



UvA-DARE (Digital Academic Repository)

Beoordelingsinstrumenten in de kunstvakken – een review

Groenendijk, T.; Damen, M.; Haanstra, F.; van Boxtel, C.

Published in:
Pedagogische Studiën

[Link to publication](#)

Citation for published version (APA):

Groenendijk, T., Damen, M., Haanstra, F., & van Boxtel, C. (2016). Beoordelingsinstrumenten in de kunstvakken – een review. *Pedagogische Studiën*, 93(2), 62-82.

General rights

It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Samenvatting

Dit paper presenteert resultaten van een reviewstudie naar beoordelingsinstrumenten en beoordelingscriteria in de kunstvakken (dans, beeldende kunst, muziek en theater) voor het basis- en voortgezet onderwijs. De onderzoeksvraag is: Welke beoordelingsinstrumenten voor de kunstvakken in het basis- en voortgezet onderwijs zijn in de literatuur tussen 2000 en 2013 beschreven en wat wordt daarin beoordeeld? De resultaten betreffen 74 instrumenten beschreven in 65 artikelen in 30 Engelstalige peer reviewed tijdschriften. Het meest voorkomende instrumentarium bevat een taak waarin leerlingen zich uiten in het artistieke medium en een scoringsinstrument in de vorm van criteria of een rubric. De instrumenten gebruiken een zeer breed spectrum aan beoordelingscriteria. De kunstdisciplines verwijzen ieder naar hun eigen vakspecifieke competenties, al zien we op een hoger abstractieniveau wel overeenkomsten. Technische vaardigheden en creativiteit zijn bijvoorbeeld competenties die bij alle disciplines – ieder op hun eigen manier - beoordeeld worden.

1 Inleiding

Volgens de Onderwijsraad en de Raad voor Cultuur (2012) ontbreekt het bij kunstvakken in het onderwijs (zoals beeldende vakken, muziek, dans en theater) aan opbrengstgericht werken. Er is volgens de raden vooral in het primair onderwijs behoefte aan leerlijnen waarin de gewenste ontwikkeling van kennis, vaardigheden en houding beschreven worden. Dit vraagt om een uitbreiding van het monitor- en beoordelingsinstrumentarium, waarmee docenten een beter inzicht krijgen in de culturele ontwikkeling van hun leerlingen. Het daarop door OCW geïnitieerde programma ‘Cultuureducatie met kwaliteit’ tracht de ontwikkeling van doorgaande leerlijnen en van een beoordelingsinstrumentarium te

stimuleren. Voor leerlijnen zijn landelijke leerplankaders ontwikkeld die als richtsnoer kunnen dienen, zoals het leerplankader kunstzinnige oriëntatie voor het primair onderwijs van de SLO (2014) en het leerplankader Cultuur in de Spiegel voor primair en voortgezet onderwijs (Van der Hoeven e.a., 2014). Voor het ontwikkelen van een beoordelingsinstrumentarium ontbreekt zo’n richtsnoer. In kunstvakken streeft men zelden naar uniforme, voorspelbare leerresultaten. Kunstzinnige processen en producten zijn verbonden met persoonlijke expressie en individuele kwaliteiten, terwijl beoordelingen objectiviteit en vergelijkbaarheid veronderstellen. Deze spanning heeft ertoe geleid dat de houding van docenten ten aanzien van beoordeling zelfs is omschreven als vijandelijk (Soep, 2004). Maar desondanks wordt er ook in de kunstvakken beoordeeld, bestaan er eindexamens voor deze vakken en in sommige landen ook nationale standaarden.

Er bestaat uiteraard ook literatuur op dit gebied. Herpin, Washington en Li (2012) verzamelden informatie over de huidige beoordelingspraktijken in kunstvakken in de Verenigde Staten en de wensen in het veld daaromtrent. Ze concludeerden dat er een tekort is aan openbaar toegankelijke, kwalitatief hoogwaardige beoordelingsinstrumenten. Hetzelfde geldt voor Nederland. Er is daarom behoefte aan een literatuuroverzicht van beoordelingsinstrumenten, waarin wordt aangegeven wat en hoe er beoordeeld wordt. De probleemstelling die we hier beantwoorden luidt dan ook: Wat en hoe wordt er beoordeeld in de beoordelingsinstrumenten voor de kunstvakken in het basis- en voortgezet onderwijs die zijn beschreven in de literatuur tussen 2000 en 2013?

1.1 Afbakening van de studie

Eind jaren negentig zijn in de Verenigde Staten reeds verschillende overzichten van beoordelingsprocedures en instrumenten gepubliceerd

(Beattie, 1997; Gardner, 1996). Recenter verschenen reviewartikelen over beoordeling in de kunsten in handboeken zoals het “New Handbook of Research in Music Teaching and Learning” (Colwell & Richardson, 2002), in het “International Handbook of Research in Arts Education” (Bresler, 2007), in bundels over historische en huidige beoordelingspraktijken in de kunsteducatie (Taylor, 2006), en in een boek over formatieve beoordeling in muziekonderwijs (Fautley, 2010). Deze literatuur beschrijft een grote variëteit aan beoordelingsinstrumenten en -vormen, maar de huidige reviews richten zich dikwijls op één kunstdiscipline of beperken zich tot instrumenten voor ofwel het primair ofwel het secundair onderwijs. Deze review dekt het brede spectrum aan instrumenten en biedt daarmee de mogelijkheid om de te beoordelen competenties over de kunstdisciplines en onderwijsniveaus heen te vergelijken. Het doel van de huidige review is om de competenties en criteria die beoordeeld worden in de kunsten in kaart te brengen. Op die manier wordt inzichtelijk gemaakt op welke criteria er beoordeeld kan worden voor de afzonderlijke disciplines en er zal verwezen worden naar specifieke publicaties zodat ontwerpers van instrumenten voort kunnen bouwen op bestaande criteria en uitwerkingen daarvan in rubrics. Analytische beoordelingen waarbij duidelijke criteria zijn geformuleerd, verhogen de betrouwbaarheid en transparantie. Maar zulke criteria worden ook gewantrouwd omdat ze de voorspelbaarheid van uitkomsten vastleggen. Voor sommigen is daarom een holistische beoordeling meer geëigend (Bijvoorbeeld Sadler, 2009 en Boughton, 2013).

Naast het onderscheid in verschillende kunstdisciplines, wordt er in de literatuur onderscheid gemaakt tussen de verschillende functies van de beoordelingsvormen: formatief en summatief (Bijvoorbeeld Sluijsmans, Joosten-Ten Brinke & Van der Vleuten, 2013). De verschillende domeinen en functies resulteren in verschillende beoordelingsvormen variërend van gestandaardiseerde toetsen over de kennis van en waardering voor de kunsten tot het creëren van uitingen in het kunstzinnige medium, en van portfolio's tot kritieken. Ook bij de beoordeling

van het leren in de kunstvakken loopt men tegen klassieke problemen aan als validiteit, betrouwbaarheid, haalbaarheid en eerlijkheid.

In deze review staan de competenties (kennis, vaardigheden en houding) zoals die beoordeeld worden in de beoordelingsinstrumenten centraal. Bij leerplannen van kunstvakken is de gangbare hoofdingeling in competenties die tussen (re)productie en receptie, dat wil zeggen tussen het zelf maken of uitvoeren enerzijds en het kijken of luisteren naar, anderzijds. Die indeling geldt voor alle kunstdisciplines. Verdere onderverdelingen in dimensies zijn vaak opsplitsingen van de genoemde tweedeling (Kirchner & Haanstra, 2015). Een internationale vergelijking van leerplannen voor kunstvakken in vijftien landen (The College Board, 2011) concludeert dat deze zijn in te delen in drie categorieën: (1) het conceptualiseren, plannen en genereren van kunst, (2) het (re)produceren van kunst en (3) het reageren op kunst. De eerste twee samen kunnen als productie worden opgevat, het derde komt overeen met receptie. In de Amerikaanse nationale assessments (NAEP) wordt in de kunstvakken een vergelijkbaar onderscheid gemaakt in drie verschillende competenties: Creating (het bedenken en ontwikkelen van nieuwe artistieke ideeën en werk), Performing (realiseren van artistieke ideeën door interpretatie en presentatie), en Responding (begrijpen en evalueren hoe de kunsten betekenis verlenen).

Er zijn ook leerplannen waarbij er naast de hoofddimensies productie en receptie een derde dimensie bestaat die op beide andere dimensies betrekking heeft, bijvoorbeeld reflectie. De driedeling productie, receptie en reflectie is ook in Nederland al lange tijd gangbaar (Oostwoud Wijdenes en Haanstra, 1997). Reflectie is een metacognitieve vaardigheid en wordt in de leerplannen vaak opgevat als het denken over het maakproces en over het kijk- en beoordelingsproces, hetzij tijdens het proces zelf hetzij achteraf. In deze studie richten we ons op actieve competenties (creatie en uitvoering), receptieve competenties om kunst te ervaren en te interpreteren en reflectieve competenties in zowel actieve als receptieve kunsteducatie.

