een volledige correctie meer tijd kost dan een onvolledige correctie verwachtten we dat docenten in 2016 meer tijd aan de tweede correctie zouden besteden dan in 2011. Dit bleek evenwel niet het geval. Anders dan verwacht, vergden de eerste en tweede correctie in 2016 evenveel tijd als in 2011. In overeenstemming hiermee is de bevinding dat docenten de (eerste en) tweede correctie in 2016 even belastend vonden als in 2011. Gegeven de gelijk gebleven tijdbesteding en ervaren belasting lijkt het niet erg waarschijnlijk dat de tweede correctie in 2016 meer volledig is uitgevoerd dan in 2011.

In 2016 was de doorlooptijd - het aantal dagen waarin docenten de correctie uitvoerden - voor de eerste en tweede correctie tezamen ruim een halve dag langer dan in 2011. De wat langere correctieperiode zou mede een gevolg kunnen zijn van de 'Adviestabel opsturen examenwerk naar de eerste en tweede corrector' zoals gepubliceerd door de VO-Raad, FvOV, CNV onderwijs en AOb. Mogelijk heeft het gebruik ervan geleid tot een efficiëntere benutting van de beperkte periode die voor de correctie beschikbaar is.

Net als in 2011 investeerden de docenten in 2016 anderhalf tot bijna twee keer zoveel tijd in de eerste correctie als in de tweede correctie. Een eerste verklaring is dat de tweede correctie sneller kan verlopen omdat de tweede corrector het werk van de 'eigen' kandidaten al eerder als

eerste corrector heeft nagekeken. Een tweede verklaring verwijst ernaar dat de eerste correctie altijd volledig plaatsvindt en de tweede correctie niet altijd.

Bij de interpretatie wordt opgemerkt dat alleen gevraagd is naar de tijdbesteding aan het nakijken als zodanig. Docenten besteden daarnaast nog tijd aan andere zaken die met de correctie samenhangen, zoals intercollegiaal overleg over de toepassing van het correctievoorschrift, raadplegen van de examenverslagen van de vakinhoudelijke verenigingen en de administratieve afhandeling. De hier gerapporteerde correctietijd vormt dan ook een onderschatting van de totale tijdbesteding aan de eerste en tweede correctie.

Facilitering van de tweede correctie en tevredenheid

Sinds het vorige onderzoek naar de correctiepraktijk heeft het belang van een volledige tweede correctie en een betere facilitering ervan door de school veel aandacht gekregen. Nog vrij recent hebben de betrokken partijen diverse acties ondernomen ter bevordering van een volledige tweede correctie. Zo vroeg de VO-raad scholen de organisatie van de correctie opnieuw te bezien zodat docenten optimaal in de gelegenheid worden gesteld om de correctie volledig uit te voeren. De bonden riepen hun leden op ervoor te zorgen dat de uitvoering van de correctie aan de vereiste kwaliteit voldoet. De verwachting was dan ook dat scholen de tweede correctie in 2016 beter zouden faciliteren dan in 2011. Anders dan verwacht blijkt de facilitering van de **tweede correctie** nagenoeg ongewijzigd te zijn. De enige verandering is dat docenten in 2016 wat minder vaak een bedrag per examenwerk ontvangen dan in 2011. In overeenstemming hiermee is de bevinding dat tegenwoordig wat meer docenten vinden dat de tweede correctie op hun school niet gefaciliteerd wordt dan destijds. Net als bij de eerste correctie zijn docenten nu even (on)tevreden met de facilitering van de tweede correctie als toen. Al met al lijkt het schoolbeleid voor de tweede correctie te weinig veranderd om van een verbetering te kunnen spreken. Als de tweede correctie in 2016 al vollediger uitgevoerd zou zijn dan in 2011, dan komt dat waarschijnlijk niet doordat scholen hun docenten nu beter in staat stellen om de tweede correctie volledig uit te voeren dan destijds.

In het schoolbeleid ter facilitering van de **eerste correctie** lijkt wel een kleine verbetering te zijn opgetreden. In 2016 zijn de uren voor de eerste correctie wat vaker opgenomen in de jaartaak dan in 2011. Mogelijk als een gevolg hiervan krijgen docenten voor de eerste correctie wat minder vaak een vast bedrag per examenwerk. De kleine verbetering lijkt door docenten ook herkend te worden. Het percentage docenten dat vindt dat hun school geen schoolbeleid voor de eerste correctie heeft, is namelijk iets afgenomen. De kleine verbetering van het schoolbeleid voor de eerste correctie uit zich niet in een toegenomen tevredenheid. In 2016 zijn docenten hier even (on)tevreden over als in 2011.

Volledigheid van de tweede correctie

Volledigheid van de tweede correctie in 2016

Evenals in 2011 komt het in 2016 nauwelijks voor dat de tweede correctie helemaal niet wordt uitgevoerd. Slechts van minder dan 1% van de kandidaten lijkt het werk niet door een tweede corrector nagekeken. De anekdote van de eerste corrector die een aantal pagina's dichtplakte met lijm en het examenwerk enkele dagen later in dezelfde staat weer terugkreeg, lijkt dus vooral een Indianenverhaal. De docenten met een examenklas die niet bij de tweede correctie betrokken waren, is gevraagd waarom zij geen tweede corrector waren. De toelichtingen maken aannemelijk dat de tweede correctie vrijwel nooit achterwege blijft.

In het onderzoek zijn eerste en tweede correctoren ook direct bevraagd naar de volledigheid van de tweede correctie. De volledigheid van de tweede correctie is een uitermate complex

onderwerp. Docenten denken verschillend over wat in de praktijk onder volledigheid moet worden volstaan. Vandaar dat we de docenten meerdere vragen over dit onderwerp hebben voorgelegd die we vervolgens met elkaar in verband gebracht hebben. Op basis hiervan zijn verschillende indicaties verkregen van de mate waarin docenten tegenwoordig volledig nakijken:

- Allereerst is nagegaan hoe docenten denken over de volledigheid van de tweede correctie. In 2016 heeft 95% van de tweede correctoren naar eigen zeggen al het examenwerk (helemaal) volledig nagekeken. Op de stelling 'Ik heb het examenwerk van alle kandidaten volledig nagekeken' antwoordden zij met 'Klopt helemaal' (72%) of 'Klopt' (23%). Een recente enquête van de Algemene onderwijsbond (AOB, 2015) en de flankerende enquête van de vakverenigingen (zie bijlage 1) komen tot vergelijkbare resultaten. Zo reageerde 77% van de 2753 bevroegde AOB-leden op de stelling 'Wanneer ik tweede correctie doe, kijk ik alles na' met 'eens', 11% met 'neutraal' en eveneens 11% met 'oneens'.
- Er zijn echter aanwijzingen om deze zelfrapportages met enige voorzichtigheid te interpreteren. Een substantieel deel van de docenten vindt namelijk dat zij (helemaal) volledig hebben nagekeken ook al hebben zij niet alle werken en/of alle vragen gecorrigeerd (bijvoorbeeld alleen de open vragen). Dat blijkt onder meer uit de toelichtingen van docenten (zie paragraaf 3.3.6). Van de tweede correctoren die bij de stellingvraag aangaven (helemaal) volledig na te kijken, heeft gezien hun toelichting 71% daadwerkelijk volledig nagekeken, 11% keek de open vragen na maar de meerkeuzevragen niet of slechts gedeeltelijk, 5% keek niet alle kandidaten en/of niet alle vragen na (en 3% gaf een toelichting die niet gecodeerd kon worden en 10% gaf een toelichting die niet met de vraagstelling in verband stond). Het gegeven dat 16% van hen toch een niet-volledige correctiewijze toepast, werpt een ander licht op de eerste stellingvraag. Kennelijk vindt een substantieel deel van de docenten dat zij (helemaal) volledig nakijken terwijl zij dat in werkelijkheid niet doen.
- Meer precieze informatie over het percentage docenten dat daadwerkelijk alle examenwerken en alle examenvragen heeft nagekeken is verkregen door de antwoorden op vier vragen over de volledigheid van de tweede correctie met elkaar in verband te brengen (zie paragraaf 3.3.8). Op basis van de gegevens van 2016 is een conservatieve en een optimistische schatting verkregen van het percentage volledig nakijkende docenten. Volgens een conservatieve berekening heeft 66% de tweede correctie volledig uitgevoerd en volgens een optimistische schatting is dat 81%. Zie voor de precieze berekeningswijze paragraaf 3.3.8. De hoofdconclusie is dat nog niet alle docenten de tweede correctie in 2016 volgens eigen opgave volledig uitvoeren.

Volledigheid van de tweede correctie in 2016 en 2011

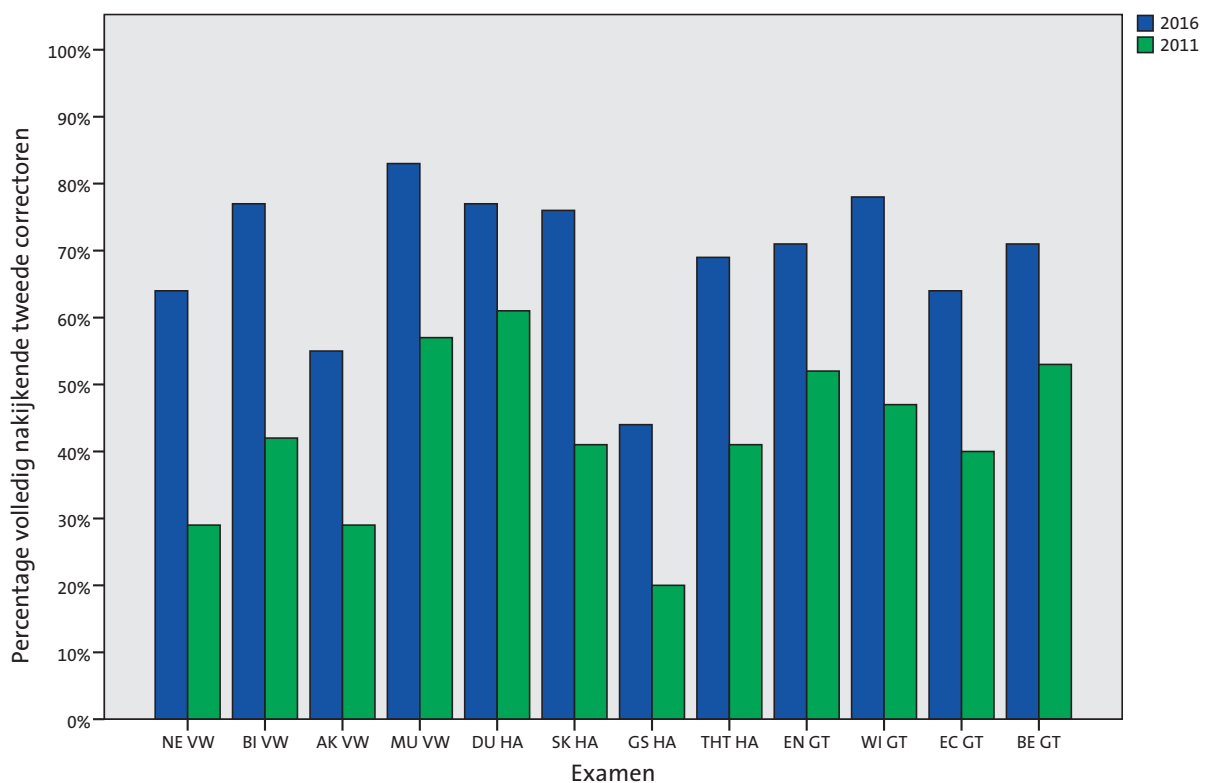
Een relevante vraag is in hoeverre de tweede correctie anno 2016 meer volledig wordt uitgevoerd dan in 2011. Drie vragen over de volledigheid van de tweede correctie zijn zowel in 2016 als 2011 aan de tweede correctoren gesteld, zodat we de uitkomsten kunnen vergelijken. Het betreft de vraag naar de volledigheid van de tweede correctie, het percentage gecorrigeerde kandidaten en het percentage nagekeken vragen. Op basis van alleen de eerste vraag heeft 37% van de tweede correctoren in 2011 volgens eigen opgave alles helemaal volledig nagekeken en in 2016 was dit percentage gestegen naar 72%. Meer precieze informatie is verkregen door de antwoorden op drie gemeenschappelijke vragen over de volledigheid van de tweede correctie met elkaar in verband te brengen. Uitgerekend is hoeveel tweede correctoren de tweede correctie naar eigen zeggen helemaal volledig uitvoeren, alle kandidaten corrigeren (voor een of meer vragen) en alle vragen nakijken (voor een of meer kandidaten). Van alle tweede correctoren voldoet in 2016 69% aan dit driedelige criterium tegenover 42% in 2011.

