

Improving **music** performance **assessment**

□ Herjan Denissen en Peter Schoenmakers

In 1991 gaf Bill Wrigley met het Brisbane Symphony Orchestra een concert. Op het programma stond het door Wrigley gecomponeerde 'Glasshouse Mountains piano concerto' waarin hij zelf als solist optrad. In wezen was die uitvoering het begin voor zijn proefschrift *Improving music performance assessment*. In april 2012, vlak voor zijn vertrek terug naar Australië, spraken wij met hem in Maastricht. Uit dat gesprek bleek dat onderzoek naar het beoordelen van muziekuitvoeringen noodzakelijk is, maar dat er ook veel aandacht moet komen voor het implementeren van resultaten

Wat was de aanleiding voor uw onderzoek naar het beoordelen van muziekuitvoeringen?

Door zelf als musicus op te treden en beoordeeld te worden kreeg Wrigley belangstelling voor het objectief beoordelen van muziekuitvoeringen en voor de invloed van het wel of niet ervaren van *flow* op het resultaat. Over *flow* later meer.

Hij vertelt dat een aantal motieven aanleiding heeft gegeven tot zijn onderzoek.

"Ten eerste bemerkte ik in Australië een sterke bewustwording van en een groeiende vraag naar verantwoording afleggen (*accountability*), ook in het onderwijs. Dit kon volgens mij alleen maar als het assessment transparant en objectief zou zijn. Op het Queensland conservatorium van de Griffith University bestond het beoordelingsformulier van muziekuit-

voeringen uit slechts één A-viertje zonder criteria. Op dit formulier schreven docenten opmerkingen over een muziekuitvoering van een student. Dat was te summier.

Ten tweede wilde ik als conservatoriumstudent weten waar experts zich bij de eindbeoordeling op focussen en wilde ik ontdekken wat de karakteristieken zijn van een goede uitvoering. Ten slotte raakte ik door het onderzoek van de psychologen Csikszentmihalyi en Seligman naar *flow*, geïnteresseerd in de invloed van *flow* op het resultaat van een muziekuitvoering."

Bill Wrigley

Klinisch psycholoog en musicus William J. Wrigley PhD is tegenwoordig directeur bij Academic Training and Consulting International in Brisbane (Australië). Voordien was hij werkzaam als onderzoeker assessment binnen de faculteit Health, Medicine and Life Sciences van de Universiteit Maastricht. Zijn werk hield onder meer in: Implementation of assessment systems, strategies, program review, professional development training in teaching and assessment in medical education. Daarvoor werkte hij onder meer bij de Universiteit Groningen en de Griffith University in Brisbane. In 2006 is hij gepromoveerd bij de Faculty of Education van de Griffith University.

U streeft naar objectief beoordelen. Er zijn echter beoordelaars die van mening zijn dat het beoordelen van een muziekuitvoering altijd subjectief is. Hoe reageert u daarop?

"Bij het globale beoordelingsformulier op het Queensland conservatorium, waarop docenten alleen opmerkingen kunnen schrijven, zou je bij de beoordeling een grote verscheidenheid in commentaar verwachten, maar dat was niet het geval. Door analyse van 800 ingevulde globale beoordelingsformulieren muziek consta-

De parameters helpen de examinatoren om consistent in hun beoordeling te zijn



teerde ik dat zij bij hun beoordeling van muziekpraktijk deductief te werk gingen. Hoewel ze een globaal beoordelingsformulier gebruikten denken ze toch in een overzichtelijk aantal parameters, zoals tempo, ritme, articulatie en de specificatie ervan. Je kunt dus zeggen dat er een gezamenlijke beoordelingstaal is. Vanuit die constatering heb ik een Performing Evaluation Report (PER) ontwikkeld."

Hoe ziet een Performing Evaluation Report (PER) eruit?

