

# 5. Beoordelingsinstrumenten in danseducatie

Talita Groenendijk, Marie-Louise Damen, Folkert Haanstra en Carla van Boxel

In de afgelopen decennia is internationaal gezien relatief veel aandacht geweest voor beoordeling in de danseducatie (Bonbright & McGreevy-Nichols, 2012). Vanaf de jaren negentig zou in de Verenigde Staten de National Assessment of Educational Progress (NAEP) dans in het primair en het voortgezet onderwijs net als de andere schoolvakken op grote schaal gaan beoordelen. Het bleef echter bij *fieldtests*, wegens een te beperkt aantal scholen met dansonderwijs. Toch betekende dit dat dans erkend werd als een van de 'geletterdheden' die ontwikkeld kunnen worden in het onderwijs (Bonbright & McGreevy-Nichols, 2012). Hoewel NAEP als doel heeft om onderwijsopbrengsten over een langere termijn in kaart te brengen en niet zozeer individuele leerlingprestaties meet, hebben deze beoordelingen veel danseducatoren en danseducatieonderzoekers geïnspireerd om instrumenten en procedures te ontwikkelen voor het beoordelen van individuele leerlingen in het primair en voortgezet onderwijs, deze te testen en daarover te publiceren.

In Nederland gaat de laatste jaren veel aandacht uit naar het ontwikkelen van doorlopende leerlijnen voor de kunstvakken (Hoeven et al., 2014). Om de voortgang van leerlingen te kunnen vaststellen zijn beoordelingsinstrumenten nodig. In deze bijdrage geven we een overzicht van beoordelingsinstrumenten voor dans die in de afgelopen decennia beschreven zijn in de wetenschappelijke literatuur.

Wat wordt er beoordeeld met de beschreven instrumenten? NAEP onderscheidt drie dimensies: *perform* (uitvoeren van dans), *create* (creëren van dans – choreografie) en *respond* (receptie van dans). Ter onderscheid van bekijken en reageren op dans door anderen (veelal professionals) en denken over het eigen dansen, kan een vierde dimensie worden toegevoegd (reflecteren). Reflecteren heeft betrekking op het denken over de eigen dans, zoals bijvoorbeeld bij zelfbeoordeling.

Eerdere publicaties bevatten al wel een overzicht van beoordeling in de danseducatie (zoals Warburton, 2006), maar deze waren niet systematisch en zeiden ook niets over de kwaliteit van de beoordelingsinstrumenten. In deze bijdrage willen we daarom wel iets zeggen over die kwaliteit: wat hebben de auteurs onderzocht ten aanzien van de kwaliteit van de beoordelingsinstrumenten?

## 5.1 Dilemma's bij het beoordelen in de danseducatie

De relatie tussen kunst en beoordeling geldt vaak als problematisch (Soep, 2004). Ook bij dans en beoordeling spelen verschillende dilemma's een rol die in dergelijke discussies vaak naar voren komen. In dans, zoals ook in andere kunsten, is sprake van vele verschillende stijlen en genres. Vraag is of beoordelingscriteria specifiek zijn voor een bepaalde stijl of techniek of dat het mogelijk is om algemeen toepasbare beoordelingsinstrumenten en – criteria te ontwikkelen. Kun je leerlingen met een achtergrond in verschillende dansstijlen zoals ballet, hiphop, jazz, volksdans en moderne dans, beoordelen met dezelfde criteria? Oliver en Sprague (2007) suggereren dat creatietaken (choreografie) gekleurd (*biased*) kunnen zijn door westerse ideeën over creativiteit en moderne dans. Dat wil zeggen: uitgaand van een westerse visie op creativiteit zouden leerlingen met een achtergrond in moderne dans wel eens bevoordeeld kunnen worden. Aandacht voor diversiteit lijkt dus een belangrijk punt te zijn als het gaat om eerlijke beoordeling in de danseducatie.

Diversiteit is er ook in niveau. In danseducatie in primair en voortgezet onderwijs kan een docent te maken hebben met grote niveaoverschillen tussen leerlingen. Sommige leerlingen volgden al danslessen buiten school, andere leerlingen niet. Moeten beoordelingen dan niet vooral gericht zijn op het meten van ontwikkeling van de leerlingen ten opzichte van het eigen aanvangsniveau? Eenmalige externe beoordelingen geven weinig zicht op de ontwikkeling van leerlingen in de tijd. Een belangrijke vraag lijkt dan

ook te zijn op welke wijze instrumenten het meten van ontwikkeling stimuleren of wat de mogelijkheden van instrumenten zijn om dit voor leerlingen zelf inzichtelijk te maken (Warburton, 2002; White, 2012).