Wellicht ten overvloede merken we op dat het percentage volledig corrigerende docenten hier waarschijnlijk overschat wordt. In het criterium zijn de antwoorden op de stelling over niet-

volledige tweede correctie immers niet meegenomen (omdat deze stelling alleen in 2016 aan respondenten is voorgelegd). Toch kunnen we concluderen dat docenten van mening zijn dat de tweede correctie nu vollediger wordt uitgevoerd dan destijds. Dit blijkt ook uit de antwoorden van de eerste correctoren op de vraag naar de volledigheid van de tweede correctie. Vond in 2011 nog slechts 31% van de eerste correctoren dat de tweede corrector zeer volledig had nagekeken (dat wil zeggen alle werken en per kandidaat alle vragen), in 2016 was dit percentage gestegen naar 44%.

Verschillen tussen examenvakken in de volledigheid van de tweede correctie

De verschillen tussen examens zijn groot, zowel in 2016 als 2011. Figuur 4.5 toont het percentage integraal nakijkende tweede correctoren per jaar per examen (berekend op basis van het driedelige criterium; zie paragraaf 3.3.8). In 2016 kijken tweede correctoren biologie en wiskunde bijvoorbeeld veel vollediger na dan hun collega's geschiedenis, aardrijkskunde en tehatex. Het verschil tussen 2016 en 2011 verschilt sterk van examen tot examen. Bij Nederlands en geschiedenis is het percentage integraal nakijkende tweede correctoren sinds 2011 meer dan verdubbeld terwijl de toename bij Engels en Duits veel minder groot is.



Figuur 4.5 Percentage volledig nakijkende tweede correctoren per examen per jaar (berekend op basis van het driedelige criterium)

Ogenscheinlijk tegenstrijdige resultaten

De toename van het percentage docenten dat naar eigen zeggen volledig nakijkt, staat op gespannen voet met de resultaten van de eerder besproken vergelijking van de tijdbesteding, de ervaren belasting en de facilitering van de tweede correctie op school. Daaruit concludeerden we immers dat de tweede correctie anno 2016 niet veel vollediger uitgevoerd kan zijn dan in 2011. Immers, de tweede correctie kost in 2016 niet meer tijd dan in 2011, docenten ervaren de

tweede correctie als even belastend, scholen faciliteren de tweede correctie niet beter en eerste correctoren beoordelen volgens de tweede correctoren even vaak te soepel (in de zin dat kandidaten bevoordeeld zijn) als in 2011. Een verklaring voor deze ogenschijnlijke discrepantie is vooralsnog moeilijk te geven. Een van de vele mogelijke verklaringen is dat de tweede correctie ten tijde van het vorige onderzoek naar de correctiepraktijk nog niet in de schijnwerpers stond. Ook hing de dreiging van omkering van de correctievolgorde toen nog niet in de lucht. Mogelijk heeft dit docenten meer bewust gemaakt van de voordelen van de huidige correctieprocedure waardoor zij de volledigheid van de tweede correctie nu positiever beoordelen dan toen.

Alternatieve verklaringen voor de ogenschijnlijke discrepantie

Vertegenwoordigers van de betrokken belangenorganisaties is gevraagd mee te denken over alternatieve verklaringen voor de geconstateerde discrepantie. Een van hen wees op de mogelijkheid dat schoolleiders de correctielast tegenwoordig over meer docenten verdelen dan in 2011. Zo corrigeerde hij destijds het werk van twee klassen in zijn eentje terwijl hij nu slechts één klas deed en een collega de andere klas. Hierdoor zouden docenten in dezelfde tijd nu vollediger kunnen nakijken dan toen. Onze reactie hierop was als volgt. Per examen per school is net als in 2011 slechts één docent in ons onderzoek vertegenwoordigd. Als de schoolleiding de correctie voor een examen nu over meer docenten verdeelt dan toen, zou het gemiddeld aantal kandidaten dat een docent nakijkt moeten dalen (ervan uitgaande dat het totale aantal kandidaten ongeveer gelijk gebleven is). Het gemiddeld aantal kandidaten dat correctoren beoordelen, is echter gelijk gebleven (zie paragraaf 3.1.1 t/m 3.1.3). De aangedragen verklaring voor de ogenschijnlijke tegenstrijdigheid tussen de gelijk gebleven tijdbesteding en de toegenomen volledigheid van de tweede correctie lijkt dus minder aannemelijk.

Een andere mogelijke verklaring voor de geconstateerde tegenstrijdigheid in de resultaten verwijst naar veranderingen in het examen. Zo zouden de twaalf onderzochte examens tegenwoordig korter kunnen zijn of minder correctie-intensieve vragen kunnen bevatten dan in 2011. Dit zou docenten in staat kunnen stellen de tweede correctie in dezelfde tijd vollediger uit te voeren. De resultaten van een vergelijking van de examens van 2011 met die van 2015 zijn weergegeven in tabel 4.1. De tabel laat zien dat de examens anno 2016 gemiddeld niet alleen langer zijn maar ook meer correctie-intensieve open vragen bevatten dan in 2011. Ook deze verklaring voor de discrepantie tussen de gelijk gebleven tijdbesteding en de toegenomen volledigheid van de tweede correctie lijkt dus minder aannemelijk.

Tabel 4.1 Vergelijking van de twaalf onderzochte examens 2011-2016

	2016	2011
Aantal opgaven	440	434
Totaal aantal scorepunten	808	788
Aantal meerkeuzevragen	100	83
Aantal voor-gestructureerde vragen	17	40
Aantal open vragen	322	309
Aftrek (taalgebruik/samenvatting) bij Nederlands vwo	1	2

Oorzaken van onvolledige tweede correctie

Dat nog niet alle docenten de tweede correctie volledig uitvoeren, lijkt meer te maken hebben met niet kunnen dan met niet willen. Ruim vier van de vijf docenten is namelijk overtuigd van de noodzaak van volledige tweede correctie en bijna driekwart staat (sterk) afwijzend tegenover niet-volledige correctie. Dat onvolledigheid meestal geen kwestie van onwil is, blijkt ook uit een

analyse van de oorzaken van niet-volledige tweede correctie. Tijdgebrek en chronische overbelasting worden bijvoorbeeld veel vaker als oorzaken genoemd dan het argument dat een steekproef van kandidaten of examenvragen voldoende garantie biedt of dat de antwoordmodellen te veel ruimte voor verschillende interpretaties bieden. Bij de vraag naar mogelijke oplossingen blijkt het opnemen van de uren in de jaartaak het meest gekozen alternatief, gevolgd door taakverlichting tijdens de examenperiode en een betere financiële compensatie. Van sterker sturende antwoordmodellen, een betere controle op de uitvoering van de tweede correctie of een betere voorlichting over het belang van tweede correctie wordt veel minder heil verwacht. Overigens zijn sterker sturende antwoordmodellen meer een wens van docenten Nederlands, biologie, economie en geschiedenis dan van docenten muziek, tehatex, wiskunde, scheikunde en Engels.