Een andersoortige indeling in competenties die in de kunstvakken worden nagestreefd

komt van Harland (2008). Hij onderscheidt kennis, technieken en vaardigheden, creativiteit en verbeeldingskracht en tenslotte het kunnen uitdrukken en analyseren van inhoudelijke betekenis door middel van kunst. Bij kennis, technieken en vaardigheden gaat het niet alleen om feitenkennis en technische vaardigheden, maar ook om interpretatieve en esthetische vaardigheden. Naast deze kunstgerelateerde competenties onderscheidt Harland vakoverstijgende competenties die worden nagestreefd op persoonlijk, sociaal en maatschappelijk gebied, zoals doorzettingsvermogen en het kunnen samenwerken. Aan deze indeling van Harland kunnen verschillende beoordelingscriteria ontleend worden die in de kunstvakken worden gehanteerd en die worden weerspiegeld in beoordelingsinstrumenten. Harland pleit voor een evenwicht tussen vaardigheid, creativiteit en betekenis. Maar op basis van onderzoek in Groot-Brittannië concludeert Harland dat er in de kunstdisciplines verschillend gewicht wordt toegekend aan criteria. Muziek en dans scoren hoog op de ontwikkeling van technische vaardigheden, bij theater is naast technische vaardigheden inhoudelijke betekenisgeving belangrijk. In de beeldende vakken zijn creativiteit en verbeeldingskracht verhoudingsgewijs belangrijk. Het bevorderen van creativiteit en creatief denken is ook in andere literatuur over kunstvakken een vaak genoemde doelstelling. Mede om die reden wil men niet alleen het beeldende eindproduct of de eindvoorstelling beoordelen, maar ook het artistieke werkproces dat heeft geleid tot het product. Hierbij speelt het bevorderen van creatief denken een rol. In het SLO leerplankader kunstzinnige oriëntatie vormt het creatieve proces (verdeeld in de fasen oriënteren, onderzoeken, uitvoeren en evalueren) het uitgangspunt om de competenties in te delen. Beoordeling van proces en product worden in deze studie onderscheiden.

2 Methode

2.1 De zoekopdracht

We startten de zoekopdracht in vijf (sociaal wetenschappelijke en kunstdiscipline specifieke) databases: ERIC, Psycinfo, Arts

& Architecture Index, RILM Abstracts of Music Literature en International Index to Performing Arts. Deze databases zijn geselecteerd op basis van relevantie voor de reviewstudie. Met het oog op actualiteit en haalbaarheid werd gezocht naar literatuur vanaf het jaar 2000 tot 2 oktober 2013.

We gebruikten zoektermen zoals gedefinieerd in de thesauri van de verschillende databases. Daarnaast werden onze eigen zoektermen ingevoerd. Gezocht werd in titel, trefwoorden en abstract van de publicatie. Zoekwoorden werden opgesteld in drie velden die gecombineerd werden: 'kunsten', 'beoordeling' en 'leerlingen'. We richtten ons op de kunstdisciplines muziek, beeldend, dans en theater. Tevens werden velden aangemaakt om 'language arts' en 'therapy' uit te sluiten, omdat deze woorden veel irrelevante publicaties opleverden. We kwamen uit op 155 zoektermen voor de verschillende databases (opvraagbaar bij de auteurs). De zoektocht leverde 7909 unieke publicaties op. Tenslotte werd nog gezocht in de database van het Landelijk Kennisinstituut voor Cultuureducatie en Amateurkunst (LKCA). Hier kon eenvoudigweg gebruik gemaakt worden van een enkele zoekterm 'toetsing'. Dit leverde nog eens 427 publicaties op.

2.2 Selectiecriteria

De selectie van de publicaties heeft plaatsgevonden in verschillende stappen door de twee eerste auteurs van dit artikel. We begonnen met brede en algemene selectiecriteria. In de vervolgstappen zijn de criteria steeds verder aangescherpt. De publicaties zijn eerst bekeken op relevantie. Op basis van titel en abstract werd besloten of een artikel werd opgenomen in de studie. Criteria bij de eerste selectie van de artikelen waren:

- 1) het artikel betreft één of meerdere kunstdisciplines;
- 2) er wordt een instrument of procedure voor beoordeling besproken en
- 3) het gaat om primair of secundair onderwijs.

Er was sprake van een instrument of procedure voor beoordeling zodra de auteurs één van de volgende zaken beschreven:

– een specifiek instrument dat in het artikel is opgenomen (of met voorbeelden van items) of dat ontwikkeld en gepubliceerd is door een professionele organisatie.

– formatieve of summatieve beoordelingsprocedures die uitgebreid beschreven worden in beoordelingsterminologie.

Veel publicaties hadden geen betrekking op kunstonderwijs. Het zoekwoord ‘art’ resulteerde in veel irrelevante ‘hits’, maar kon uiteraard niet uit de zoekopdracht verwijderd worden. Tevens hadden veel publicaties betrekking op hoger onderwijs. Ook deze werden niet in de studie opgenomen. Eveneens uitgesloten werden heel specifieke beoordelingsinstrumenten die alleen als toets bij één specifieke les over een bepaald onderwerp bruikbaar zijn. Tevens werden docentenvragenlijsten waarbij men inventariseerde welke beoordelingsinstrumenten docenten gebruiken, niet opgenomen. Deze instrumenten werden te beperkt beschreven. Als het een artikel betrof over aanleg als aangeboren eigenschap, dan werd de publicatie niet opgenomen wegens geringe relevantie voor het onderwijs. Artikelen die twijfel oproepen werden in een aparte map bewaard en door de tweede onderzoeker nogmaals beoordeeld op relevantie.

Deze eerste selectieronde waarin alle 8336 publicaties werden gescand op basis van titel en abstract resulteerde in 624 relevante publicaties. Onder de publicaties waren zowel artikelen waarbij het beoordelingsinstrument of de beoordelingsprocedure onderwerp van de publicatie was, als artikelen waarbij het ging om de meting van een afhankelijke variabele waarbij een artistiek product beoordeeld werd om de effectiviteit van een interventie te bepalen.

Om tot een verdere inperking van het aantal resultaten te komen werden de publicaties in een volgende ronde ingedeeld in ‘type’ publicatie. We onderscheidde boeken, dissertaties, collected work, conference proceedings, rapporten en journal artikelen. Er werd besloten om enkel peer reviewed journal artikelen op te nemen in de definitieve studie. Dit deden we enerzijds om de grote hoeveelheid literatuur in te perken, anderzijds fungeert opname in een peer reviewed

tijdschrift als een extra kwaliteitsindicator. Dit betrof 202 publicaties. De 16 relevante peer reviewed artikelen die we selecteerden uit de LKCA search zaten allen reeds binnen de 202 artikelen die we vonden in de overige databases.

Tijdens de analyse werden de 202 artikelen voor het eerst in full text gelezen. Dit betekende dat er wederom artikelen afvielen wegens irrelevantie. Artikelen van louter beschouwende aard (bijvoorbeeld over mogelijke functies van beoordeling) werden geschrapt. Het beoordelingsinstrument moet het hoofdonderwerp van het artikel zijn en het artikel moet gegevens over de kwaliteit van het instrument of de procedure bevatten. Uiteindelijk werden 65 publicaties geanalyseerd en opgenomen in de studie. Deze artikelen zijn afkomstig uit 30 peer reviewed tijdschriften en ze beschrijven gezamenlijk 74 beoordelingsinstrumenten. De tijdschriften waren zonder uitzondering Engelstalig, ze werden uitgegeven in de Verenigde Staten of Groot-Brittannië.

2.3 Analyse

Er is een codeersysteem ontwikkeld om de gevonden artikelen te coderen. Ook de codering werd door de eerste twee auteurs van het artikel gedaan. Codeercategorieën betroffen: bibliografische gegevens, kunstdiscipline, onderwijsniveau, beoordelaar, object van beoordeling, doel van beoordeling, toetstaak, scoringsinstrument, soort competentie en beoordelingscriteria. Telkens konden meerdere codes toegekend worden of gekozen voor ‘onbekend’ indien iets niet duidelijk werd uit de publicatie. In Tabel 1 zijn de codeercategorieën, de variabelen en definities van enkele variabelen opgenomen.

Bij het coderen is uitgegaan van de woorden van de auteur(s). Beoordelingscriteria uit de beoordelingsinstrumenten werden in letterlijke bewoordingen overgenomen uit de artikelen. Tevens werden de onderzoeksgegevens uit de geselecteerde publicaties aangehouden. Als een instrument in het betreffende artikel ingezet en getest werd in het primair onderwijs, werd dat onderwijsniveau ingevoerd. Mogelijkerwijs heeft een instrument een bredere toepasbaarheid, maar

Tabel 1
Codeercategorieën, variabelen en definities

Categorieën	Variabelen	definitie
bibliografische gegevens	auteur(s), jaartal, titel, journal, volume, issue, paginanummers, doi	
kunstdiscipline	beeldend, theater, dans, muziek	
onderwijsniveau	primair, secundair onderwijs	
beoordelaar	docent, zelf, peer, onafhankelijke beoordelaar	
object van beoordeling	product	Dat wat gemaakt is; de dans, het beeldend werkstuk, de uitvoering. In het onderwijs is dit meestal eindpunt van een proces.
	proces	Alles wat voorafgaat aan het eindproduct; hoe is men tot het eindproduct gekomen?
doel van beoordeling	formatief	Doel is feedback geven gericht op verbeteren.
	summatief	Doel is cijfer geven/ een oordeel vellen.
toetstaak	uiting in artistiek medium	De leerling maakt iets of voert iets uit in een artistiek medium. Dit wordt beoordeeld.
	mondelinge/schriftelijke vragen	De leerling beantwoordt vragen. Deze antwoorden worden beoordeeld.
	houding of gedrag; geen specifieke taak	Er is geen duidelijke taak. De houding van een leerling in het algemeen wordt beoordeeld.
	portfolio	De auteur geeft zelf aan dat het om portfolio gaat en gebruikt letterlijk het woord 'portfolio'.
scoringsinstrument	combinatietaak	Een combinatie van bovenstaande taken.
	criteria lijst	Een lijst met aspecten waarop beoordeeld wordt.
	rubric	Een lijst met aspecten waarop beoordeeld wordt én daarbij concrete beschrijvingen van beheersingsniveau's.
	zelfrapportage vragenlijst	Vragenlijst met antwoordschalen op basis waarvan leerlingen hun eigen gedrag/houding inschatten/beoordelen.
	commentaar/feedback	Mondelinge of schriftelijke feedback, kwalitatief.
soorten competenties	juryprocedure	Beschrijving van procedure voor een jury om tot een (holistisch) oordeel te komen.
	ankers	Voorbeeldproducten van leerlingen die een bepaalde score representeren.
	perform	Realiseren van artistieke ideeën door interpretatie en presentatie.
	create	Het bedenken en ontwikkelen van nieuwe artistieke ideeën en werk.
	respond	Begrijpen en evalueren hoe de kunsten betekenis verlenen.
	reflect	Reageren op eigen artistiek product of proces.