Authenticiteit van de beoordeling is een ander belangrijk issue (Bonbright & McGreevy-Nichols, 1999). In hoeverre zijn de toetstaken en contexten betekenisvol voor leerlingen? Als dat niet het geval is, kun je je afvragen wat er dan gemeten wordt. Bij authenticiteit van de toets kan gekeken worden naar de taak, de sociale en fysieke context, de beoordelingscriteria en de vorm van de toets (Gullikers, 2011). In hoeverre komen deze overeen met hoe en waarop mensen in de professionele wereld of in de wereld buiten school beoordeeld worden? Bij dans is bijvoorbeeld samenwerking een gebruikelijk kenmerk van de professionele praktijk. Veel dansers dansen in gezelschappen; leerlingen dansen in hun vrije tijd samen. Het is echter niet altijd makkelijk om individuen in een groepssetting te beoordelen. In sommige gevallen kiezen docenten dan toch voor individuele toetsafname, wat wellicht als minder authentiek gezien kan worden. Hetzelfde geldt voor opwarmingsoefeningen. Ook in toetssituaties hebben leerlingen opwarming nodig, maar dit kost wel extra tijd die voor toetsafname niet altijd beschikbaar is.

Een ander dilemma betreft de registratie van dans. Verschillende beoordelingen vereisen video-opname om dans vast te leggen voor externe beoordeling. Het is echter de vraag of alle ruimtelijke beweging op een plat scherm kan worden vastgelegd en dans dus daadwerkelijk van scherm eerlijk beoordeeld kan worden.

In deze bijdrage geven we een overzicht van de publicaties van de afgelopen twintig jaar en gaan we tevens met enkele voorbeelden in op hoe ontwikkelaars van instrumenten omgaan met bovenstaande dilemma's. We beantwoorden de volgende vragen:

1. Welke beoordelingsinstrumenten voor dans in het basis- en voortgezet onderwijs zijn in de wetenschappelijke literatuur van 1993 tot 2013 beschreven?
2. Hoe worden in enkele voorbeeld-instrumenten dilemma's bij beoordeling in danseducatie (omgaan met diversiteit, authenticiteit, registratie) benaderd?
3. Wat wordt er met de gevonden instrumenten beoordeeld?
4. Wat rapporteren auteurs over de kwaliteit van de instrumenten?

## 5.2 Methode

### Het zoekproces

We hebben gezocht naar wetenschappelijke literatuur vanaf 1993 tot september 2013, bijna twintig jaar. Daarbij zijn de volgende zoektermen werden gebruikt en gecombineerd (AND) in het zoekstelsel van de Universiteit van Amsterdam (Catalogue Plus) waarin verschillende zoekmachines zijn gecombineerd (zoals ERIC, Web of science en JSTOR):

- assess\* OR evaluat\* OR reflect\* OR measur\* (subject)
- dance OR ballet OR choreography (subject)
- education OR learning NOT 'higher education' (all words)

Dit leverde negentig publicaties op die we beoordeelden op relevantie. Publicaties over andere vakken of over hoger onderwijs lieten we buiten beschouwing. Van de overgebleven 29 publicaties bleken er veertien betrekking te hebben op een specifiek beoordelingsinstrument of specifieke beoordelingsmethode. Deze artikelen werden aangevuld met de zoekresultaten over dans uit de grootschalige NWO-reviewstudie (Groenendijk, Damen, Haanstra & Van Boxtel, ingediend). Zo vonden we nog twee artikelen en kwamen we op een totaal van zestien publicaties, waarvan vijftien peer reviewed artikelen en een dissertatie (Ericsson, 1996).

We gingen uit van een beoordelingsinstrument zodra de auteurs één van de volgende dingen beschreven of testten:

- een specifiek instrument dat in de publicatie was opgenomen of een specifiek instrument dat al eerder door experts gemaakt en gepubliceerd was
- een formatieve of summatieve beoordelingsmethode of procedure die uitgebreid beschreven werd in beoordelingsterminologie

### Codeerprocedure

De zestien publicaties hebben we gecodeerd op verschillende onderdelen:

1. onderwijsniveau (primair of voortgezet onderwijs)
2. taak (taak om te dansen, taak om vragenlijst te beantwoorden, portfolio samenstellen, etc.)
3. scoringsinstrument (rubrics, criterialijst, etc)
4. object van beoordeling (uitvoering, creatie, receptie, reflectie)
5. beoordelingscriteria
6. informatie over de kwaliteit van de instrumenten
  - o betrouwbaarheid (in hoeverre is onderzocht of het instrument leidt tot stabiele oordelen ongeacht beoordelaar, context, etc?)*
  - o validiteit (in hoeverre is onderzocht of het instrument geldig is; meet het wat het moet meten?)*
  - o eerlijkheid (zeggen de auteurs iets over de eerlijkheid van het instrument; benadeelt het instrument niet onbedoeld bepaalde groepen leerlingen?)*
  - o bruikbaarheid (zeggen de auteurs iets over de praktische bruikbaarheid van het instrument in de klas?)*
  - o transparantie (zeggen de auteurs iets over de transparantie voor leerlingen?)*

Bij het coderen is uitgegaan van de woorden van de auteur(s). Beoordelingscriteria namen we letterlijk over uit de publicaties. Tevens hielden we de onderzoeksgegevens uit de geselecteerde publicaties aan. Als een instrument in het betreffende artikel ingezet en getest werd in het primair onderwijs, voerden we dat onderwijsniveau in. Mogelijkerwijs heeft een instrument een bredere toepasbaarheid, maar dat werd niet als zodanig gecodeerd. Ten slotte hebben we een secundaire analyse uitgevoerd. De ingevoerde beoordelingscriteria werden gecodeerd volgens het codeersysteem voor dans, ontwikkeld in voornoemde NWO-reviewstudie.