Hoe onvolledig is onvolledig?

Hiervoor concludeerden we dat nog niet alle docenten de tweede correctie volledig uitvoeren en dat degenen die dat wel zeggen te doen toch niet allemaal alles helemaal nakijken. Dit impliceert natuurlijk geenszins dat deze docenten de tweede correctie schromelijk verwaarlozen. Het onderzoek heeft ook laten zien hoe docenten te werk gaan als zij niet alles helemaal nakijken. Dat kunnen we opmaken uit een analyse van de inhoud van de toelichtingen die 280 docenten (overeenkomend met 8% van de 3540 respondenten) gaven bij hun antwoord op de vraag naar de volledigheid van de tweede correctie. Ruim de helft (56%) van deze 280 docenten benadrukt in hun toelichting dat zij de tweede correctie volledig uitvoeren of dat zij eerst een deel van de kandidaten of vragen nakijken om vervolgens toch alles te corrigeren. Dertien procent keek de open vragen na maar sloeg (een deel van) de meerkeuzevragen over, 21% corrigeerde niet alle kandidaten en/of niet alle vragen, 2% gaf een toelichting die niet met het categorieënsysteem geassocieerd kon worden en 8% gaf een toelichting die geen relatie heeft met de vraagstelling. Voor een meer gedetailleerde beschrijving van de aangetroffen correctiestrategieën verwijzen wij de geïnteresseerde lezer naar paragraaf 3.3.6. Deze gegevens doen vermoeden dat een aanzienlijk deel van degenen die onvolledig nakijken de open vragen wel volledig nakijkt en alleen de meerkeuzevragen niet of slechts gedeeltelijk corrigeert.

Objectiviteit van de correctie

Behalve volledig en zorgvuldig moet de correctie ook objectief zijn. Eerste en tweede correctoren mogen kandidaten bijvoorbeeld niet bevoordelen of benadelen door een te soepele of te strenge beoordeling. Een meer volledige tweede correctie betekent niet automatisch dat de correctie objectiever wordt (Oberon, 2014). In een eerder onderzoek is informatie verzameld over de objectiviteit van de correctie (Kuhlemeier, Van Rijn & Kremers, 2011). In totaal dertig onafhankelijke correctoren keken het examenwerk van 803 examenkandidaten van 127 scholen opnieuw na. Daarbij werden de scores van de 'derde' correctoren vergeleken met die van de 'eigen' docent. Eerste correctoren bleken gemiddeld 6% hogere scores toe te kennen dan 'onafhankelijke' derde correctoren. Er bleken grote verschillen tussen scholen te bestaan in de soepelheid van de beoordeling. Op de ene school profiteren leerlingen daar veel meer van dan op de andere school. De conclusie was destijds dat kandidaten met eenzelfde vaardigheidsniveau op de ene school gemakkelijker een voldoende behalen dan op de andere school. Daarnaast bleken er bij geschiedenis, tehatex en Nederlands docenten te zijn die hun kandidaten onverklaarbaar toegeeflijk beoordelen zonder dat de tweede correctie dat corrigeert (Kuhlemeier, Gitsels, Boom & Van de Kerkhof, 2011). In het huidige onderzoek naar de correctiepraktijk is eerste correctoren gevraagd naar de soepelheid van de tweede corrector en tweede correctoren is gevraagd naar de soepelheid van de eerste corrector. De oordelen van de eerste en tweede correctoren over de soepelheid van elkaars beoordelingen lopen ver uiteen. Van de tweede correctoren vindt 9% dat de eerste corrector te soepel beoordeeld heeft in de zin

dat kandidaten bevoordeeld zijn. Daarentegen vindt vrijwel geen enkele eerste corrector (0%) de tweede corrector te soepel. Het omgekeerde patroon zien we bij de percentages strenge en te strenge beoordelingen. Slechts 8% van de tweede correctoren vindt de eerste corrector streng of te streng. Daarentegen is 28% van de eerste correctoren van mening dat de tweede corrector streng of te streng was. Een mogelijke verklaring veronderstelt dat docenten kritischer staan tegenover de objectiviteit van het werk van een onbekende collega dan tegenover hun eigen werk. Een andere verklaring stelt dat docenten als eerste corrector minder objectief beoordelen dan als tweede corrector. Als eerste corrector beoordelen docenten immers hun 'eigen' kandidaten terwijl zij met de kandidaten die zij als tweede corrector beoordelen geen persoonlijke band hebben.

In 2016 vinden tweede correctoren de eerste corrector even vaak (te) soepel als in 2011. Dit doet vermoeden dat eerste correctoren nu niet objectiever beoordelen dan destijds. Dit vermoeden is gebaseerd op (zelf)rapportages van docenten die ook hun 'eigen' kandidaten beoordeelden en niet op onafhankelijke beoordeling door derden. Het verdient daarom aanbeveling het onderzoek van 2009 waarbij onafhankelijke correctoren het examenwerk van een representatieve steekproef van examenkandidaten opnieuw beoordeelden (Kuhlemeier, van Rijn & Kremers, 2012) te herhalen. Door de resultaten met die van 2011 te vergelijken kan worden nagegaan in hoeverre de objectiviteit van de correctie verbeterd is. Het is immers van groot belang dat elke kandidaat het cijfer krijgt dat hij of zij verdient.