Tabel 2

Verdeling van de instrumenten over kunstdisciplines en schoolniveaus (aantallen en percentages)

	beeldend	muziek	dans	drama	instrumenten voor meerdere disciplines	totaal
primair	4 21%	14 42%	3 21%	3 75%	4 100%	28 38%
secundair	6 32%	14 42%	10 71%	1 25%	-	31 42%
primair én secundair	7 37%	4 12%	1 7%	-	-	12 16%
schoolniveau niet benoemd	2 11%	1 3%	-	-	-	3 4%
totaal	19 100%	33 100%	14 100%	4 100%	4 100%	74 100%

werd dat niet als zodanig gecodeerd. In sommige gevallen ging een artikel over meerdere instrumenten. In dat geval werd de publicatie vaker ingevoerd in het databestand. De eerste vijf artikelen zijn gezamenlijk gecodeerd. Daarna werden de artikelen verdeeld op basis van alfabetische volgorde van auteur. Tijdens het coderen werden logboeken bijgehouden met twijfelpunten. De logboeken werden regelmatig doorgesproken om een gelijke aanpak door beide onderzoekers te garanderen.

Tenslotte werd een secundaire analyse uitgevoerd op de beoordelingscriteria. Experts uit de vier verschillende kunstdisciplines hebben de beoordelingscriteria onafhankelijk van elkaar gecodeerd. Zij hadden allen een kunstopleiding in de betreffende kunstdiscipline gevolgd, zijn ervaren docenten en waren als onderzoeker verbonden aan het lectoraat van de Amsterdamse Hogeschool voor de Kunsten. Codeercategorieën ontstonden bottom up vanuit de data.

2.4 Overzicht van de instrumenten per kunstdiscipline en schoolniveau

We hebben beoordelingsinstrumenten geselecteerd die, ontworpen, gebruikt of getoetst zijn in het primair en secundair onderwijs. Enkele instrumenten zijn gebruikt bij jongere kinderen (vanaf 3 jaar). Omdat de scheidslijn tussen voorschools - en primair onderwijs internationaal niet eenduidig is, hebben we ervoor gekozen die instrumenten ook onder

het primair onderwijs te scharen. De instrumenten hadden betrekking op de kunstdisciplines beeldend, muziek, dans en drama. De verdeling van de instrumenten over de schoolniveaus en kunstdisciplines is als volgt. Zestien procent van de instrumenten is beschreven voor zowel primair als secundair onderwijs, 42% is alleen beschreven voor secundair onderwijs en 38% is alleen beschreven voor primair onderwijs. Bij 4% (N=3) van de instrumenten is het onderwijsniveau niet genoemd in het artikel. De verdeling van de disciplines over de schoolniveaus is min of meer gelijk.

Opvallend is dat bijna de helft (45%, N=33) van de instrumenten beschreven is voor de discipline muziek, 26% (N=19) voor de discipline beeldend, 19% (N=14) voor de discipline dans, 4 instrumenten (5%) voor de discipline drama en 4 instrumenten (5%) zijn beschreven voor alle disciplines samen. Een aantal van de instrumenten is specifiek gericht op één type muziekinstrument (bv. snaar- of blaasinstrumenten), genre (bv. ballet) of medium (bv. tekenen). In het secundair onderwijs zijn vaker specifieke instrumenten beschreven dan in het primair onderwijs.

3 Resultaten

Vraag 1: *Welke beoordelingsinstrumenten voor de kunstvakken in het basis- en voortgezet onderwijs zijn in de literatuur sinds 2000 beschreven?*

3.1 Soorten instrumenten

We maken in de beschrijving van de beoordelingsinstrumenten een onderscheid tussen de taak die een leerling moet doen (bv. een beeldend werkstuk maken) en het scoringsinstrument om die taak te beoordelen. Deze twee onderdelen samen vormen het beoordelingsinstrument. Het scoringsinstrument kan door de docent, maar ook door de leerling zelf of door een leeftijdsgenoot gebruikt worden om de taak te beoordelen. In sommige gevallen is de taak heel specifiek beschreven (kies drie ideeën, beelden of gevoelens op basis van een gegeven beeldend werk om daarmee een dans van max. 1 minuut samen te stellen) (Englebright & Mahoney, 2012). In dat geval is het scoringsinstrument vaak ook taakspecifiek. In andere gevallen is de taak heel open (bv. ‘dansen in de klas tijdens de les’) en is het scoringsmodel algemener toepasbaar (bv. een observatielijst op basis van criteria). In sommige gevallen is alleen het scoringsinstrument of alleen de taak beschreven. Ook maken we een onderscheid tussen instrumenten bedoeld voor formatieve of summatieve beoordeling.

3.2 Taken

De meeste taken in onze review betroffen ‘uitingen in het artistieke medium’. Het gaat om het toepassen of gebruiken van kennis en vaardigheden waarbij iets gemaakt wordt, zoals een dans, muziek, theater of beeldend werk. Oreck (2005) en Oreck, Owen en Baum (2003) ontwikkelden een performance based assessment voor muziek, dans en drama om potentieel talent te identificeren in heterogene groepen leerlingen in het primair onderwijs. Leerlingen krijgen vijf lessen in de betreffende discipline. Bij dans komen verschillende stijlen aan bod. Drama gaat vooral over improvisatieopdrachten en bij muziek komen vooral percussie en zang aan de orde. De leerlingen worden tijdens deze lessen geobserveerd door onafhankelijke observanten.

Een andere categorie taken is het beantwoorden van schriftelijke of mondelinge vragen. In Miller (2009) krijgen leerlingen bijvoorbeeld contrast- en vergelijkingsopgaven. Leerlingen krijgen hetzelfde muziekstuk in verschillende versies te horen en worden gevraagd om de stukken te vergelijken op

punten als melodie, harmonie, ritme en vorm.

Er zijn ook cases waarin de leerlingen geen specifieke taak krijgen, maar waarin ze als het ware worden ‘geobserveerd’ in hun bestaande gedrag of houding. Dunbar (2011) heeft een methode ontwikkeld om leerlingen in 30 seconden of minder te beoordelen. Deze methode past ze elke muzikles toe, waardoor de gezamenlijke observaties een compleet beeld van de leerling geven. Per les bepaalt ze welke competenties, kennis of vaardigheden ze wil beoordelen. Vervolgens scoort ze op een namenlijst van de klas of een leerling de betreffende vaardigheid goed, voldoende of onvoldoende laat zien. Ze observeert de leerlingen dus tijdens de les, zonder dat ze specifieke taken hebben gekregen.

Vooraf in de beeldende vakken is het samenstellen van een portfolio een veel gebruikte taak. Een portfolio bevat vaak verschillende onderdelen. Zo kan het werkstuk (beeldend) opgenomen worden, maar daarnaast ook bewijsmateriaal ten aanzien van het proces of geschreven reflecties. In dat geval bevat het portfolio zowel uitingen in het artistieke medium als andere taken. In het geval van Lindström (2006) bestaat het portfolio uit schetsen, ontwerpen en uiteindelijke producten, uit logboeken, inspiratiebronnen en videoverslagen van opgenomen gesprekken met leerlingen. Ten slotte zijn er nog taken of opdrachten waarin verschillende componenten gecombineerd worden.

In tabel 3 staat vermeld hoe de verschillende taken verdeeld zijn over de disciplines. Zoals gezegd geldt voor alle disciplines dat de ‘uiting in het artistieke medium’ de meest voorkomende taak is. Voor drama geldt zelfs dat dit het enige type taak is in de artikelen die zijn opgenomen in deze review. In de discipline beeldend is het samenstellen van een portfolio een belangrijke taak, deze komt minder voor bij de andere disciplines. Bij muziek wordt relatief vaker bestaand gedrag of een houding bekeken. Ook het beantwoorden van mondelinge of schriftelijke vragen komen relatief vaker voor bij muziek.

3.3 Scoringsinstrumenten

In tabel 4 staan de scoringsinstrumenten genoemd en hun verdeling over de disciplines.

Tabel 3

Taken en disciplines (aantallen en percentages)

	muziek	beeldend	dans	drama	instrument voor meerdere disciplines	totaal
uiting in het artistieke medium	21 64%	7 37%	9 64%	4 100%	-	41 56%
mondelijke/schriftelijke vragen	6 18%	1 5%	1 7%	-	-	8 11%
houding of gedrag; geen specifieke taak	3 9%	1 5%	2 14%	-	3 100%	9 12%
portfolio	2 6%	5 26%	1 7%	-	1	9 12%
combinatietaak	1 3%	5 26%	1 7%	-	-	7 9%
totaal	33 100%	19 100%	14 100%	4 100%	4 100%	74 100%

We zien dat criterialijsten en rubrics de meest voorkomende scoringsinstrumenten zijn. Bij de podiumkunsten wordt gebruik gemaakt van zelfrapportages, maar dit geldt niet voor de beeldende vakken. Behalve bij drama zijn we scoringsinstrumenten tegengekomen in de vorm van commentaar of opmerkingen (feedback). In de disciplines muziek en beeldend zijn we jurybeoordelingen tegengekomen, maar niet bij dans of drama. Het gebruik van ankers zijn we één keer tegengekomen, bij

beeldend. De scoringsinstrumenten werden ingevuld door de leerkracht (27%), door de leerling zelf (8%), door peers (3%), door een onafhankelijke beoordelaar (bijv. de onderzoeker, 22%), door een ander (4%) of het was onduidelijk wie de beoordelaar was (9%). In 27% van de gevallen waren meerdere van bovenstaanden beoordelaars.