## 5.3 Resultaten

*Onderzoeksvraag 1: Welke beoordelingsinstrumenten voor dans in het basis- en voortgezet onderwijs zijn in de wetenschappelijke literatuur van 1993 tot 2013 beschreven?*

Alle zestien publicaties zijn weergegeven in tabel 1, evenals de namen van de beschreven beoordelingsinstrumenten, de typen instrumenten en geschiktheid voor primair of voortgezet onderwijs. Uit de tabel blijkt dat zes instrumenten in het primair onderwijs ingezet werden, negen instrumenten in het voortgezet onderwijs (veelal zowel onder- als bovenbouw) en één instrument in zowel primair als voortgezet onderwijs. Twee publicaties hadden betrekking op DTAP (Oreck, 2005; Oreck, Owen & Baum, 2003), een instrument dat speciaal is ontwikkeld voor talentidentificatie en tevens beschikbaar is voor theater en muziek (met aangepaste beoordelingscriteria).

Voor taken waarbij leerlingen een dans moeten uitvoeren of creëren, zijn relatief veel rubrics gepubliceerd (10) en enkele criterialijsten (3). In beide gevallen is sprake van beoordelingscriteria, maar in het geval van de rubric worden er tevens competentieniveaus beschreven voor elk criterium. Voorbeelden van rubrics zijn de Growth Grade Rubric (White, 2012) en een rubric voor het meten van originaliteit in het creëren van beweging (Kranicke & Pruitt, 2012). Voorbeelden van criterialijsten zijn de Folk Dance Performance Checklist (Slettum et al., 2001) en DTAP (Oreck et al., 2003).

We zoomen verder in op een van de voorbeelden. Kranicke en Pruitt (2012) beschreven een voorbeeldrubric voor het beoordelen van originaliteit in het creëren van beweging. Aan VO-leerlingen werd gevraagd om originaliteit te laten zien bij het toepassen van tijd, ruimte en energie binnen een dansfrase. Ze moesten daarbij ook uitleggen waarom hun toepassingen anders en nieuw zijn. De criteria staan in de eerste

kolom van de rubric onder elkaar: originaliteit in toepassing van tijd, originaliteit in toepassing van ruimte, originaliteit in toepassing van energie en de uitleg van hoe tijd, ruimte en energie vernieuwend en anders zijn ingezet. De rubric kent vier competentieniveaus, van 1 (niets geleerd) tot 4 (boven verwachting-onverwacht product). In vier kolommen staan de beschrijvingen per niveau van ieder criterium. Niveau 1 beschrijft dat een leerling geen duidelijke tempoverschillen liet zien, de ruimte niet gebruikte, geen krachtvariaties gebruikte en niet kon uitleggen waarom zijn toepassingen vernieuwend waren. Op niveau 2 gebruikt de leerling enkele aspecten van tijd, ruimte en energie, op niveau 3 gebruikt de leerling alle aspecten van tijd, ruimte en energie. De kolom van competentieniveau vier is echter leeg gelaten, omdat een dermate origineel product onverwacht is en dus niet vooraf door de dansdocent beschreven kan worden. Er staat wel dat de docent hier voorbeelden kan geven van zeer originele producten. De rubric kan de docent zowel voor groepen als voor individuen gebruiken, formatief (bij zelf- en *peer*-beoordeling) en summatief (door de docent voor een cijfer).

In sommige gevallen is de rubric of criterialijst taakspecifiek, dan wordt ook de bijbehorende taak beschreven. Voorbeelden van beschreven taken zijn: 'Creëer een 1 minuut (solo) dans die ongebruikelijke manieren laat zien om op en af een stoel te komen. Gebruik verschillende posities, niveaoverschillen, en "traveling" en "non traveling" bewegingen en een begin, midden en einde' (Oliver & Sprague, 2007, p. 76) of het creëren van een korte (1 minuut) choreografie waarin drie ideeën, beeldend of gevoels die ontstaan zijn naar aanleiding van beeldende kunst, naar voren komen (Englebright & Robson Mahoney, 2012, p88). Soms ging het om specifieke dansstijlen, zoals volksdans (Slettum, Fox, Looney & Jay, 2001) of ballet (White, 2009). En in sommige gevallen ging het expliciet om algemene instrumenten die met veel aandacht voor diversiteit ontwikkeld werden (bv. Oreck, Owen & Baum, 2003). De criteria in de rubrics en criterialijsten zijn onder onderzoeksvraag 3 beschreven.