"Ik ben natuurlijk alleen uitgegaan van wat de diverse experts vonden. Hun beoordelingstaal, de door hen genoemde muzikale parameters en specificaties heb ik in het beoordelingsmodel opgenomen. En omdat er tussen de experts van de verschillende instrumentengroepen verschil van mening was over op te nemen parameters, is er voor elke instrumentengroep een aparte PER ontwikkeld. Overigens zijn de verschillen tussen die PER's niet heel groot. Mijn doel was om examinatoren te helpen om met het PER betrouwbaarheid, validiteit en transparantie te verbeteren. Het nieuwe van dit model ten opzichte van andere beoordelingsmodellen schuilt vooral hierin dat er overeenstemming is tussen de examinatoren over de in het PER opgenomen parameters en specificaties (beoordelingscriteria). Dit helpt de examinatoren om consistent in hun beoordeling te zijn.

Een PER (zie figuur 1) bevat in de eerste kolom een aantal parameters met korte specificatie gegroepeerd, afhankelijk van instrumentgroep, in twee tot vier *dimensions*. De horizontale regels bevatten de drie niveaus: Needs Attention, Satisfactory en Excellent. Deze drie niveaus bevatten elk een specificatie, bijvoorbeeld bij Needs Attention: 'generally inadequate throughout' en 'limited throughout'. De beoordelaars geven geen cijfer, maar kruisen aan wat volgens hen bij de beoordeling van toepassing is. Het is dus geen cijfermatig model, maar een beschrijvend model, vooral ook geschikt voor feedback.

Wat zijn voordelen en nadelen van het werken met een PER?

"Naast een meer consistente beoordeling, is een voordeel dat voor iedereen (de student en de maatschappij) duidelijk is welke parameters bij de beoor-

Geen rubrics

Op de vraag waarom Wrigley geen rubrics gebruikt antwoordt hij:

"Bij gebruik van rubrics zou de PER gevuld zijn met uitgebreide beschrijvingen. Nu bevat het model de parameters en een specificatie van enkele woorden, bijvoorbeeld Articulation: clarity, freedom. Ik ben geen voorstander van rubrics in een beoordelingsmodel, want iedere uitgebreide beschrijving leidt tot discussie, en dat leidt tot bias. De beoordelaarsovereenstemming zal daardoor afnemen. Rubrics zijn goed om richting te geven. Mijn voorkeur gaat uit naar parameters met korte specificatie zoals in de PER."

De redactie vult hierbij het volgende aan. In Nederland zijn door de SLO voor het beoordelen van vaardigheden rubrics ontwikkeld. SLO schrijft daarover: 'Er is behoefte aan een instrument dat de kwaliteit van de vaardigheden meet, op een zodanige manier dat de uitkomsten een zo groot mogelijke mate van geldigheid hebben. SLO heeft de oplossing gezocht in een bestaand instrument: rubrics (ofwel rubrieken in gewoon Nederlands). Rubrics zijn ontwikkeld door Bernie Dodge [...] De invulling van rubrics is gebaseerd op genoemde definitie van een competentie, namelijk iemands vermogen om adequaat gedrag te vertonen in bepaalde situaties. Ze zijn om te beginnen uitgewerkt voor de algemene vaardigheden samenwerken, communiceren, plannen en organiseren, informatie verwerven en verwerken en presenteren.' (SLO, 2006)

deling gebruikt worden. Door de PER ook tijdens het leerproces te gebruiken wordt steeds duidelijk wat de sterke en zwakke punten van een muziekstudent zijn, waardoor verbetering mogelijk is. De ingevulde PER kan als discussiestuk gebruikt worden met de student zelf en de medestudenten.