Oreck (2005; Oreck et al., 2003) gebruikt voor muziek, dans en theater een criterialijst

Tabel 4

Scoringsinstrumenten en disciplines (aantallen en percentages)

	muziek	beeldend	dans	drama	instrumenten voor meerdere disciplines	totaal
criterialijst	16 48%	5 26%	3 21%	2 50%	4 100%	30 41%
rubric	13 39%	4 21%	7 50%	1 25%	-	25 34%
zelfrapportage vragenlijst	1 3%	-	2 14%	1 25%	-	4 5%
commentaar/feedback	1 3%	3 16%	2 14%	-	-	6 8%
juryprocedure	1 3%	1 5%	-	-	-	2 3%
onbekend	1 3%	5 26%	-	-	-	6 8%
ankers	-	1 5%	-	-	-	1 1%
totaal	33 100%	19 100%	14 100%	4 100%	4 100%	74 100%

voor vaardigheden, motivatie en creativiteit. Voorbeelden van criteria in de categorie vaardigheden voor dans zijn: 'fysieke controle', 'coördinatie en behendigheid' en 'ruimtelijk bewustzijn'. Bij muziek wordt motivatie bijvoorbeeld gemeten aan de hand van categorieën als: 'enthousiasme' en 'doorzettingsvermogen'. En een criterium voor creativiteit is bij drama bijvoorbeeld: 'verbeeldingskracht'. Later in dit artikel gaan we dieper in op de verschillende criteria die beschreven worden in de artikelen. Als ook competentieniveaus van verschillende criteria beschreven worden, dan is het scoringsinstrument een rubric. In het geval van Lindström (2006) zijn expliciet zowel processen als producten beoordeeld met behulp van een rubric.

Een voorbeeld van een zelfrapportagevragenlijst is de Dance Attitude Scales (Sanderson, 2000). De vragenlijst die gevalideerd wordt in de studie bevat verschillende schalen met stellingen op basis waarvan adolescenten hun eigen attitude beoordelen ten aanzien van dans en ballet. Leerlingen gaven op een vijf-punt schaal aan in hoeverre ze het met de stellingen eens of oneens waren.

Een andere vorm van scoring is het geven van commentaar of opmerkingen door docenten bij de taken die de leerlingen hebben uitgevoerd. Dit is in alle gevallen een kwalitatieve vorm van feedback geven. Watts (2010) beschrijft bijvoorbeeld dat veel groepsleerkrachten in het primair onderwijs de tekeningen van leerlingen niet beoordelen, omdat ze niet goed weten hoe. Watts heeft een lijst met vragen opgesteld die een docent kan gebruiken om met een leerling over een tekening in gesprek te gaan en zodoende de leerling te stimuleren om zich verder te ontwikkelen. Een voorbeeldvraag is: "*How are the people in your picture feeling? How might you show in a drawing the way one person feels about another?*" (p. 149).

Twee instrumenten maakten expliciet gebruik van een jury als scoringsinstrument. Chan, Chan, en Chau (2009) lieten de tekeningen van leerlingen bijvoorbeeld beoordelen door onafhankelijke juryleden als ondergemiddeld, gemiddeld of boven gemiddeld. Bergee (2007) liet muziekuitvoeringen van leerlingen door juryleden beoordelen en

deed dit op een schaal van 0 (slecht) tot 100 (excellent). Opgemerkt moet worden dat ook juryleden altijd minstens één beoordelingscriterium hanteren om de leerlingen te beoordelen. Bij jurybeoordelingen wordt gebruik gemaakt van een holistisch oordeel (op basis van 1 criterium) of van een analytisch oordeel (meerdere criteria).

Een laatste categorie kwam slechts eenmaal voor in deze review: het gebruik van ankers. Myford (2002) ontwikkelde de 'descriptive graphic'; een continue beoordelingschaal met ankers (voorbeelden van leerlingwerkstukken geplaatst in het continuüm van de beoordelingschaal). De ankers dienden als referentiepunten in de training van de beoordelaars.

3.4 Formatief/summatief toetsen

In principe bepaalt de docent of de beoordelaar of een instrument formatief of summatief ingezet wordt in het onderwijs; de docent kiest namelijk of een instrument wordt ingezet om vast te stellen of een niveau gehaald is en op basis daarvan een cijfer te geven (summatief gebruik) of om een instrument in te zetten voor feedback waarbij vooral aandacht is voor de volgende stap (formatief gebruik). Uiteraard zijn veel instrumenten zowel formatief als summatief te gebruiken. Niet alle instrumenten lenen zich echter voor summatief gebruik, zoals bijvoorbeeld de feedbackvragen van Watts (2010).

In de beschreven artikelen vinden we voor 43% beoordelingsinstrumenten die vooral summatief werden ingezet en voor 15% beoordelingsinstrumenten die formatief werden ingezet. Tweeëndertig procent van de instrumenten zijn in de beschreven artikelen voor beide type beoordelingen gebruikt. Voor 7 beoordelingsinstrumenten was niet duidelijk welke vorm van beoordeling het betrof.

Vraag 2: *Wat wordt er met deze instrumenten beoordeeld?*

In de beschreven artikelen wordt er in 62% van de gevallen naar een product gekeken (zoals een tekening of een muziekuitvoering), in vier procent van de gevallen naar een proces (bijvoorbeeld het maakproces van een tekening) en in 20% van de gevallen

naar beiden. Voor 10 beoordelingsinstrumenten kon op basis van het artikel niet vastgesteld worden of het om een product- of procesbeoordeling ging. De verhoudingen verschillen enigszins over de disciplines. Bij muziek, dans en drama gaat het hoofdzakelijk om productbeoordelingen, bij beeldend is er in bijna de helft van de instrumenten sprake van zowel een proces- als een productbeoordeling.

3.5 Soorten competenties

Zoals in de inleiding beschreven, maken we in navolging van de NAEP onderscheid in de competenties: *perform*, *create*, *respond* en vullen dit aan met *reflect*. In tabel 5 staat een overzicht van het voorkomen van de competenties in de beschreven artikelen over de verschillende disciplines.

De meeste instrumenten toetsen competenties in de domeinen *perform* en/of *create*. In de podiumkunsten (muziek, dans en drama) ligt de nadruk op *perform*, terwijl de beeldende vakken zich uitsluitend richten op *create*. *Respond* en *reflect* worden in mindere mate los getoetst in de beschreven instrumenten, maar komen wel voor in combinatie met de andere competenties.

3.6 Beoordelingscriteria

Vervolgens bekijken we de criteria die gehanteerd worden bij de beoordeling. Niet bij elk instrument is overigens sprake van criteria. Sommige artikelen beschrijven holistische beoordelingsprocedures, waarbij de beoordelaar een algemeen oordeel over een geheel geeft, gebaseerd op ervaring, gevoel of intuïtie. Ook bij formatieve procedures waarbij vooral kwalitatieve feedback wordt gegeven is niet altijd sprake van criteria. De criteria die wel genoemd zijn, zijn door ons per discipline opgesplitst en voorgelegd aan vier vakexperts met de vraag om de criteria te groeperen naar deelcompetenties of dimensies. In tabel 6 is de indeling in dimensies zoals gemaakt door de experts weergegeven in kolom 1. In kolom 2 zijn de criteria die daar onder vallen, soms in iets ingekorte vorm, weergegeven. In kolom 3 zijn veelal minder criteria vermeld dan het aantal in kolom 4 aangeeft, omdat identieke criteria slechts eenmaal zijn genoemd. In kolom 4 zijn de artikelen waaruit de criteria afkomstig zijn weergegeven.

Bij muziek meten de beoordelingsinstrumenten vooral aspecten van het uitvoeren: de interpretatie en mate van expressiviteit, de technische en ritmische beheersing. Ook

Tabel 5
Soorten competenties (aantallen en percentages)

	muziek	beeldend	dans	drama	instrumenten voor meerdere disciplines	totaal
perform	16 48%	-	5 36%	1 25%	-	22 30%
create	1 3%	6 32%	-	-	3 75%	10 14%
respond	8 24%	2 11%	1 7%	-	-	11 15%
reflect	1 3%	-	1 7%	-	-	2 3%
alles	2 6%	1 5%	1 7%	-	-	4 5%
perform+create	4 12%	-	3 21%	2 50%	-	9 12%
overige combinaties	1 3%	10 53%	3 21%	1 25%	1 25%	16 22%
totaal	33 100%	19 100%	14 100%	4 100%	4 100%	74 100%

veel criteria op het gebied van houding (zoals inzet, doorzettingsvermogen, nauwkeurigheid) hebben betrekking op het uitvoeren van muziek. Andere dimensies betreffen het gehoor/ klankvoorstelling en muzikaal vocabulaire. Hoewel muziek maken meestal een groepsgebeuren is, zijn er weinig instrumenten die aspecten van samenspel beoordelen. Ook weinig beoordelingsinstrumenten hebben betrekking op muzieknotatie en op het scheppen van muziek: componeren/ creativiteit.

Bij de beeldende vakken staat het beoordelen van beelddaspecten en originaliteit/creativiteit voorop. Bij beelddaspecten gaat het uitsluitend om producten beoordelen. Hoewel het bij beeldend niet om het uitvoeren van bestaande werken gaat, zoals in de podiumkunsten, zijn de beelddaspecten wel aspecten van de uitvoering van een idee of concept in materiaal. Echter, bij originaliteit/creativiteit in product en proces gaat het juist om het creërende aspect. Dat is ook het geval bij de dimensie onderzoek/experiment en concept/idee. Weinig instrumenten hebben betrekking op kunstbeschouwing, hoewel dat zeker in het voortgezet onderwijs een belangrijk onderdeel van het curriculum is.