Er waren twee publicaties over portfolio-beoordelingen (Giguere, 2012; Oliver & Sprague, 2007). Portfolio's bevatten vaak meer taken, zoals in het geval van Oliver & Sprague; leerlingen moesten zowel een dans uitvoeren, een choreografie samenstellen als danskritieken schrijven. Video's van de dans en de geschreven dan wel mondelinge opdrachten vormen samen dit portfolio. Bij Giguere ging het voornamelijk om beschrijvingen van doelen voor het vak, wekelijks geschreven reflecties en een samenvattende eindreflectie. De publicatie bevatte ook een lijstje met voorbeelden van aanmoedigingen voor de docent om reflecties in de portfolio's te stimuleren. Bijvoorbeeld: 'Write about one "correction" you or someone else received today' (Giguere, 2012, p. 101). In drie andere gevallen was sprake van vergelijkbare lijstjes met open reflectievragen (Englebright & Robson Mahoney, 2012; Harding, 2012; Harding & Haven, 2009). Dit waren onderdelen van een beoordeling met bijvoorbeeld ook een uitvoerings- of choreografie-onderdeel. Een vraag was bijvoorbeeld: 'What was your intention as a performer in this piece?' (Harding, 2009, p. 97). In het geval van Harding & Haven (2009) ging het om een formulier voor *peer coaching*, vragen die leerlingen kunnen gebruiken om elkaar te coachen, zoals 'What do you want the performer to learn?' (Harding & Haven, 2012, p. 46). We vonden twee zelfbeoordelingsvragenlijsten met gesloten vragen (Sanderson, 2000; Wong et al., 2012). Sanderson beschrijft twee schalen om houdingen te meten van jongeren tegenover ballet en tegenover mannelijke dansers. Hierbij is dus de jongere zelf de 'beoordelaar' van de eigen houding. Een voorbeelditem is: 'Ballet can look so pure, so beautiful' (p. 99). Leerlingen konden antwoorden geven op een vijfpuntschaal (eens-neens). Bij de Ballet Executive Scale (Wong et al., 2012) gaat het om een vragenlijst die leerlingen invullen over hun planning van dans leren, organisatie van dansgedrag, motivatie, impulscontrole en empathie voor andere dansers. Ook hier 'beoordeelden' leerlingen hun eigen houding en gedrag op een vijfpuntschaal.

Ten slotte betroffen enkele instrumenten danskritiek of herkenning van betekenis of vorm in dans (receptietaken; Oliver & Sprague, 2007; Schmid, 2003; Warburton, 2000; 2002). Een voorbeeld is de taak om een danskritiek te schrijven als onderdeel van het portfolio (Oliver & Sprague, 2007). Schmid 2003 beschrijft een specifieke rubric voor het beoordelen van danskritieken. Bij het instrument van Warburton moeten leerlingen vorelementen en betekenissen identificeren in dansvideofragmenten van professionele dansers.

**Tabel 1:** Overzicht van instrumenten en onderwijsniveau

Publicaties	Naam van het instrument	(Danstaak en) rubric			(Danstaak en) criteriaalijst	Portfolio	Open (reflectie) vragen	Gesloten vragenlijst	Receptietaak (en/of scoring-systeem)
		taakspectiek	stijlspectiek	algemeen					
Englebright & Robson Mahoney, 2012 <sup>1</sup>	OSPI (Office of the Superintendent of Public Instruction) developed performance assessment	X					X		
Ericson, 1996 <sup>1</sup>	Assessment of schoolchildren's performance in dance	X							
Giguere, 2012 <sup>2m3</sup>	Self Reflective Portfolio				X		X		
Harding, 2012 <sup>2m3</sup>	1) Dans techniek rubrics 2) Reflectievragen		X (1)				X (2)		
Harding & Haven, 2009 <sup>3</sup>	1) Rubrics 2) Peer coaching formulier		X (1)				X (2)		
Kranjick & Pruitt, 2012 <sup>2</sup>	Rubric voor het beoordelen van originaliteit in het creëren van beweging			X					
Oliver & Sprague, 2007 <sup>2 m3</sup>	Portfolio ingrediënten en scoring rubrics voor eindexamens in Rhode Island		X			X			X
Oreck, 2005 <sup>1</sup>	DTAP: Dance Talent Assessment Process					X			
Oreck, Owen & Baum, 2003 <sup>1</sup>	DTAP: Dance Talent Assessment Process					X			
Sanderson, 2000 <sup>2</sup>	Dance Attitude Scales							X	
Schmid, 2003 <sup>2m3</sup>	Voorbeeld rubrics voor het beoordelen van dans (productie en receptief)			X					X
Slettum, Fox, Looney & Jay, 2001 <sup>1</sup>	Folk Dance Performance Checklist						X		
Warburton, 2000 <sup>1</sup>	AFACM: Assessment of Prime Actions in Creative Movement	X							X
Warburton, 2002 <sup>2m2</sup>	MAID: Multi dimensional Assessment in Dance	X							X
White, 2012 <sup>2m3</sup>	Growth Grade Rubric (ballet)		X						
Wong et al., 2012 <sup>2m3</sup>	BES: Ballet Executive Scale							X	