Maatregelen ter verbetering van de correctie

Docenten zijn in overgrote meerderheid tegen maatregelen waarbij de huidige correctie-systematiek ingrijpend gewijzigd wordt. Dat geldt zowel voor omkering van de correctievolgorde als onafhankelijke correctie waarbij de eigen docent er niet meer aan te pas komt. De massale weerstand tegen omkering van de correctievolgorde is ook gevonden in recente enquêtes van de Algemene Onderwijsbond (AOB, 2015) en de vakinhoudelijke verenigingen. Zo stemde 72% van de ondervraagden van de AOB-enquête tegen omkering van de correctievolgorde en wenste 74% de eerste correctie door de eigen docent te behouden.

De aanvullende adviezen voor het corrigeren die de vakinhoudelijke verenigingen tijdens de correctieperiode publiceren, kunnen op brede ondersteuning van docenten rekenen. De overgrote meerderheid van de docenten zou het op prijs stellen als de adviezen voor de correctie uiterlijk twee werkdagen na de landelijke bespreking gepubliceerd worden.

Gedifferentieerd schoolbeleid voor de correctie

Vastgesteld is dat de correctielast en de tijdbesteding aan de correctie voor het ene examen veel groter is dan voor het andere. Voor zover scholen de correctie al via expliciete maatregelen faciliteren, lijkt er sprake van een voor iedereen geldende aanpak. Te overwegen valt om scholen ertoe te bewegen het huidige 'one-size-fits-all'-beleid te vervangen door een gedifferentieerde aanpak afhankelijk van de omvang van de correctielast. Zo zouden docenten van correctie-intensieve examens met veel kandidaten zoals Nederlands en geschiedenis beter gefaciliteerd kunnen worden dan nu het geval is.

Omkering van de correctievolgorde een te zwaar middel?

In zijn brief van juni 2015 aan de Tweede Kamer gaf de staatssecretaris aan bereid te zijn om het besluit tot omkering van de correctievolgorde nog een jaar aan te houden. Eerst wilde hij bezien of de huidige werkwijze met de inzet van alle betrokken partijen toch goed ingevuld kon worden. Daarbij kondigde hij aan de correctie van de centrale examens in 2016 opnieuw door

Cito te laten onderzoeken en hierover te rapporteren. Als zou blijken dat de tweede correctie nu wel integraal wordt uitgevoerd zou hij afzien van omkering. Mocht het gewenste resultaat uitblijven, zou hij de eerder gemelde omkering alsnog in procedure brengen. Uit ons onderzoek komt naar voren dat nog niet alle docenten de tweede correctie (helemaal) volledig uitvoeren. Afhankelijk van het gehanteerde criterium variëren de schattingen van 5% tot 34% (zie paragraaf 3.3.8). Uit de toelichtingen bij de vragen kunnen we ook opmaken hoe docenten te werk gaan als zij niet alles helemaal nakijken. Daarbij blijkt dat 21% niet alle kandidaten en/of niet alle vragen corrigeerde en 13% alleen de meerkeuzevragen niet of slechts gedeeltelijk nakeek. Dit roept de vraag op of omkering van de correctievolgorde niet een te zwaar middel is om deze docenten ertoe te bewegen de tweede correctie alsnog helemaal volledig uit te voeren. Hierbij merken we op dat volledige tweede correctie (en omkering van de correctievolgorde) geen doel op zich is, maar een middel om ervoor te zorgen dat de eerste correctie objectief wordt uitgevoerd (en alle kandidaten het cijfer krijgen dat zij verdienen). Hierbij doelen wij vooral op de relatief kleine groep van eerste correctoren die te soepel beoordeelt in de zin dat kandidaten bevoordeeld worden. De precieze omvang van deze kleine groep is vooralsnog moeilijk te bepalen. Volgens de tweede correctoren uit dit vragenlijstonderzoek en een AOb-enquête (2015) gaat het om ongeveer 9% van de eerste correctoren. Een vergelijkbaar percentage werd na omrekening ook gevonden in het Panelonderzoek Vierde Correctie bij de vakken geschiedenis en tehatex (Kuhlemeier, Gitsels, Boom & Van de Kerkhof, 2011). Hoe kunnen we er dan toe bijdragen dat deze kleine groep objectiever ofwel minder toegeeflijk gaat nakijken? Het is de vraag of omkering van de correctievolgorde daarvoor een effectief en proportioneel middel is. De omkering brengt namelijk ook ongewenste neveneffecten met zich mee (AOb, 2015). Er zijn minder ingrijpende en waarschijnlijk effectievere middelen voorhanden die de huidige correctiesystematiek intact laten. Hierbij denken wij vooral aan bemiddeling door vakinhoudelijk onderlegde deskundigen op het moment dat een tweede corrector constateert dat de eerste corrector onverklaarbaar toegeeflijk beoordeeld heeft. Een tweede mogelijkheid is onaangekondigde steekproefsgewijze controle van examenwerk na afloop van de examenperiode die ongetwijfeld preventief zal werken.

Literatuur

Literatuur

AOb (2015). *Een motie van wantrouwen. Enquête omdraaien tweede correctie eindexamens.*

Krippendorff, K. (1980). *Content analysis: An introduction to its methodology.* Beverly Hills, CA: Sage Publications.

Krosnick, J.A. (1999). Survey research. *Annual Review of Psychology*, 50, 1, 537–567.

Kuhlemeier, H., Gitsels, H., Boom, S., & Kerkhof, A. van de (2012). *Examenkenmerken en verschillen tussen correctoren. Verslag van een panelonderzoek bij het CSE geschiedenis, tehatex en Nederlands.* Cito: Arnhem.

Kuhlemeier, H., Hemker, B., & Bergh, H. van den (2013). Impact of verbal scale labels on the elevation and spread of performance ratings. *Applied Measurement in Education*, 26, 1, 16-33.

Kuhlemeier, H. & Kremers, E. (2013a). *De praktijk van de eerste en tweede correctie. Verslag van een landelijke enquête.* Arnhem: Cito.

Kuhlemeier, H. & Kremers, E. (2013b). *De praktijk van de eerste en tweede correctie. Samenvatting van onderzoek naar het functioneren van het CSE.* Arnhem: Cito.

Kuhlemeier, H., Rijn, P. van & Kremers, E. (2012). *Eerste, tweede en derde correctie van examenwerk: Wat is het verschil?* Arnhem: Cito.

Loeffen, J., & Eikelboom, M. (in voorbereiding). *Antwoorden van docenten op de open vragen van het onderzoek naar de correctiepraktijk van het CSE. Verslag van een inhoudsanalyse (tentatieve titel).* Arnhem: Cito.