Beoordelingsinstrumenten bij dans hebben net als bij muziek vooral betrekking op het uitvoeren: op technische aspecten, expressiviteit, muzikaal ritmische aspecten,

ruimtegebruik, improvisatie en fysieke aspecten. De weinige instrumenten op het gebied van theater hebben vooral betrekking op spel, zoals bijvoorbeeld expressie, geloofwaardigheid en fysiek bewustzijn. Er zijn geen instrumenten gevonden die zich richten op het schrijven van theaterteksten en op regie.

De te meten dimensies en de gebruikte criteria zijn in meerderheid disciplinespecifiek, maar er zijn op een hoger abstractieniveau veel overeenkomsten aan te wijzen. In alle kunstdisciplines richten instrumenten zich op technische beheersing, originaliteit/creativiteit, zelfreflectie, samenwerking en inzet/houding. De instrumenten beperken zich niet tot makkelijk te meten kennis of technische vaardigheden, maar hebben ook betrekking op kwalitatieve aspecten die typerend zijn voor de kunstvakken, zoals bijvoorbeeld expressie, bewegingskwaliteit, toonkwaliteit, sfeer, originaliteit, interpretatie en visie.

Ellis en Lawrence (2009) beschrijven een instrument met criteria dat bruikbaar is voor alle kunstdisciplines. De criteria zijn daarom in alle tabellen met criteria opgenomen. Het instrument, voor ‘Creative Learning Assessment’ bevat de volgende criteria: 1) confidence, independence and enjoyment, 2) collaboration and communication, 3) creativity, 4) strategies and skills, 5) knowledge and understanding, 6) reflection and evaluation.

Tabel 6a
Beoordelingscriteria muziek

dimensie	criteria	% van totaal aantal criteria (N=121)	auteurs
expressie/ interpretatie	consequent fraseren; dynamiek; een idee of gevoel uitdrukken; esthetische aantrekking; expressie; expressiviteit; expressiviteit (mate waarin de compositie muzikaal expressief is en esthetische gevoeligheid laat zien); frase; geschreven interpretatie; individuele toon; interpretatie; interpretatie fermaten; interpretatie/ muzikaal effect; melodie; melodisch motief; melodische inconsequentie; muzikale interpretatie; ontwikkeld melodisch motief; ontwikkeld ritmisch motief; ritmische inconsequentie; tempo; vibrato; vloeiende lijn	28 (23%)	Barbot & Lubart, 2012; Brophy, 2005; Byo, 2001; Crochet & Green, 2012; Fautley, 2005; Hewitt, 2001; Hickey, 2001; Latimer, Bergee, & Cohen, 2010; Norris & Borst, 2007; Nyland & Acker, 2012; Oreck, 2005; Oreck, Owen & Baum, 2003; Scott, 2001; Silveira, 2013; Zdzinsky & Barnes, 2002
technische beheersing	ademhaling; akkoorden correct wisselen; akkoorden stevig indrukken; articulatie/toon; boog; coördinatie; dictie; handen correct plaatsen; linkerhand; melodische nauwkeurigheid; noot nauwkeurigheid; noot nauwkeurigheid en intonatie; positie; ritme- en nootnauwkeurigheid; slag over alle snaren; stabiel tempo volhouden; stembereik; techniek/articulatie; techniek/dictie; vibrato	21 (17%)	Byo, 2001; Chiodo, 2001; Crochet & Green, 2012; Hewitt, 2001; Latimer, Bergee, & Cohen, 2010; Norris & Borst, 2007; Oreck, 2005; Oreck, Owen & Baum, 2003; Smith & Barnes, 2007; Zdzinsky & Barnes, 2002
gehoor/ klankvoorstelling	aandachtig luisteren; accuraat identificeren van overeenkomsten en verschillen (melodie, harmonie, timbre, ritme, vorm en expressiemiddelen); balans; geluidperceptie; intonatie; klankmenging; melodie; nauwkeurigheid in toon patroon reproductie; toon; toonkwaliteit	18 (15%)	Byo, 2001; Chiodo, 2001; Hewitt, 2001; Latimer, Bergee, & Cohen, 2010; Miller, 2009; Norris & Borst, 2007; Norris, 2000; Oreck, 2005; Oreck, Baum & Owen, 2003; Scott, 2001; Silveira, 2013; Zdzinsky & Barnes, 2002
ritmische beheersing	maat/ritme; metronomisch tempo; ontregeling van de slag in waarneembaar veranderde structuren; puls; ritme; ritme/ tempo; ritmische complexiteit; ritmische nauwkeurigheid; stabiel tempo volhouden; tempo	18 (15%)	Brophy, 2005; Byo, 2001; Chiodo, 2001; Crochet & Green, 2012; Fautley, 2005; Hewitt, 2001; Latimer, Bergee, & Cohen, 2010; Norris & Borst, 2007; Oreck, 2005; Oreck, Baum & Owen, 2003; Scott, 2001; Silveira, 2013; Smith & Barnes, 2007; Zdzinsky & Barnes, 2002
muzikaal vocabulaire/ grammatica	accuraat identificeren van overeenkomsten en verschillen (melodie, harmonie, timbre, ritme, vorm en expressiemiddelen); compositie en improvisatie; consequent fraseren; correct gebruik van muzikale terminologie; melodisch motief; nauwkeurigheid in muziek beschrijven; ontwikkeld melodisch motief ontwikkeld ritmisch motief; ritmisch motief; technische kwaliteit (mate waarin tonische en ritmische elementen technische competentie laten zien m.b.t. tonal center en ritmische regulariteit); toegenomen verticale opbouw; type kennis (feiten, conceptueel, procedureel, metacognitief; cognitief proces: onthouden, begrijpen, toepassen, analyseren, evalueren en creëren); uitwerking in muziek beschrijven; verwachte cadensformules en doorbreken van het stereotype; vorm	16 (13%)	Barbot & Lubart, 2012; Brophy, 2005; Chiodo, 2001; Fautley, 2005; Hanna, 2007; Miller, 2009; Oreck, 2005; Oreck, Owen & Baum, 2003; Scott, 2001

Tabel 6a (vervolg van vorige pagina)
Beoordelingscriteria muziek

dimensie	criteria	% van totaal aantal criteria (N=121)	auteurs
attitude	basisvaardigheden; presenteren; betrokken zijn; doorzetten bij moeilijkheid; doorzettingsvermogen; enthousiasme; interesse hebben; netheid; opdracht voltooien; overtuiging; presentatie; vaardigheid om te focussen; verantwoordelijkheid nemen; zekerheid	16 (13%)	Chiodo, 2001; Crochet & Green, 2012; Norris, 2000; Nyland & Acker, 2012; Oreck, 2005; Oreck, Owen & Baum, 2003; Smith & Barnes, 2007; Thorpe, 2012
creativiteit	creativiteit; creativiteit (mate van originaliteit en coherentie van de compositie); innovatief gebruik van ritmische elementen; originaliteit; verwachte cadensformules en doorbreken van het stereotype	6 (5%)	Barbot & Lubart, 2012; Chiodo, 2001; Fautley, 2005; Hickey, 2001; Scott, 2001; Thorpe, 2012
samenspel	balans; klankmenging; samenspel	5 (4%)	Latimer, Bergee, & Cohen, 2010; Norris & Borst, 2007; Smith & Barnes, 2007
muzieknotatie	muzieknotatie, netheid, partituur, schrijftechniek	3 (3%)	Chiodo, 2001; Silveira, 2013
rest/ onbekend	algemene kwaliteit; anders; muziektechnologie; kwaliteit van de inhoud (reflectief denken over, accurate beoordeling van de individuele kennis, vaardigheden en werk); vakmanschap; voltooiing inhoud	6 (5%)	Bergee, 2007; Chiodo, 2001; Hickey, 2001; Latimer, Bergee, & Cohen, 2010; Mills, 2009

Criteria muziek tellen op tot meer dan 100% omdat enkele criteria in meer dan één dimensie werden geplaatst

Tabel 6b
Beoordelingscriteria beeldend

dimensies	criteria	% van totaal aantal criteria (N=89)	auteurs
beeldaspecten (product)	compositie; contourversmelting; detail; esthetische organisatie; Hoe effectief gaat de leerling om met verre, nabije vormen en overlap?; Hoe effectief integreert de gehele compositie de manier waarop het hier geplaatst is?; In welke mate is het hier in de omgeving geplaatst?; kleur, vorm en compositie; lijnvoering; ontwerpruimte gebruik; organisatie (totale compositie); platte of ruimtelijke weergave; representatie; sfeer; veel details in het gezicht geplaatst; verhoudingen die kloppen met de werkelijkheid; vormdifferentiatie	20 (22%)	Beattie, 2000; Chan, Chan & Chau, 2009; Lindström, 2006; Myford, 2002; Tobias, 2000; Watts, 2010
originaliteit/ creativiteit (proces en product)	creativiteit; divergent denken; flexibel denken; flexibiliteit; identificatie van verre associaties; in welke mate zijn de ideeën origineel (inventief, ingenieus, niet-standaard) in concept en vorm; inventiviteit; originaliteit; risico nemen; speelsheid; uitwerking; vloeiendheid; vloeiendheid van ideeën; voorkeur voor nieuwheid	20 (22%)	Bahia & Trindade, 2012; Beattie, 2000; Clark & Zimmerman, 2001; Ellis & Lawrence, 2009; Lindström, 2006; Myford, 2002

Verder op volgende pagina.