1=primair onderwijs, 2=voortgezet onderwijs (12-15), 3= voortgezet onderwijs 16-18

*Onderzoeksvraag 2: Hoe worden in enkele voorbeeldinstrumenten dilemma's bij beoordeling in danseducatie (omgaan met diversiteit, authenticiteit, registratie) benaderd?*

Om deze vraag te illustreren beschrijven we hier de instrumenten DTAP van Oreck, Owen & Baum (2003) en MAiD van Warburton (2002) uitgebreider. De voorbeelden laten zien dat verschillende instrumenten een verschillende focus hebben als het gaat om beoordeling in dans en hoe ontwikkelaars van de instrumenten verschillende dilemma's verschillend benaderen.

### **DTAP: een stijlonafhankelijk instrument**

Het Dance Talent Assessment Process (DTAP) werd ontwikkeld voor talentidentificatie op scholen. De ontwikkelaars waren verschillende dansspecialisten die geselecteerd waren vanwege hun diverse stilistische achtergronden. Deze heterogene groep specialisten had als doel om 'dans-stijl-onafhankelijke beoordelingscriteria' te ontwikkelen. Een centraal idee was dat leerlingen beoordeeld worden in hun gebruikelijke klaslokaal, gedurende danslessen die door de heterogene groep specialisten ontwikkeld waren en waarin ze in groepen dansen; zo authentiek mogelijk. Specialisten vanuit verschillende achtergronden observeerden de leerlingen gedurende verschillende sessies.

De DTAP-criteria zijn gegroepeerd in drie hoofddimensionen: vaardigheden, motivatie en creativiteit:

Vaardigheden: fysieke controle, coördinatie & lenigheid, ruimtelijk bewustzijn, observatie & geheugen, ritme

Motivatie: vaardigheid om te focussen, doorzettingsvermogen

Creativiteit: expressiviteit, bewegingskwaliteiten, improvisatie

Per criterium zijn enkele indicatoren ontwikkeld. Geturfd werd of bepaalde indicatoren voorkwamen bij een specifieke leerling. Tevens kreeg elke leerling per les een holistisch oordeel. De beoordelaars bediscussieerden direct na de les wat ze hadden gezien en keken of ze tot overeenstemming konden komen.

We kunnen concluderen dat DTAP bijzonder gericht is op eerlijke beoordeling onder authentieke omstandigheden, rekening houdend met diversiteit tussen leerlingen. Er zijn ook daadwerkelijk metingen gedaan om die eerlijkheid te onderzoeken (zie onderzoeksvraag 4). Er is in ieder geval een serieuze poging gedaan om om te gaan met stilistische diversiteit. Er lijkt rekening gehouden te zijn met authenticiteit van de toets-situatie. Echter, de beoordelingen meten geen ontwikkeling over langere tijd. Diverse live observatoren omzeilen registratieproblemen. De vraag is echter of een docent in een gewone klassensituatie iets aan dit instrument heeft. Wellicht zijn de beproefde criteria bruikbaar.

### **MAiD: het meten van ontwikkeling**

Het Multidimensional Assessment Instrument in Dance (MAiD, Warburton, 2002) kan beschouwd worden als een verder ontwikkelde versie van de APACM (Warburton, 2000). Het instrument is bruikbaar voor 2 tot 15-jarigen en wordt afgenomen bij individuele leerlingen. Het is gebaseerd op de theorie van Laban en zijn bewegingsprincipes. Het doel is om productie en receptie in dans te meten en vooral de ontwikkeling van leerlingen op die verschillende onderdelen vast te leggen.

Het instrument bestaat uit videofragmenten van bewegingsfrases van toenemende complexiteit. Leerlingen moeten de bewegingen herkennen, produceren en de betekenis ervan uitdrukken. Ze doen dit voor drie bewegingselementen: actie, kracht en ruimtegebruik. Dit levert dus negen subtesten op. Bovendien is er sprake van vijf complexiteitsniveaus en de leerling kan voor ieder niveau op een subtest slagen of zakken. Met andere woorden, een leerling kan op elk niveau zakken of slagen op de herkennings-, productie- of de expressietaak en ook voor iedere dimensie (actie, kracht of ruimtegebruik). In rubrics (niet opgenomen in de publicatie) is vervolgens beschreven wanneer een leerling een bepaalde subtest haalt of niet haalt. Het productiegedeelte gaat om dansen naar aanleiding van een videofragment. Het dansen van de leerling wordt met video opgenomen voor beoordeling.

Vanuit de beschrijvingen in de tijdschriftartikelen is niet duidelijk om wat voor dansstijlen het gaat en wat nu precies een voldoende op een productietaak betekent

(gaat het dan bijvoorbeeld om accurate reproductie?). Zowel de herkenning- als de expressietaak lijken receptietaken te zijn, waarbij de eerste gaat om het herkennen van bewegingen en de tweede om uitleggen wat een danser probeerde uit te drukken.

