

Voor sommige docenten wordt het de eerste keer, velen deden het al vaker, maar het blijft spannend om straks de centrale eindexamens te moeten nakijken. Het is niet alleen een indirecte test of je je leerlingen goed genoeg hebt voorbereid, maar ook en vooral moet je zorgen dat er resultaten uitkomen die zowel je vak alsook de kandidaten recht doen.

Dat nakijken is nog niet zo eenvoudig. De centrale eindexamens kunst en de examens in de beeldende vakken bestaan bijna helemaal uit open vragen met antwoorden die nogal eens 'de volgende strekking' moeten hebben. Dat komt doordat vaak naar inzichten wordt gevraagd of doordat interpretaties of verklaringen moeten worden gegeven.

Op de evaluatiebijeenkomsten voor docenten is er dan ook veel discussie over het al dan niet goed mogen rekenen van alternatieve antwoorden of nuances die niet in het correctievoorschrift (cv) zijn opgenomen. De meest geuite kritiek op deze examens is al jarenlang dat 'het antwoordmodel niet volledig' zou zijn.

Het is natuurlijk altijd mogelijk dat er goede antwoorden bestaan die de examenmakers over het hoofd hebben gezien. Maar veel alternatieven die in eerste instantie juist lijken, zijn dat niet en zijn daarom dan ook niet in het cv opgenomen. Als gevreesd wordt dat een bepaald fout antwoord massaal wordt gegeven, wordt soms expliciet vermeld dat voor dat antwoord géén scorepunt mag worden toegekend. Het is echter ondoenlijk om bij elke vraag alle voordehand liggende foute antwoorden op te nemen. De examenmakers beperken zich dus tot de juiste antwoorden met een enkele keer een 'nul-punten-antwoord'.

Toch zijn er nog allerlei nuances mogelijk. Kandidaten geven namelijk niet altijd letterlijk het antwoord dat in het cv staat. Dat kan bijvoorbeeld komen doordat het gewenste antwoord in het officiële vakjargon staat, zodat u als docent precies begrijpt wat wordt bedoeld. Uiteraard streven we ernaar om zo exact en helder mogelijk te formuleren, met de kandidaat en het niveau in gedachten. Maar toch zullen docenten altijd moeten wegen in hoeverre de antwoorden van hun leerlingen overeenkomen met het correctiemodel. En in hoeverre de kandidaat duidelijk maakt dat de kwestie goed begrepen is, ook al staat het er misschien wat onhandig.

Dat wegen is vooral voor beginnende docenten moeilijk. Niemand wil zijn leerlingen tekort doen, maar het vak moet goed getoetst worden en de ene kandidaat weet of begrijpt gewoon meer dan de andere. Wat is eerlijk? Hoe objectief kan een docent zijn?

Nu houdt de ene docent zich strenger en letterlijker aan het cv dan de ander en zo is het denkbaar dat een kandidaat voor dezelfde antwoorden bij de ene examinerator een hoger cijfer krijgt dan bij een ander. Het systeem van de tweede correctie moet dit helpen voorkomen, maar het is de vraag of daarmee alle problemen worden opgelost.

De 'eigen' docent moet nog steeds wegen en zorgen dat hij dat eerlijk doet. Doe je de onhandig formulerende kandidaat die het toch goed begrepen heeft tekort en ben je dus te streng? Of kijk je structureel (te) soepel na en ken je bij twijfelachtige antwoorden eerder wel dan niet een punt toe? En wat betekent het als je dat doet?

Het gevolg van (te) soepel nakijken is dat je in zekere zin de 'goede' kandidaten tekort doet. Als de zwakkeren meer punten krijgen dan ze verdienen, komen ze met hun cijfer dichterbij de buurt van degenen die het echt goed begrijpen. De spreiding wordt kleiner. De betere kandidaat kan immers niet méér dan de maximumscore voor een antwoord behalen en profiteert dus minder van de 'soepele' docent.

Een toets moet betrouwbaar, valide en van het juiste niveau zijn. Maar een toets moet ook zo goed mogelijk discrimineren tussen betere en zwakkere kandidaten. De 'soepele' docent discrimineert minder dan een 'strengere' docent.

Als bijvoorbeeld leerling A, B en C respectievelijk een 4, 6 en 8 waard zijn (gemiddeld 6, spreiding 2), komen zij bij een soepel nakijkende docent bij wijze van spreken uit op 5, 6.5 en 8. Het gemiddelde cijfer stijgt en de spreiding, de discriminatie wordt kleiner. Krijgt die 8 er niets bij? Waarschijnlijk niet of nauwelijks. De betere leerling geeft meestal niet zoveel twijfelachtige antwoorden en er valt dan ook minder te wikken en te wegen.

Stel dat ineens op grote schaal het havo-examen veel soepeler zou worden nagekeken. Dan zou het totale gemiddelde cijfer natuurlijk stijgen en het percentage onvoldoendes dalen.

Bij de landelijke normering wordt dan geconcludeerd dat het examen te gemakkelijk was in vergelijking met voorgaande jaren en met het zogenaamde referentie-examen. De norm (de N-term) wordt dus aangepast. Het is niet aannemelijk dat de hele havo-populatie ineens veel intelligenter is geworden. Dus de N-term zal dalen om de resultaten van het 'te gemakkelijke' examen alsnog te corrigeren. De cijfers van alle kandidaten pakken dan lager uit.

En wat betekent dat binnen de klas? Door de lagere N-term zal het (te hoge) cijfer van de leerlingen A en B dalen, maar het kan altijd nog hoger uitpakken dan wat ze in feite waard zijn. Erger is dat de goede kandidaat C zijn of haar verdiende 8 nu niet krijgt!

Het gevolg van te soepel nakijken is dus dat de beste leerlingen worden benadeeld. Dat zullen noch de soepelste, noch de strengste docenten op hun geweten willen hebben. Alles pleit er dus voor om 'precies goed' na te kijken in mei. Vakbekwaam wegend wat uw kandidaten gepresteerd hebben. Ik wens u daartoe de nodige wijsheid en natuurlijk veel succes in de komende examentijd.

Marieke Wensing

Voorjaar 2015

(toetsdeskundige bij Cito en betrokken bij de examens in de kunstvakken)