Oberon (2014). *Onderzoek naar de pilots tweede correctie centraal examen vo.* Utrecht: VO-raad.

Weber, R.P. (1990). *Basic content analysis* (Second edition). Newbury Park, CA: Sage Publications.

Bijlagen

Bijlage 1: Resultaten flankerende enquêtes NVvW, NVON en Levende Talen

Heleen van der Ree en Henry van Bergen (namens de vakverenigingen NVvW, NVON en LT)

Na de centrale eindexamens van mei 2016 hebben drie onderwijsvakverenigingen een enquête onder hun leden uitgezet die qua vragen grotendeels overeenkwam met de vragen die in het onderzoek van Cito werden gesteld. De enquête van de vakverenigingen is digitaal uitgevoerd.

De resultaten van deze enquêtes komen overeen met de resultaten van het door Cito uitgevoerde onderzoek. In deze bijlage zijn de resultaten van de enquête van de vakverenigingen opgenomen.

1a Aantallen respondenten

Tabel 1 Aantal respondenten en aantal leerlingen voor wie zij werk hebben nagekeken

	Aantal respondenten 1e correctie	Aantal kandidaten	Aantal respondenten 2e correctie	Aantal kandidaten
Nederlands vwo	91	3.121	69	2491
Frans vwo	106	2.192	80	1789
Engels vwo	6	271	6	335
Duits vwo	1	28	2	59
Biologie vwo	76	2.116	72	2.106
Natuurkunde vwo	175	5.263	159	4.967
Scheikunde vwo	180	5.557	158	5.438
Wiskunde A vwo	130	3.182	100	2.648
Wiskunde B vwo	140	3.380	129	3.480
Wiskunde C vwo	24	197	22	178
Nederlands havo	95	4.091	74	3297
Frans havo	51	1.090	46	1085
Engels havo	8	417	8	492
Duits havo	1	36	2	58
Biologie havo	90	2.818	82	1.129
Natuurkunde havo	121	3.125	105	1.829
Scheikunde havo	146	4.826	138	1.210
Wiskunde A havo	103	3.601	109	3.970
Wiskunde B havo	73	1.668	63	1.502
Nederlands gt	77	3.574	60	2882
Frans gt	25	385	15	294
Engels gt	9	392	7	292
Duits gt	3	90	3	135
Biologie gt	41	1.278	36	1.129
Natuurkunde gt	82	2.299	76	1.829
Scheikunde gt	41	1.400	46	1.210
Totaal	1.804	56.397	1.667	45.834

Gezien het kleine aantal respondenten bij sommige vakken, zullen deze niet mee genomen worden wanneer er naar aparte groepen gekeken wordt. Wanneer er naar de totale groep gekeken wordt, dan worden deze wel meegenomen.

1b Correctielast en tijdsbesteding

Tabel 2 Correctielast en tijdsbesteding aan de eerste en tweede correctie

	Gem aantal kandidaten	Gem tijd per kandidaat (in min)	Gem aantal kandidaten	Gem tijd per kandidaat (in min)
	1e correctie	1e correctie	2e correctie	2e correctie
Nederlands vwo	34,3	61	36,1	30
Frans vwo	21,1	12	22,4	8
Engels vwo	45,2	9	55,8	5
Biologie vwo	27,8	34	29,3	23
Natuurkunde vwo	30,1	46	31,2	27
Scheikunde vwo	31,0	48	34,4	26
Wiskunde A vwo	24,5	43	26,5	26
Wiskunde B vwo	24,1	50	27,0	27
Wiskunde C vwo	5,5	62	8,1	37
Nederlands havo	43,1	42	44,6	20
Frans havo	21,4	13	23,6	7
Engels havo	52,1	8	61,5	8
Biologie havo	31,3	28	31,2	16
Natuurkunde havo	25,8	43	27,6	24
Scheikunde havo	33,1	37	32,7	23
Wiskunde A havo	35,3	34	36,4	22
Wiskunde B havo	22,8	43	23,8	26
Nederlands gt	46,4	37	48,0	30
Frans gt	21,1	12	19,6	9
Engels gt	45,2	16	41,7	11
Biologie gt	31,2	22	31,4	16
Natuurkunde gt	28,0	27	24,1	22
Scheikunde gt	23,3	23	26,3	17

Wat opvalt is dat de belasting per vak sterk verschilt. Dit geldt zowel voor de 1e als de 2e correctie.

1c Ervaren bting eerste en tweede correctie

Tabel 3 Mate waarin docenten de eerste correctie belastend vinden (rijpercentages tellen op tot 100%)

	Niet belastend	Enigszins belastend	Belastend	Zeër belastend
Nederlands vwo	2	9	34	55
Frans vwo	63	30	7	0
Engels vwo	33	33	33	0
Biologie vwo	11	20	42	28
Natuurkunde vwo	1	19	42	38
Scheikunde vwo	4	26	47	23
Wiskunde A vwo	13	38	41	8
Wiskunde B vwo	9	29	42	20
Wiskunde C vwo	63	38	0	0
Nederlands havo	3	24	46	26
Frans havo	75	25	0	0
Engels havo	25	50	25	0
Biologie havo	10	44	37	9
Natuurkunde havo	5	36	42	17
Scheikunde havo	3	34	48	15
Wiskunde A havo	12	42	40	7
Wiskunde B havo	15	38	37	10
Nederlands gt	1	25	49	25
Frans gt	56	36	8	0
Engels gt	22	67	0	11
Biologie gt	15	44	39	2
Natuurkunde gt	21	48	27	5
Scheikunde gt	33	48	15	3

Tabel 4 *Mate waarin docenten de tweede correctie belastend vinden (rijpercentages tellen op tot 100%)*

	Niet belastend	Enigszins belastend	Belastend	Zeer belastend
Nederlands vwo	1	23	42	33
Frans vwo	53	40	6	1
Engels vwo	33	33	33	0
Biologie vwo	0	28	44	28
Natuurkunde vwo	4	25	41	31
Scheikunde vwo	3	34	41	22
Wiskunde A vwo	6	40	33	21
Wiskunde B vwo	4	36	35	26
Wiskunde C vwo	41	45	14	0
Nederlands havo	4	34	32	30
Frans havo	52	41	4	2
Engels havo	38	25	25	13
Biologie havo	12	35	35	17
Natuurkunde havo	6	35	39	20
Scheikunde havo	3	37	39	21
Wiskunde A havo	5	37	42	17
Wiskunde B havo	5	43	25	27
Nederlands gt	13	22	38	27
Frans gt	27	47	27	0
Engels gt	43	43	0	14
Biologie gt	6	42	33	19
Natuurkunde gt	11	43	29	17
Scheikunde gt	17	35	33	15

1d Schoolbeleid en tevredenheid

De resultaten van deze vraag zijn minder goed te interpreteren dan in het Cito-onderzoek, omdat er in dit onderzoek meerdere docenten van 1 school in de steekproef kunnen zitten. Wanneer er sprake is van facilitering/compensering, dan is dit vaak het niet hoeven surveilleren bij CE's en het niet opvullen van de vrijgekomen uren.