Omdat de test vaak herhaald kan worden en verschillende complexiteitsniveaus kent, kan men met deze test ontwikkeling van leerlingen heel goed in kaart brengen. Om een voorbeeld te geven: een leerling die al veel danservaring heeft haalt misschien alles direct tot complexiteitsniveau 3 en productie zelfs op niveau 4 voor actie en kracht. Bij een herhaling van de test een half jaar later haalt hij wellicht alles op niveau 4. Een andere leerling start mogelijk op complexiteitsniveau 1 als het gaat om productie maar scoort al wel hoger op receptie. Op die manier kan voor ieder leerling de vooruitgang in kaart worden gebracht. Het vereist echter wel tijd voor afname met individuele leerlingen.

We kunnen dus concluderen dat dit instrument goed ontwikkeling in kaart brengt. Het is echter onduidelijk of er aandacht is voor stilistische diversiteit van leerlingen. Gezien de individuele afname kunnen we enkele vragen zetten bij de authenticiteit van deze toetsing. Dit instrument lijkt ook niet direct voor anderen bruikbaar, omdat de specifieke video's noodzakelijk zijn. Bij het meten van de effectiviteit van interventies bleek het instrument wel effectief (Warburton, 2000).

### *Onderzoeksvraag 3: Wat wordt er met deze instrumenten beoordeeld?*

In tabel 2 is weergegeven wat er met de verschillende instrumenten gemeten wordt. Uitvoering van dans lijkt in de meeste gevallen beoordeeld te worden (twaalf van de zestien publicaties). Creatie of choreografie komt in slechts zes publicaties beoordeeld, evenals receptie en reflectie. De door Sanderson (2000) beschreven 'dansattitudeschaal'; plaatsten we onder receptie, omdat het vooral meningen over en houdingen tegenover dans in het algemeen en ballet in het bijzonder betrof. Dit instrument meet echter niet direct receptievaardigheden.

**Tabel 2:** Welke competenties worden er beoordeeld?

Wat wordt er beoordeeld?	Aantal publicaties	Publicaties
Uitvoeren	4	Ericson, 1996; Harding & Haven, 2009; Slettum, Fox, Looney, & Jay, 2001; White, 2012
Creëren	1	Kranicke & Pruitt, 2012
Receptie	1	Sanderson, 2000
Reflectie	2	Giguere, 2012; Wong et al., 2012
Uitvoeren & creëren	2	Oreck, 2005; Oreck, Owen & Baum, 2003
Uitvoeren & receptie	2	Warburton, 2000; Warburton, 2002
Uitvoeren & reflecteren	1	Harding, 2012
Uitvoeren, creëren, receptie & reflectie	3	Oliver & Sprague, 2007; Englebright & Robson Mahoney, 2012; Schmid, 2003
<b>Totaal</b>	<b>16</b>	

Tabel 3 bevat een overzicht van de beoordelingscriteria, de hoofdcategorieën waaronder deze geschaard zijn en de bijbehorende publicaties. In slechts vier gevallen was er geen sprake van gerapporteerde beoordelingscriteria (Giguere, 2012; Sanderson, 2000; Warburton, 2000; Wong et al., 2012). De meeste criteria (23%) betreffen danstechniek. Daarop volgen expressie (16%), muzikaliteit (15%), houding (12%), creativiteit (10%) ruimtegebruik (7%), samenwerking (6%), receptie (4%), bewegingsgeheugen (3%) en fysieke fitheid (3%).

De nadruk ligt dus erg op de uitvoering van dans en minder op choreografische kwaliteiten van leerlingen. Ook houdingen spelen een grote rol, maar samenwerking betreft slechts 6% van de criteria. Criteria over het kijken naar en waarderen van dans (receptie) kwamen weinig in de publicaties naar voren.