Overigens sluit het model ook aan bij de beleving



*Als docent wil je voor de student
succeservaringen creëren*

Appendices

Appendix 4E Voice Performance Evaluation Report

VOICE Queensland Conservatorium Performance Evaluation Report							
Student:		Year Level:					
TECHNIQUE Desired Qualities	Needs Attention		Satisfactory			Excellent	
	generally inadequate throughout	limited throughout	inconsistent throughout	moderately consistent throughout	mostly consistent throughout	consistent command throughout	complete mastery throughout
ARTICULATION: clarity, freedom							
REGISTRATION: low, high, balanced, tessitura							
INTONATION: accuracy							
TONE/COLOUR: chiaro scuro, vibrancy, clarity							
AIR/BREATH: appoggio, energy							
TENSION: body alignment, ease							
FREEDOM							
INTERPRETATION Desired Qualities	generally inadequate throughout	limited throughout	inconsistent throughout	moderately consistent throughout	mostly consistent throughout	consistent command throughout	complete mastery throughout
TEXT: accuracy, clarity							
INTERPRETATION: expressive, insightful, tone colour							
CHARACTERIZATION: stage presentation, convincing							
INSIGHT: meaningful, imaginative							
MUSICALITY Desired Qualities	generally inadequate throughout	limited throughout	inconsistent throughout	moderately consistent throughout	mostly consistent throughout	consistent command throughout	complete mastery throughout
STYLE: insightful, musical integrity							
PHRASE: direction, shape							
DYNAMICS: contrast, choice							
RHYTHM & TEMPO: accuracy, choice							
COMMUNICATION Desired Qualities	generally inadequate throughout	limited throughout	inconsistent throughout	moderately consistent throughout	mostly consistent throughout	consistent command throughout	complete mastery throughout
COMMUNICATION: expressive, honest, committed, engaging							
POISE/CONFIDENCE							

Figuur 1. Performance Evaluation Report voor zang

van de examinatoren dat muziek beschrijvend en niet cijfermatig is. De beoordelaar zet immers alleen kruisjes in de diverse vakjes en krijgt op die manier overzicht over het resultaat.

Een nadeel is dat sommige examinatoren de PER blijven afwijzen omdat ze de uitvoering als geheel willen beoordelen en niet de onderdelen, de afzonderlijke parameters. Maar met mijn studie heb ik aangetoond dat docenten altijd in parameters denken, ook als ze alleen de uitvoering als geheel zouden willen beoordelen.

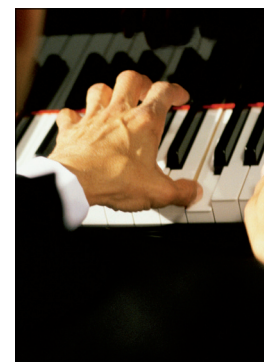
Een nadeel is wellicht ook dat de PER te uitgebreid is. Alhoewel, de praktijk leert dat de muziekexaminator heel snel de PER invult omdat de parameters de docent

zeer vertrouwd zijn. Enige verfijning zou echter wel goed zijn."

Uit uw studie blijkt dat er bij een beoordeling met een globaal beoordelingsformulier grotere beoordelaar-overeenstemming is dan bij gebruik van de PER. Hoe verklaart u dat?

"Uit sommige onderzoeken blijkt inderdaad dat het makkelijker is om overeenstemming te krijgen met een globale beoordeling. Echter met meer beoordelaarstraining en door vaak gebruik te maken van de PER zal de overeenstemming toenemen. Bovendien zijn de diagnostische en feedback doeleinden van de PER belangrijk."

Muziekstudenten zouden op meerdere momenten en in verschillende contexten beoordeeld moeten worden



Het tweede thema in uw onderzoek is het effect van flow op een uitvoering. Wat voor invloed heeft flow op de kwaliteit van een uitvoering?

"Flow ontstaat wanneer de student evenwicht ervaart tussen de eigen vaardigheden (*skills*) en de uitdaging (*challenge*). Vaak is een student nerveus. Spanning ontstaat wanneer de perceptie van de uitdaging hoger is dan die van de vaardigheden. De perceptie van de uitdaging en de perceptie van vaardigheden moeten matchen. Het model van Csikszentmihalyi laat dat zien (figuur 2).