Tabel 5 *Tevredenheid met de facilitering van de eerste correctie*

	Zeervrededen	Tevrededen	Noch tevrededen noch ontevrededen	Ontevrededen	Zeervrededen
Nederlands vwo	2	16	30	31	22
Frans vwo	8	35	42	12	2
Engels vwo	0	33	33	33	0
Biologie vwo	0	24	29	32	16
Natuurkunde vwo	4	20	39	26	10
Scheikunde vwo	3	21	43	18	15
Wiskunde A vwo	5	25	46	17	7
Wiskunde B vwo	3	19	54	18	6
Wiskunde C vwo	17	25	46	4	8
Nederlands havo	5	16	38	25	16
Frans havo	10	33	41	14	2
Engels havo	0	0	63	25	13
Biologie havo	2	20	48	17	13
Natuurkunde havo	4	18	46	22	9
Scheikunde havo	3	19	48	18	12
Wiskunde A havo	2	22	54	16	6
Wiskunde B havo	3	21	53	15	8
Nederlands gt	8	21	39	21	12
Frans gt	20	16	52	8	4
Engels gt	33	33	22	11	0
Biologie gt	7	27	32	22	12
Natuurkunde gt	2	38	43	13	4
Scheikunde gt	8	23	43	22	3

Tabel 6 Tevredenheid met de facilitering van de tweede correctie

	Zeervrededen	Tevrededen	Noch tevrededen noch ontevrededen	Ontevrededen	Zeervrededen
Nederlands vwo	0	13	30	35	22
Frans vwo	6	35	43	15	1
Engels vwo	0	33	17	50	0
Biologie vwo	0	20	28	27	25
Natuurkunde vwo	2	18	34	27	19
Scheikunde vwo	4	13	38	22	24
Wiskunde A vwo	4	15	41	22	17
Wiskunde B vwo	2	7	50	28	14
Wiskunde C vwo	5	27	50	9	9
Nederlands havo	1	18	30	36	15
Frans havo	13	26	41	13	7
Engels havo	0	14	43	29	14
Biologie havo	2	17	34	28	18
Natuurkunde havo	4	15	40	26	15
Scheikunde havo	4	17	36	28	16
Wiskunde A havo	1	15	42	25	18
Wiskunde B havo	3	3	56	24	14
Nederlands gt	3	22	33	28	13
Frans gt	20	27	27	27	0
Engels gt	43	43	0	14	0
Biologie gt	0	20	28	27	25
Natuurkunde gt	2	18	34	27	19
Scheikunde gt	4	13	38	22	24

Het valt op dat de ontevredenheid over de facilitering bij de tweede correctie sterker is.

1e Volledigheid van de tweede correctie

Tabel 7 Hoe volledig heeft de tweede corrector het werk van uw kandidaten naar uw mening nagekeken?

	Zeer volledig	Volledig	Min of meer volledig	Onvolledig	Zeer onvolledig	Ik kan niet overzien hoe volledig de tweede corrector heeft nagekeken
Nederlands vwo	53	24	14	1	2	5
Frans vwo	55	30	0	1	0	14
Engels vwo	33	0	17	17	0	33
Biologie vwo	66	23	4	0	0	7
Natuurkunde vwo	64	21	6	1	1	7
Scheikunde vwo	56	29	9	1	0	6
Wiskunde A vwo	53	26	5	0	3	13
Wiskunde B vwo	58	27	6	0	0	9
Wiskunde C vwo	52	22	4	0	0	22
Nederlands havo	48	24	17	1	1	8
Frans havo	66	20	2	0	0	12
Engels havo	50	38	0	0	0	13
Biologie havo	62	28	7	1	1	1
Natuurkunde havo	58	26	6	2	1	8
Scheikunde havo	60	27	9	1	1	3
Wiskunde A havo	58	24	5	1	1	12
Wiskunde B havo	51	26	4	4	4	10
Nederlands gt	48	32	9	3	3	5
Frans gt	44	36	8	0	4	8
Engels gt	44	33	22	0	0	0
Biologie gt	44	34	10	0	0	12
Natuurkunde gt	57	32	2	1	0	7
Scheikunde gt	55	35	3	2	0	5

De eerste correctoren hebben de indruk dat de tweede correctoren het werk volledig hebben nagekeken.

1f Zorgvuldigheid van de tweede correctie

Tabel 8 Hoe zorgvuldig of slordig heeft de tweede corrector het werk van uw kandidaten naar uw mening nagekeken?

	Ze er zorg vuldig	Zorg vuldig	Min of meer zorg vuldig	Slordig	Ze er slordig	Ik kan niet overzien hoe zorg vuldig de tweede corrector heeft nagekeken
Nederlands vwo	52	26	8	0	6	9
Frans vwo	53	26	1	2	0	18
Engels vwo	17	67	0	0	0	17
Biologie vwo	51	35	7	1	0	7
Natuurkunde vwo	53	35	6	1	1	5
Scheikunde vwo	45	43	5	0	0	7
Wiskunde A vwo	43	34	7	0	1	15
Wiskunde B vwo	46	41	4	0	0	8
Wiskunde C vwo	50	29	0	0	0	21
Nederlands havo	41	39	12	1	1	6
Frans havo	63	20	8	0	0	10
Engels havo	50	50	0	0	0	0
Biologie havo	53	40	4	1	0	1
Natuurkunde havo	43	41	4	2	1	8
Scheikunde havo	47	42	5	1	1	4
Wiskunde A havo	46	34	7	1	1	12
Wiskunde B havo	39	40	7	1	1	11
Nederlands gt	48	36	12	0	1	3
Frans gt	38	42	4	4	4	8
Engels gt	56	33	11	0	0	0
Biologie gt	41	37	10	2	0	10
Natuurkunde gt	49	37	5	0	0	9
Scheikunde gt	38	48	7	2	0	5

De eerste correctoren hebben de indruk dat de tweede correctoren het werk zorgvuldig hebben nagekeken.