**Tabel 3:** Beoordelingscriteria

Dimensies	Criteria	Publicaties	% van totaal aantal criteria (N=137)
Danstechniek/ bewegings- uitvoering/ actie	bewegings)actie; accuraatheid; ademgebruik; articulatie van de voet; basale danspas 'slide'; basale danspas: boog; basale danspas: 'castoff'; basale danspas: 'yesh' handen <sup>1</sup> ; bewegingsnauwkeurigheid; coördinatie; coördinatie en lenigheid; coördinatie en verbinding; fysieke controle; gebruik van gewicht; gebruik van plié; gebruik van port de bras; gebruik van rotatie in benen en bekken; houding; nauwkeurigheid / afstemming; pirouettes; placering van de ruggengraat; sprongen en lifts; stijl consistentie; stilistische vloeiendheid in twee verschillende genres; technische bekwaamheid; voet- en arposities	Ericson, 1996; Harding & Haven, 2011; Harding, 2012; Oliver & Sprague, 2007; Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003; Slettum, Fox, Looney & Jay, 2001; Warburton, 2002; White, 2012	31 (23%)
Kracht/ intentie/ expressie	artistieke intentie uitstralen; bewegingskarakter; bewegingskwaliteit; bewegingskwaliteiten; communiceert ideeën, emoties of culturele esthetiek; expressie van een kwaliteit, stijl of muzikaliteit; expressiviteit; gebaar; gebruik van focus en energie; gebruikt intentionele energie gedurende hele uitvoering; helderheid; kinesthetisch bewustzijn en bewegingsvaardigheden; kracht; richten van de focus; uitvoeringskwaliteit; voert alle beweging in volle mogelijkheid uit; voert begin- en eindvormen helder uit	Englebright & Robson Mahony, 2012; Ericson, 1996; Harding, 2012; Harding & Haven, 2011; Oliver & Sprague, 2007; Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003; Slettum, Fox, Looney & Jay, 2001; Warburton, 2002; White, 2012	22 (16%)
Muzikaliteit/ ritme/ bewegings- dynamiek	correcte dynamiek; dynamisch bereik; frasering; frasering en gebruik van overgangen; muzikaliteit; muzikaliteit/ritme; opeenvolging; ritme; ritme & pulsering; ritmische scherpte; vermogen tot fraseren	Ericson, 1996; Harding, 2012; Harding & Haven, 2011; Oliver & Sprague, 2007; Oreck, Owen & Baum, 2003; Schmid, 2003; Slettum, Fox, Looney & Jay, 2001; White, 2012	20 (15%)
Persoonlijke attitude	bereidwilligheid; betrokkenheid; doorzettingsvermogen; houdt focus en concentratie gedurende de dans; oplettendheid; plezier; toewijding in 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; wil en vaardigheid om risico te nemen; zelf regulatie tijdens creatief proces; zelfbewustzijn/ metacognitie tijdens creatief proces	Englebright & Robson Mahony, 2012; Ericson, 1996; Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003; White, 2012	16 (12%)



Vermogen tot improviseren, creativiteit, choreografie	benadrukt hoe toepassing van tijd, ruimte en energie in de dansfrase anders en nieuw is; bewegingskwaliteiten; creëert beginnen eindvormen; creëert bewegingsfrases om idee, beeld of gevoel uit het kunstwerk uit te drukken; gebruikt specifieke bewegingselementen; identificatie en implementatie van creatieve mogelijkheden; improvisatie; innovatie; laat originaliteit zien bij het toepassen van energie in gecreëerde dansfrase; laat originaliteit zien bij het toepassen van ruimte in gecreëerde dansfrase; laat originaliteit zien bij het toepassen van tijd in gecreëerde dansfrase; lost bewegingsprobleem op; presenteert een originele beweging	Englebright & Robson Mahony, 2012; Kranicke & Pruitt, 2012; Oliver & Sprague, 2007; Oreck, Owen & Baum, 2003; Schmid, 2003	14 (10%)
Ruimtegebruik	correcte ruimtelijke oriëntatie; driedimensionaliteit van het lichaam; gebruik van ruimte; ruimtegebruik; ruimtelijk bewustzijn; ruimtelijke nauwkeurigheid; ruimtelijke relatie	Harding, 2012; Harding & Haven, 2011; Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003; Slettum, Fox, Looney & Jay, 2001; Warburton, 2002; White, 2012; White, 2012	10 (7%)
Samenwerking	betrokkenheid in creatief proces; integratie; positieve en open houding; gemeenschapszin; samenwerking; samenwerking/ ruimtelijk bewustzijn; sociale sfeer; toewijding aan gemeenschap van de klas; vorm-groep	Ericson, 1996; Oreck, 2005; Schmid, 2003; White, 2012	8 (6%)
Receptie	beschrijft hoe beweging gebruikt wordt om ideeën, beelden of gevoelens uit te drukken; esthetische vragen over dans beantwoorden; identificeert ideeën, beelden of gevoelens uit het kunstwerk; kennis van dans als kunstvorm; kennis van proces van danskritiek	Englebright & Robson Mahony, 2012; Oliver & Sprague, 2007	5 (4%)
Bewegingsgeheugen	observatie en herinnering van ritme; observatie en geheugen; reproductie; vermogen om combinaties te onthouden en om te keren	Ericson, 1996; Oreck, 2005; Oreck, Owen & Baum, 2003; White, 2012	4 (3%)
Uithoudingsvermogen/ fysieke fitheid/ flexibiliteit	bereik in beweging en extensie; fitheid; kwalitatief bewegingsbereik; lichaamsextensie	Ericson, 1996; White, 2012	4 (3%)
Overig	voert dans uit zonder onderbreking; vorm-individueel; zelf evaluatie tijdens creatief proces	Englebright & Robson Mahony, 2012; Ericson, 1996; Schmid, 2003	3 (2%)
Overig	voert dans uit zonder onderbreking; vorm-individueel; zelf evaluatie tijdens creatief proces	Englebright & Robson Mahony, 2012; Ericson, 1996; Schmid, 2003	3 (2%)

<sup>1</sup> Specifieke volksdansterm, evenals boog en castoff

*Onderzoeksvraag 4: Wat rapporteren auteurs over de kwaliteit van deze instrumenten?*