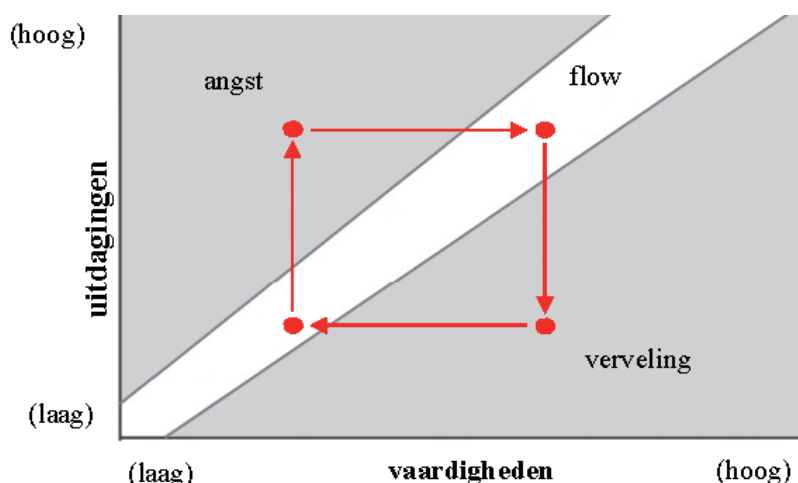
Je moet een student niet in een situatie brengen waarbij de uitdaging verder reikt dan de perceptie van het niveau van de vaardigheden. Als docent wil je voor de student succeservaringen creëren. Ik heb veel gewerkt met mensen die last hadden van spanning en heb altijd gekeken naar hun ervaringen met uitvoeringen. In de meeste gevallen was nervositeit sterk verbonden met het feit dat ze te weinig succesvolle uitvoeringen hadden gehad in met name de eerste optredens. Wat ik gedaan heb, is succesvolle uitvoeringen creëren door hen veel op te laten treden in situaties met een lage uitdaging, maar waar wel hoge vaardigheden vereist waren."

Het onderzoek is uitgevoerd op de Griffith University in Australië. Zijn de PER's ook bruikbaar op andere conservatoria, bijvoorbeeld hier in Maastricht?

"Ik denk dat de parameters in de PER's universeel zijn. De PER is echter een ecologisch model, bestemd voor de cultuur, de situatie, het curriculum op het Queensland conservatorium. Als deze in Maastricht gelijk zijn aan die in Brisbane dan kun je de PER ook in Maastricht gebruiken. Zo niet, dan moeten de betreffende examinatoren/docenten overeenstemming bereiken over de parameters op de betreffende school."

Hoe zou uw studie een vervolg kunnen krijgen?

"De afsluiting van een conservatoriumopleiding is meestal gebaseerd op één toetsmoment. Muziekstudenten zouden op meerdere momenten en in verschillende contexten beoordeeld moeten worden en niet alleen op dat ene moment, het eindconcert. Ik denk dat de betrouwbaarheid en validiteit zullen toenemen wan-



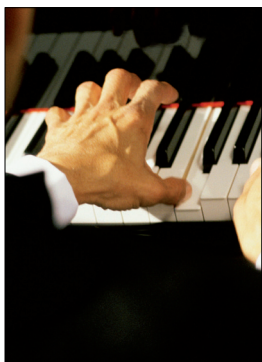
Figuur 2. Het flow model (Csikszentmihalyi, 1990)

Implementatie is niet eenvoudig

Op de vraag of de resultaten van zijn onderzoek op enigerlei wijze zijn geïmplementeerd op het Queensland conservatorium, antwoordt Wrigley:

"Zover ik weet niet. Dat heeft ten eerste te maken met de traditie van beoordelen. Docenten willen graag hun wijze van beoordelen behouden, oftewel: wat onbekend is, blijft onbemand. Ook moet het conservatorium de toetspraktijk willen veranderen. En de idee blijft bij sommigen dat je een muziekuitvoering niet kunt beoordelen op de wijze die ik voorstel. Het juiste klimaat was er in 2006 nog niet, vermoed ik."