Tabel 6b (vervolg van vorige pagina)
Beoordelingscriteria beeldend

dimensies	criteria	% van totaal aantal criteria (N=89)	auteurs
onderzoek/ experiment (proces)	dagdromen als incubatie activiteit; kritisch bronnen uit visuele cultuur analyseren en demonstreer begrip van doel, betekenis en context; onderzoekswerk; ontwikkel ideeën door experiment, exploratie en evaluatie; problemen formuleren en conceptualiseren; vaardigheid om vastleggen van artistiek probleem uit te stellen; vaardigheid om vastleggen van artistiek probleem uit te stellen; vaardigheid om voorbeelden te gebruiken; zoekt stimulans en kennis	9 (10%)	Beattie, 2000; Lindström, 2006; Torres Pereira de Eca, 2005
concept/idee (product)	combineren of herordenen van concepten; Hoe geschikt zijn de vormen voor de uitgedrukte ideeën?/Drukken de vormen de bedoelde betekenis uit? ; persoonlijke ideeën, intenties, ervaringen, informatie en meningen in visuele en andere vorm opslaan; toepassen van abstracte relaties om nieuw begrip te genereren; visie; zichtbaarheid van intentie	7 (8%)	Beattie, 2000; Lindström, 2006; Myford, 2002; Torres Pereira de Eca, 2005
zelfreflectie (proces)	capaciteit voor zelfbeoordeling; evalueren en verfijnen; kwaliteiten van het werk evalueren; persoonlijke heuristiek; presenteren van een coherente en georganiseerde selectie van werken en eindproducten die een persoonlijke en doordachte keuze laat zien en overeenkomt met intenties; reflectie en evaluatie; reflectiviteit	6 (7%)	Beattie, 2000; Ellis & Lawrence, 2009; Lindström, 2006; Torres Pereira de Eca, 2005
nieuwsgierigheid/ openheid (proces)	nieuwsgierigheid; openheid voor ervaring; tolerantie voor onbekende sensorische stimulatie; vragen stellend	5 (6%)	Beattie, 2000
inzet/houding (proces)	doorzettingsvermogen; toewijding; vertrouwen, onafhankelijkheid en plezier; zelfmotivatie	4 (4%)	Beattie, 2000; Ellis & Lawrence, 2009
overig	bewustzijn van gaten of missende delen; elegante oplossing; expressiviteit; feedback gebruiken; impact van het product; inspiratie; kennis en begrip; productiviteit; samenwerking en communicatie; sensitiviteit; strategieën en vaardigheden; synthese; vaardigheid om dingen te transformeren; vaardigheid om te abstraheren en van geheel naar delen te gaan; vakmanschap; voorkeur voor cognitieve complexiteit	18 (20%)	Bahia & Trindade, 2012; Beattie, 2000; Ellis & Lawrence, 2009; Lindström, 2006

Tabel 6c
Beoordelingscriteria dans

dimensies	criteria	% van totaal aantal criteria (N=101)	auteurs
dans techniek/ bewegings- uitvoering/actie	accuraatheid; (bewegings)actie; ademgebruik; articulatie van de voet; bewegingsnauwkeurigheid; coördinatie; coördinatie en lenigheid; coördinatie en verbinding; fysieke controle; gebruik van gewicht; gebruik van plie; gebruik van port de bras; gebruik van rotatie in benen en bekken; nauwkeurigheid / afstemming; pirouettes; placing van de ruggengraat; sprongen en lifts; stijl consistentie; technische bekwaamheid; voet- en armposities	24 (24%)	Harding, 2012; Harding & Haven, 2011; Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003; Warburton, 2002; White, 2012
kracht/ intentie/ expressie	artistieke intentie uitstralen; bewegingskwaliteit; bewegingskwaliteiten; expressiviteit; gebruik van focus en energie; gebruikt tijdens de voorstelling; houdt focus en concentratie tijdens de dans; intentionele energie; helderheid; kracht; richten van de focus; voert begin- en eindvormen helder uit; voert bewegingen zo volledig mogelijk uit;	17 (17%)	Englebright & Robson Mahoney, 2012; Harding, 2012; Harding & Haven, 2011; Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003; Warburton, 2002; White, 2012
muzikaliteit/ ritme/ bewe- gings-dynamiek	correcte dynamiek; dynamisch bereik; frasering en gebruik van overgangen; muzikaliteit; ritme; ritmische scherpte; vermogen tot fraseren; voert dans uit zonder onderbreking	15 (15%)	Englebright & Robson Mahoney, 2012; Harding, 2012; Harding & Haven, 2011; Oreck, Owen & Baum, 2003; Schmid, 2003; White, 2012
persoonlijke attitude	doorzettingsvermogen; toewijding t.a.v. training en leerproces; vaardigheid om correcties te ontvangen, toe te passen en te onthouden; vaardigheid om doelen te stellen, na te volgen en te bereiken; vaardigheid om te focussen; vertrouwen, onafhankelijkheid en plezier; wil en vaardigheid om risico te nemen; zelf regulatie tijdens creatief proces; zelfbewustzijn/ metacognitie tijdens creatief proces	11 (11%)	Ellis & Lawrence, 2009; Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003; White, 2012
ruimtegebruik	correcte ruimtelijke oriëntatie; driedimensionaliteit van het lichaam; gebruik van ruimte; ruimtegebruik; ruimtelijk bewustzijn; ruimtelijke accurateid	9 (9%)	Harding, 2012; Harding & Haven, 2011; Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003; Warburton, 2002; White, 2012
vermogen tot improviseren, creativiteit, choreografie	creativiteit; creëert 3 bewegingsfrases om idee, beeld of gevoel over een kunstwerk uit te drukken; creëert begin- en eindvormen; identificatie en implementatie van creatieve mogelijkheden; improvisatie; innovatie; strategieën en vaardigheden	8 (8%)	Ellis & Lawrence, 2009; Englebright & Robson Mahoney, 2012; Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003
samenwerking	betrokkenheid in creatief proces; positieve en open houding; gemeenschapszin; samenwerking; samenwerking en communicatie; samenwerking/ruimtelijk bewustzijn; toewijding aan gemeenschap van de klas	6 (6%)	Ellis & Lawrence, 2009; Oreck, 2005; Schmid, 2003; White, 2012

Tabel 6c (vervolg van vorige pagina)
Beoordelingscriteria dans

dimensies	criteria	% van totaal aantal criteria (N=101)	auteurs
bewegings-geheugen	observatie en herinnering van ritme; observatie en geheugen; vermogen om combinaties te onthouden en omkeren	3 (3%)	Oreck, 2005; Oreck, Owen & Baum, 2003; White, 2012
Uithoudings-vermogen/ fysieke fitheid/ flexibiliteit	bereik in beweging en extensie; fitheid; kwalitatief bewegingsbereik	3 (3%)	White, 2012
overig	beschrijft hoe beweging gebruikt is om een idee, beeld of gevoel uit te drukken; identificeert ideeën, beeldend of gevoelens die gekozen zijn [uit het kunstwerk]; kennis en begrip; reflectie en evaluatie; zelf evaluatie tijdens creatief proces	5 (5%)	Ellis & Lawrence, 2009; Englebright & Robson Mahoney, 2012; Schmid, 2003

Tabel 6d
Beoordelingscriteria theater

dimensies	criteria	% van totaal aantal criteria (N=37)	auteurs
spel	controle in lichaam en beweging (reizende en niet reizende bewegingen, bewustzijn van ruimte en tijdsconcept of ritmische beweging); creëren en implementeren van plot; expressieve kwaliteit; fysiek bewustzijn; geloofwaardige karakters neerzetten; mate van controle van stemgeluid; publieksetiquette; rol spelen; verbeelding	11 (30%)	Lan & Morgan, 2003; Lin, 2013; Oreck, 2005; Oreck, Owen & Baum, 2003
persoonlijke eigenschappen	betrokkenheid; betrokkenheid in dramatische situaties; focus/ toewijding; onafhankelijk werk; onthouden of reageren op beelden/ zintuigelijke stimuli; persoonlijke ervaring; vertrouwen, onafhankelijkheid en plezier	8 (22%)	Lin, 2013; Oreck, 2005; Oreck, Owen & Baum, 2003
kennis en begrip	herkennen van karakters; herkennen van plotontwikkeling; herkennen van 'special effects'; kennis en begrip; script kennen; tekstregels onthouden	6 (16%)	Ellis & Lawrence, 2009; Lan & Morgan, 2003; Lin, 2013
samenwerken	samenwerken; samenwerking	4 (11%)	Ellis & Lawrence, 2009; Lin, 2013; Oreck, 2005; Oreck, Owen & Baum, 2003
vormgeving en theatertechniek	creëren en implementeren van geluidseffecten; creëren en implementeren van visuele effecten; decorstukken gebruiken	3 (8%)	Lan & Morgan, 2003; Lin, 2013
creativiteit	creatieve kwaliteit; creativiteit	2 (5%)	Ellis & Lawrence, 2009; Lin, 2013
reflectie	reflectie en evaluaties; reflecties	2 (5%)	Ellis & Lawrence, 2009; Lin, 2013
overig	strategieën en vaardigheden	1 (3%)	Ellis & Lawrence, 2009

4 Samenvatting en conclusie

Dit artikel geeft een overzicht van 74 beoordelingsinstrumenten voor de kunstvakken (muziek, beeldend, drama en dans) in het primair en secundair onderwijs, zoals beschreven in de internationale peer reviewed literatuur tussen 2000 en oktober 2013. De huidige reviews richtten zich dikwijls op één kunstdiscipline of beperkten zich tot instrumenten voor ofwel het primair ofwel het secundair onderwijs. Deze review dekt het brede spectrum aan instrumenten.