1g Soepelheid of strengheid van de tweede correctie

Tabel 9 Hoe soepel of streng heeft de tweede corrector het werk van uw kandidaten naar uw mening nagekeken?

	Te soepel	Soepel	Noch te soepel, noch te streng	Streng	Te streng	Ik kan niet overzien hoe soepel of streng de tweede corrector heeft nagekeken
Nederlands vwo	0	2	55	24	10	9
Frans vwo	0	7	71	13	2	8
Engels vwo	0	17	50	33	0	0
Biologie vwo	1	5	48	28	13	4
Natuurkunde vwo	1	5	72	16	4	3
Scheikunde vwo	0	1	77	15	6	2
Wiskunde A vwo	0	4	66	16	3	12
Wiskunde B vwo	0	5	67	22	4	3
Wiskunde C vwo	0	4	79	0	0	17
Nederlands havo	0	5	61	20	9	5
Frans havo	0	6	60	12	6	16
Engels havo	0	0	75	13	0	13
Biologie havo	0	6	56	29	9	1
Natuurkunde havo	0	5	70	13	4	7
Scheikunde havo	1	3	64	23	6	3
Wiskunde A havo	0	7	74	7	2	10
Wiskunde B havo	3	7	61	17	1	11
Nederlands gt	4	5	65	19	4	3
Frans gt	0	8	64	16	0	12
Engels gt	0	22	44	22	11	0
Biologie gt	5	5	54	15	15	7
Natuurkunde gt	0	0	70	20	2	9
Scheikunde gt	0	8	66	17	7	2

1h Correct toepassen van correctievoorschrift door eerste corrector

Tabel 10 In hoeverre heeft de eerste corrector het correctievoorschrift naar uw mening correct toegepast?

	Volledig correct	Correct	Min of meer correct	Niet correct	Helemaal niet correct
Nederlands vwo	9	41	26	17	7
Frans vwo	33	43	18	5	3
Engels vwo	17	33	50	0	0
Biologie vwo	13	31	32	18	7
Natuurkunde vwo	13	36	31	17	3
Scheikunde vwo	10	54	25	8	2
Wiskunde A vwo	13	45	30	7	4
Wiskunde B vwo	12	43	27	15	4
Wiskunde C vwo	10	62	19	10	0
Nederlands havo	7	36	33	19	5
Frans havo	27	41	23	9	0
Engels havo	0	63	38	0	0
Biologie havo	15	37	32	12	5
Natuurkunde havo	16	50	27	7	1
Scheikunde havo	2	47	30	17	3
Wiskunde A havo	11	45	29	14	1
Wiskunde B havo	11	57	21	11	0
Nederlands gt	3	42	25	17	13
Frans gt	13	33	47	7	0
Engels gt	14	14	43	29	0
Biologie gt	6	37	34	14	9
Natuurkunde gt	9	46	26	16	3
Scheikunde gt	4	43	26	15	11

1i Zorgvuldigheid van de eerste correctie

Tabel 11 Hoe zorgvuldig of slordig heeft de eerste corrector het werk van zijn/haar kandidaten naar uw mening nagekeken?

	Zeer zorgvuldig	Zorgvuldig	Min of meer zorgvuldig	Slordig	Zeer slordig
Nederlands vwo	16	36	16	22	9
Frans vwo	39	40	14	6	1
Engels vwo	33	33	33	0	0
Biologie vwo	15	47	15	17	6
Natuurkunde vwo	23	39	19	16	4
Scheikunde vwo	22	46	16	11	4
Wiskunde A vwo	18	44	19	14	5
Wiskunde B vwo	20	37	29	11	2
Wiskunde C vwo	23	55	18	5	0
Nederlands havo	18	36	16	19	11
Frans havo	29	38	27	7	0
Engels havo	0	75	25	0	0
Biologie havo	27	41	22	6	4
Natuurkunde havo	24	49	13	10	5
Scheikunde havo	12	41	22	17	7
Wiskunde A havo	16	47	21	14	3
Wiskunde B havo	19	52	13	11	5
Nederlands gt	12	27	28	17	17
Frans gt	7	47	33	13	0
Engels gt	14	57	14	14	0
Biologie gt	11	31	37	14	6
Natuurkunde gt	20	42	22	12	4
Scheikunde gt	11	48	17	17	7

1j Soepelheid of strengheid van de eerste correctie

Tabel 12 Hoe soepel of streng heeft de eerste corrector het werk van zijn/haar kandidaten naar uw mening nagekeken?

	Te soepel	Soepel	Noch te soepel, noch te streng	Streng	Te streng
Nederlands vwo	16	18	51	13	1
Frans vwo	9	14	65	11	1
Engels vwo	17	17	50	17	0
Biologie vwo	31	21	39	10	0
Natuurkunde vwo	15	23	54	6	2
Scheikunde vwo	7	15	68	9	1
Wiskunde A vwo	15	12	61	10	2
Wiskunde B vwo	16	19	53	12	0
Wiskunde C vwo	14	0	86	0	0
Nederlands havo	18	23	45	12	3
Frans havo	5	14	70	9	2
Engels havo	0	0	100	0	0
Biologie havo	16	16	57	10	1
Natuurkunde havo	6	15	68	10	1
Scheikunde havo	15	26	51	7	1
Wiskunde A havo	9	18	56	13	4
Wiskunde B havo	14	21	48	17	0
Nederlands gt	22	23	42	10	3
Frans gt	20	20	47	13	0
Engels gt	57	14	29	0	0
Biologie gt	19	33	39	6	3
Natuurkunde gt	12	17	61	9	1
Scheikunde gt	17	17	59	7	0

De praktijk van de eerste en tweede correctie van het CSE | Technische rapportage van de landelijke enquête van 2016 in vergelijking met 2011

Cito

Amsterdamseweg 13
Postbus 1034
6801 MG Arnhem
T (026) 352 11 11
klantenservice@cito.nl
www.cito.nl

Fotografie: Ron Steemers