Bij navraag op het Queensland conservatorium reageerde professor Huib Schippers, directeur van het Queensland Conservatorium Research Centre: "Ik herinner me Bills werk. Interessant, maar ik kan niet zeggen dat ik het helemaal eens ben met zijn ideeën. We zijn inmiddels een stuk verder met denken over assessment." Dr. Don Lebler, onderdirecteur Learning & Teaching van het conservatorium, laat weten dat het conservatorium naast allerlei vormen van toetsing zoals portfolioassessment en peer assessment voor onder meer popmuziek, nog steeds gebruikmaakt van een globaal beoordelingsformulier.



Studenten treden te weinig op

neer je met een portfoliosysteem met meerdere toetsmomenten en toetscontexten werkt. Studenten treden te weinig op. Met een portfoliosysteem dat aangeeft waaraan studenten moeten voldoen (bijvoorbeeld x-aantal optredens in verschillende situaties), zullen ze vaker optreden onder verschillende omstandigheden, zoals op een bruiloft, in een park of in de concertzaal: 'to match the real world.'"

Literatuur

Csikszentmihalyi, M. (1990). *Flow: The psychology of optimal experience*. New York, NY: Harper & Row). In

Nederlandse vertaling: Csikszentmihalyi, M. (2010). *Psychologie van de optimale ervaring*. 9e druk. Amsterdam: Boom.

Seligman, M.E.P. & Csikszentmihalyi, M. (2000). *Positive psychology: An introduction*. American Psychologist, 55(1), 5-14.

Wrigley, W.J. (2006). *Improving music performance assessment*. Faculty of Education, Griffith University, Brisbane.

SLO (2006). *Rubrics als beoordelingsinstrument voor vaardigheden*. Vmbo reeks. Enschede.

□ De heer drs. H.J.M. Denissen en de heer drs. P.A. Schoenmakers zijn toetsdeskundigen muziek bij Cito. E-mail: herjan.denissen@cito.nl.

Gesignaleerd

Extra tijd invoering referentieniveaus taal en rekenen 19 december 2012 - <http://www.mborraad.nl/?news/>

Scholen krijgen extra tijd voor de invoering van de referentieniveaus taal en rekenen. Dit maken minister Bussemaker en staatssecretaris Dekker bekend in de 'Voortgangsrapportage implementatie referentiekader taal en rekenen' die zij op 19 december naar de Tweede Kamer stuurden.

Hiermee komen ze tegemoet aan de zorgen die in het onderwijs leven over de consequenties van een te snelle invoering van de slaag-/zakregeling voor vo-leerlingen en mbo-studenten. Op 12 december verwoordden MBO Raad, AOC Raad, VO-raad, de NRTO, SPV en de JOB deze zorgen in een brief aan minister en staatssecretaris.

Belang sectoroverstijgende aanpak

De MBO Raad is blij met de extra tijd, maar vindt met name de afstemming in de maatregelen voor de verschillende onderwijssectoren belangrijk. Het verhogen van het taal- en rekenniveau van leerlingen en studenten kost tijd. Maatregelen die hieraan tegemoet komen moeten er niet toe leiden dat een leerling op het vo profiteert van een tijdelijke maatregel (uitstel van de slaag-/zakregeling) om hierop in het mbo alsnog te stranden. Het blijft daarom belangrijk de resultaten van leerlingen en studenten in de verschillende onderwijssectoren zorgvuldig te blijven volgen en de slaag-/zakregeling pas in te voeren als ook het instroomniveau op orde is. Het belang van een sectoroverstijgende aanpak zullen wij ook in de toekomst, samen met de andere sectorraden, blijven benadrukken.

Complexiteit in regelgeving

De regelgeving rond de taal en rekenen is erg complex. Het feit dat in de slaag-/zakregeling Nederlands en rekenen in het mbo losgekoppeld worden maakt deze complexiteit nog groter. De MBO Raad zal hier aandacht voor vragen bij het ministerie van Onderwijs.

Beschikbare middelen

Tot en met 2013 is er jaarlijks ruim € 51 miljoen beschikbaar voor de implementatie van de referentieniveaus. De minister is voornemens deze middelen vanaf 2014 onder te brengen in de prestatiebox.