Ook in de kunstvakken worden leerlingen steeds vaker beoordeeld. De taken en scoringsinstrumenten zijn echter divers en er is veel discussie over het nut en de noodzaak van het beoordelen. De verschillende benaderingen in de verschillende kunstvakken vragen om verschillende beoordelingsinstrumenten. In deze review selecteerden we 65 peer reviewed artikelen uit 30 internationale tijdschriften die 74 beoordelingsinstrumenten beschreven. Hiermee pretenderen we niet een volledig en up to date overzicht te hebben van de beoordelingsinstrumenten zoals ze in de praktijk toegepast worden, maar laten we wel zien wat er in de wetenschappelijke literatuur verschenen is over beoordelingsinstrumenten in de kunstvakken. De vaak genoemde diversiteit vinden we ook in deze literatuur terug. Beoordelingsinstrumenten voor muziek zijn duidelijk in de meerderheid, voor drama vinden we maar heel weinig instrumenten beschreven in de literatuur. Voor beeldend vinden we meer instrumenten dan voor dans. Er worden ongeveer evenveel instrumenten beschreven voor het primair onderwijs als voor het secundair onderwijs.

In deze selectie van beoordelingsinstrumenten maakten we een onderscheid in verschillende taken en diverse scoringsinstrumenten. Het meest voorkomende instrumentarium is een uiting in het artistieke medium (zoals een muziekkuitvoering of een beeldend product) dat beoordeeld wordt aan de hand van criterialijsten of rubrics. Dit geldt zonder uitzondering voor alle kunstdisciplines.

Om de vraag te beantwoorden wat er met de instrumenten wordt beoordeeld hebben we

een overzicht gemaakt van de belangrijkste competenties die de instrumenten bestrijken: perform, create, respond en reflect. Het blijkt dat de meeste instrumenten beschreven worden voor competenties die vallen binnen het spectrum van perform (podiumkunsten) en create (beeldend). Beoordelingscriteria die te maken hebben met het beschouwen van kunst (respond) en vooral met reflecteren komen verhoudingsgewijs weinig voor.

De beoordelingscriteria die worden gehanteerd om deze competenties (en bijbehorende deelcompetenties) te beoordelen laten een breed spectrum zien. Het is duidelijk dat elke kunstdiscipline zijn eigen terminologie en eigen vakspecifieke competenties beoordeelt in de verschillende instrumenten. Toch vinden we op een hoger abstractieniveau wel degelijk overeenkomsten tussen de verschillende disciplines. Algemene competenties die we bij meerdere disciplines tegenkomen zijn: technische beheersing, interpretatie (perform), originaliteit/creativiteit, inzet/houding, zelfreflectie en samenwerking. Dit zijn blijkbaar belangrijke dimensies bij de beoordeling in alle kunstdisciplines.

De reviewgegevens bevestigen de bevindingen van Harland dat er bij de verschillende kunstdisciplines sprake is van een onevenwichtige verdeling van criteria op het gebied van (1) kennis, technieken en vaardigheden, (2) creativiteit en verbeeldingskracht en (3) inhoudelijke betekenis. Beoordelingsinstrumenten op het gebied van technieken en vaardigheden zijn er veel, op het gebied van creativiteit en verbeeldingskracht veel minder en op het gebied van inhoudelijke betekenis van kunst nauwelijks. Zo ligt bij dans en muziek het zwaartepunt bij technische beheersing en interpretatie/expressie. Dit zijn duidelijke 'perform' dimensies. Er zijn nauwelijks criteria die gaan over componeren of choreografie. Ook bij theater zijn vooral perform-criteria gevonden; die betrekking hebben op spel. Bij beeldend lijkt er meer evenwicht door zowel veel criteria die gaan over de uitvoering van werk (beeldaspecten) als criteria die gaan over creativiteit en onderzoek. Ook zijn er beoordelingscriteria op het gebied van inhoudelijke concepten en ideeën. Het proces is bij beeldend

ook vaker object van beoordeling dan bij de andere disciplines.

De lijst met artikelen en de metagegevens over deze artikelen zoals we die voor deze review verzameld hebben zijn toegankelijk gemaakt via de website van het Landelijk Kennisinstituut voor Cultuureducatie en Amateurkunst (LKCA²). We hopen daarmee voor zowel leerkrachten, onderzoekers als vakexperts een bron van mogelijke beoordelingsinstrumenten te ontsluiten. De instrumenten kunnen ofwel (deels) direct bruikbaar zijn in de praktijk ofwel ter inspiratie dienen bij de constructie van beoordelingsinstrumenten in de kunstvakken. Daarbij moeten enkele kanttekeningen worden geplaatst. Zoals boven beschreven zijn in de verzamelde instrumenten niet alle soorten beoordelingscriteria vertegenwoordigd, die men op basis van een evenwichtige verdeling van kunst disciplines en gewenste competenties zou verwachten. Om meer evenwicht te bereiken zullen de verzamelde instrumenten aangevuld moeten worden. We weten uit de praktijk dat er veel meer beoordelingsinstrumenten beschikbaar zijn en gebruikt worden, maar deze instrumenten vinden kennelijk niet hun weg naar de wetenschappelijke literatuur. Bovendien zijn de in de review geraadpleegde tijdschriften zonder uitzondering afkomstig uit de Verenigde Staten of Groot-Brittannië. Veel artikelen zijn gevonden via de Amerikaanse database ERIC. Uit een proefsearch is gebleken dat in de Duitstalige literatuur diverse relevante publicaties op het gebied van muziek en beeldende kunst zijn verschenen.

Vervolgonderzoek zou zich enerzijds kunnen richten op een uitbreiding van deze review in de richting van anderstalige literatuur en handboeken voor de praktijk. Anderzijds zou onderzoek naar bruikbaarheid, validiteit en betrouwbaarheid van instrumenten wenselijk zijn. Er wordt in peer reviewed artikelen nog weinig gerapporteerd over instrumenten die gericht zijn op beoordelen bij theater en van kunstbeschuwing en reflectie. Onderzoek waarin dergelijke instrumenten geconstrueerd of onderzocht worden op kwaliteit is hier tevens gewenst.

Noot

- 1 Dit onderzoek werd gefinancierd door de Nederlandse Organisatie voor Wetenschappelijk Onderzoek (NWO-projectnummer 411-11-635). We bedanken Janneke Staaks, Lot Siebe, Ellen van Hoek, Debbie Klarenbeek en Borius van der Meulen voor hun bijdrage aan deze studie.
- 2 <http://www.lkca.nl/kennisdossiers/beoordelen/reviewstudie-assessment-in-kunsteducatie>

Literatuur

- Bahia, S., & Trindade, J.P. (2012). Entwining psychology and visual arts: A Classroom experience. *Psychology Research*, 2(2), 89-98.
- Barbot, B., & Lubart, T. (2012). Creative thinking in music: Its nature and assessment through musical exploratory behaviors. *Psychology of Aesthetics, Creativity & the Arts*, 6(3), 231-242. doi: 10.1037/a0027307
- Beattie, D. K. (1997). *Assessment in art education*. Worcester, MA: Davis Publications.
- Beattie, D.K. (2000). Creativity in art: the feasibility of assessing current conceptions in the school context. *Assessment in Education*, 7(2), 175-192. doi: 10.1080/713613331
- Bergee, M. J. (2007). Performer, rater, occasion, and sequence as sources of variability in music performance assessment. *Journal of Research in Music Education*, 55(4), 344-358. doi: 10.1177/0022429408317515
- Boughton, D. (2013). Assessment of Performance in the Visual Arts: What, How, and Why. In A. Karpati & E. Gaul Eds.), *From Child Art to Visual Culture of Youth - New Models and Tools for Assessment of Learning and Creation in Art Education* (pp.134-157). Bristol: Intellect Press.
- Bresler, L. (Ed.) (2007). *International handbook of research in arts education*. Dordrecht: Springer Verlag.
- Brophy, T. S. (2005). A longitudinal study of selected characteristics of children's melodic improvisations. *Journal of Research in Music Education*, 53(2), 120-133. doi: 10.2307/3345513
- Byo, J. L. (2001). Designing substantive playing tests - A model. *Music Educators Journal*, 88(2), 39-44. doi: 10.2307/3399741
- Chan, D. W., Chan, L., & Chau, A. (2009). Judging Drawing Abilities of Hong Kong Chinese Gifted

- Students: Could Nonexperts Make Expert-Like Judgments? *Gifted Child Quarterly* 53(1), 15-24. doi: 10.1177/0016986208326555
- Chiodo, P. (2001). Assessing a cast of thousands. *Music Educators Journal*, 87(6), 17-23. doi: 10.2307/3399687
- Clark, G., & E. Zimmerman (2001). Identifying artistically talented students in four rural communities in the United States. *Gifted Child Quarterly*, 45(104), 104-114. doi: 10.1177/001698620104500204
- College Board (2011). *International arts education standards: A survey of the arts education standards and practices of fifteen countries and regions*. New York: National Coalition of Core Arts Standards. [http://nccas.wikispaces.com/file/view/int%20standards%208%201%20\(final\).pdf](http://nccas.wikispaces.com/file/view/int%20standards%208%201%20(final).pdf)
- Colwell, R. & Richardson, C. (Eds.) (2002). *The New Handbook of Research in Music Teaching and Learning*. Oxford: Oxford University Press.
- Crochet, L. S., & Green, S.K. (2012). Examining progress across time with practical assessments in ensemble settings. *Music Educators Journal*, 98(3), 49-54. doi: 10.1177/002743111435276
- Dunbar, L. (2011). Performance assessment of the masses in 30 seconds or less. *General Music Today*, 25(2), 31-35. doi: 10.1177/104871311406266
- Ellis, S., & Lawrence, B. (2009). The influence of the Creative Learning Assessment (CLA) on children's learning and teachers' teaching. *Literacy*, 43 (1), 3-10. doi: 10.1111/j.1741-4369.2009.00509.x
- Englebright, K., & Mahoney, M. R. (2012). Assessment in elementary dance education. *Journal of Dance Education* 12(3), 87-92. doi: 1080/15290824.2012.701176
- Fautley, M. (2010). *Assessment in music education*. Oxford music education series.
- Gardner, H. (1996). The assessment of student learning in the arts. In D. Boughton, E. Eisner, & J. Ligtoet (Eds.) *Evaluating and assessing the visual arts in education* (pp 131-155). New York: Teachers College Press.
- Hanna, W. (2007). The New Bloom's Taxonomy: Implications for Music Education. *Arts Education Policy Review*, 108(4), 7-16. doi: 10.3200/AEPR.108.4.7-16
- Harding, M. (2012). Assessment in the high school technique class: Creating thinking dancers. *Journal of Dance Education*, 12(3), 93-98. doi :10.1080/15290824.2012.701172
- Harding, M., & Haven, B. (2009). The effects of peer coaching in the secondary arts classroom: Intentional watching in the dance studio. *Journal of Dance Education*, 9(2), 41-51. doi: 10.1080/15290824.2009.10387385
- Harland, J. (2008). Voorstellen voor een evenwichtiger kunsteducatie. *Cultuur+Onderwijs* 23,12-52. Utrecht: Cultuurnetwerk Nederland.
- Herpin, S., Washington, A.Q., & Li, J. (2012). *Improving the assessment of student learning in the arts- State of the field and recommendations*. Washington: National Endowments for the Arts.
- Hewitt, M. P. (2001). The effects of modeling, self-evaluation, and self-listening on junior high instrumentalists' music performance and practice attitude. *Journal of Research in Music Education* 49(4), 307-322. doi: 10.2307/3345614
- Hickey, M. (2001). An application of Amabile's consensual agreement technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education*, 49(3), 234-244. doi: 10.2307/3345709
- Kirchner, C., & Haanstra, F. (2015) *Expertenbefragung zu europäischen Lehrplänen im Fach Kunst*. ENViL. <http://envil.eu/europaeische-kunstlehrplaene-und-kompetenzdimensionen-im-vergleich/>
- Lan, W. Y., & Morgan, J. (2003). Videotaping as a means of self-monitoring to improve theater students' performance. *The Journal of Experimental Education*, 71(4), 371-381.
- Latimer, M. E., Bergee, M.J., & Cohen, M.L. (2010). Reliability and perceived pedagogical utility of a weighted music performance assessment rubric. *Journal of Research in Music Education*, 58(2), 168-183. doi: 10.1177/0022429410369836
- Lin, M.C. (2013). The development of a performance assessment with performing arts teachers in Taiwan--from national policy to classroom practice. *Research in Drama Education*, 18(3), 296-312. doi: 10.1080/13569783.2013.810926
- Lindström, L. (2006). Creativity: What Is It? Can You Assess It? Can It Be Taught? *International Journal of Art & Design Education*, 25(1), 53-66. doi: 10.1111/j.1476-8070.2006.00468.x
- Miller, B. A. (2009). Mom and Little Richard: The

- benefits of contrast-and-compare activities. *General Music Today*, 23(1), 4-11. doi: 10.1177/1048371309334234
- Mills, M. M. (2009). Capturing student progress via portfolios in the music classroom. *Music Educators Journal*, 96(2), 32-38. doi: 10.1177/0027432109351463
- Myford, C. M. (2002). Investigating design features of descriptive graphic rating scales. *Applied Measurement in Education* 15(2), 187-215. doi: 10.1207/s15324818AME1502_04
- Norris, C. E. (2000). Factors related to the validity of reproduction tonal memory tests. *Journal of Research in Music Education*, 48(1), 52-64. doi: 10.2307/3345456
- Norris, C. E., & Borst, J. (2007). An examination of the reliabilities of two choral festival adjudication forms. *Journal of Research in Music Education*, 55(3), 237-251. doi: 10.1177/002242940705500305
- Nyland, B., & Acker, A. (2012). Young children's musical explorations: The potential of using learning stories for recording, planning and assessing musical experiences in a preschool setting. *International Journal of Music Education*, 30(4), 328-340. doi: 10.1177/0255761412459162
- Onderwijsraad & Raad voor Cultuur (2012). *Cultuureducatie: leren, creëren, inspireren!* Den Haag: Onderwijsraad.
- Oostwoud Wijdenes, J., & Haanstra, F. (1997). *Over actief, receptief en reflectief*. Utrecht: LOKV.
- Oreck, B. A. (2005). A Powerful Conversation: Teachers and Artists Collaborate in Performance-Based Assessment. *Teaching Artist Journal*, 3(4), 220-227. doi: 10.1207/s1541180xtaj0304_2
- Oreck, B. A., Owen, S. V., & Baum, S. M. (2003). Validity, Reliability, and Equity Issues in an Observational Talent Assessment Process in the Performing Arts. *Journal for the Education of the Gifted* 27(1), 62-94.
- Sadler, D.R. (2009). Transforming holistic assessment and grading into a vehicle for complex learning. In G. Joughin (Ed.), *Assessment, learning and judgement in higher education* (pp1-19). Dordrecht, The Netherlands: Springer Science.
- Sanderson, P. (2000). The Development of Dance Attitude Scales. *Educational Research*, 42(1), 91-99. doi: 10.1080/001318800363944
- Schmid, D. (2003). Authentic assessment in the arts: Empowering students and teachers. *Journal of Dance Education*, 3(2), 65-73. doi: 10.1080/15290824.2003.10387231
- Scott, S. J. (2001). A Task-Centered Approach to Performance-Based Assessments. *General Music Today*, 14, 10-14. doi: 10.1080/15290824.2003.10387231
- Silveira, J. M. (2013). Idea Bank: Portfolios and assessment in music classes. *Music Educators Journal*, 99(3): 15-24. doi: 10.1177/0027432112470071
- SLO (2014). Leerplankader kunstzinnige oriëntatie. Geraadpleegd op 6 augustus 2014 van <http://www.kunstzinnigeoriëntatie.slo.nl>
- Sluijsmans, D., Joosten-ten Brinke, D., & Van der Vleuten, C. (2013). *Toetsen met leerwaarde. Een reviewstudie naar de effectieve kenmerken van formatief toetsen*. NWO-PROO rapport.
- Smith, B. P. & Barnes, G.V. (2007). Development and validation of an orchestra performance rating scale. *Journal of Research in Music Education*, 55(3), 268-280. doi: 10.1177/002242940705500307
- Soep, E. (2004). Assessment and visual arts education In E. W. Eisner & M.D.Day (Eds.), *Handbook of research and policy in art education*. (pp.579 - 583). New Jersey, London: Lawrence Erlbaum Publishers.
- Taylor, P. (Ed) (2006). *Assessment in arts education*. Porthmouth: Heinemann.
- Thorpe, V. (2012). Assessment rocks? The assessment of group composing for qualification. *Music Education Research*, 14(4), 417-429. doi: 10.1080/14613808.2012.699957
- Tobias, A. (2000). A test for the visual arts. *Visual Arts Research*, 26(1), 30-37.
- Torres Pereira De Eça, M. T. (2005). Using portfolios for external assessment: An experiment in Portugal. *International Journal of Art & Design Education*, 24(2), 209-218.
- Van der Hoeven, M., Jacobse, A., Van Lanschot Hubrecht, V., Rass, A., Roozen, L. Sluijsmans, R., . . . Ekster, W. (2014). *Cultuur in de Spiegel in de praktijk. Een leerplankader voor cultuuronderwijs*. Enschede: SLO.
- Watts, R. (2010). Responding to Children's Drawings. *Education 3-13: International Journal of Primary, Elementary and Early Years Education*, 38(2), 137-153. doi: 10.1080/03004270903107877

- Warburton, E.(2002). From talent identification to multidimensional assessment: Toward new models of evaluation in dance education. *Research in Dance Education*, 3(2), 103-121. doi:10.1080/1464789022000050480
- White, J. H. (2012). Motivating and evaluating growth in ballet technique. *Journal of Dance Education*, 12(3), 109-112. doi:10.1080/15290824.2012.701173
- Zdzinski, S. F., & Barnes, G.V. (2002). Development and validation of a string performance rating scale. *Journal of Research in Music Education*, 50(3), 245-255. doi: 10.2307/3345801

Auteurs

T.Groenendijk is docent en onderzoeker aan de Amsterdamse Hogeschool voor de Kunsten. **M. Damen** is onderzoeker bij het Noors agentschap voor kwaliteitszorg in het hoger onderwijs (NOKUT). **F. Haanstra** is lector Kunsten en Cultuureducatie aan de Amsterdamse Hogeschool voor de Kunsten. **C. van Boxtel** is hoogleraar vakdidactiek, in het bijzonder geschiedenisonderwijs, aan de Universiteit van Amsterdam.

Correspondentieadres:

T. Groenendijk,
Amsterdamse Hogeschool voor de Kunsten
Master Kunsteducatie
Waterlooplein 211-213
1011 PG Amsterdam
Nederland
talita.groenendijk@ahk.nl

Abstract

Assessment in Arts Education

This paper presents the results of a review study on assessment instruments and assessment criteria in arts education (dance, visual art, music and theater) for primary and secondary education. The research question is: What assessment instruments for arts education in primary and secondary education are described in the literature between 2000 and 2013, and what is being assessed? The results concern 74 instruments which are described in 65 articles in 30 peer reviewed journals. Most frequent instruments combine a task in which students express themselves in an artistic medium and a scoring instrument such as a list of criteria or a rubric. A broad variety of criteria is used in these instruments. Specific competencies are found for each artistic discipline, but on a more general level there are similarities. Technical skills and creativity are assessed in all artistic disciplines, but for each discipline in a different way.