In tabel 4 is weergegeven wat de auteurs van de publicaties hebben gerapporteerd over de kwaliteit van de instrumenten. In veel gevallen is in ieder geval sprake van pilot-testing of langdurig praktijkgebruik (negen publicaties). Empirische toetsing van validiteit en betrouwbaarheid van de instrumenten kwam veel minder voor (minder dan de helft van de gevallen). Bij betrouwbaarheid kon het zowel gaan om interbeoordelaars-betrouwbaarheid (komen beoordelaars onafhankelijk van elkaar tot hetzelfde oordeel?) als schaalbetrouwbaarheid (consistentie van de schaal: dragen verschillende criteria/vragen van het instrument stabiel bij aan het oordeel?) dan wel stabiliteit van het instrument over de tijd (wordt bij beoordeling op verschillende momenten hetzelfde oordeel gegeven?). Ook bij de validiteit ging het om verschillende toetsen, zoals factoranalyses om constructvaliditeit (wordt er daadwerkelijk één ding gemeten) vast te stellen of andere vormen van validiteit om te toetsen of de meting daadwerkelijk het beoogde meet.

**Tabel 4:** Kwaliteit van de beoordelingsinstrumenten

Kwaliteitscriteria	Aantal publicaties	Publicaties
Betrouwbaarheid (empirisch getoetst)	6	Ericson, 1996; Oreck, Owen & Baum, 2003; Sanderson, 2000; Slettum, Fox, Looney, & Jay, 2001; Warburton, 2000; Wong et al, 2012
Validiteit (empirisch getoetst)	3	Oreck, Owen & Baum, 2003; Sanderson, 2000; Wong et al, 2012
Pilottest of langdurig praktijkgebruik	9	Englebright & Robson Mahoney, 2012; Ericson, 1996; Harding & Haven, 2009; Oliver & Sprague, 2007; Oreck, Owen, & Baum, 2003; Sanderson, 2000; Slettum, Fox, Looney, & Jay, 2001; Warburton, 2000; White, 2012
Gebaseerd op theorie	4	Oreck, 2005; Warburton, 2000; 2002; Wong et al, 2012
Expertraadpleging	3	Englebright & Robson Mahoney, 2012; Oreck, Owen, & Baum, 2003; Slettum, Fox, Looney, & Jay, 2001
Authenticiteit	4	Oreck, 2005; Oreck, Owen & Baum, 2003; Schmid, 2003; Warburton, 2002
Eerlijkheid	4	Oliver & Sprague, 2007; Oreck, 2005; Oreck, Owen & Baum, 2003; Warburton, 2002
Haalbaarheid	3	Oreck, 2005; Oreck, Owen & Baum, 2003; Slettum, Fox, Looney, & Jay, 2001
Transparantie	3	Harding, 2012; Harding & Haven, 2009; Oreck, 2005

Een interessante publicatie in dit verband is die van Oreck, Owen & Baum (2003) die veel verschillende kwaliteitstoetsen uitvoerden. Ten eerste deden ze een factoranalyse om constructvaliditeit vast te stellen. Vervolgens keken ze of de score verschilde van andere metingen zoals een wiskundetest (discriminant validiteit) en of gender of etniciteit de score niet onbedoeld beïnvloedde, dus of bijvoorbeeld jongens systematisch anders beoordeeld worden dan meisjes. Ze onderzochten ook of er een positief verband bestond met andere docentoordelen (concurrent validiteit). Ten slotte keken ze of de als talentvol geselecteerde leerlingen anders scoorden op het instrument dan de niet als talentvol geselecteerde leerlingen (contrasting groups evidence) en of door het instrument geselecteerde leerlingen later succesvol waren in vervolgonderwijs (predictieve validiteit). Voor de betrouwbaarheid onderzochten ze of de beoordelaars het met elkaar eens waren (interrater reliability) en of de meting stabiel was (op verschillende momenten vergelijk-

bare scores gemeten). Het onderzoek naar DTAP van Oreck en collega's is dus een mooi voorbeeld van een empirisch grondig onderzocht beoordelingsinstrumentarium.

In sommige publicaties zeiden auteurs enkel iets over praktische haalbaarheid, transparantie of authenticiteit van het instrument of de procedure. We kunnen concluderen dat lang niet alle instrumenten uitvoerig onderzocht werden.

## 5.4 Conclusie

Deze bijdrage geeft een overzicht van beoordelingsinstrumenten in dans in primair en voortgezet onderwijs zoals de afgelopen twintig jaar beschreven in de wetenschappelijke literatuur. Het blijkt om erg verschillende instrumenten te gaan. In de meeste gevallen gaat het om taken waarbij leerlingen daadwerkelijk moeten dansen en een rubric om die dans te beoordelen. In sommige gevallen zijn er criterialijsten voor de scoring. Twee publicaties betroffen portfolio's en nog eens twee zelfbeoordelingsvragenlijsten. Ten slotte kwamen er reflectievragen voor en uitgebreide danskriektaken. Sommige instrumenten betroffen specifieke dansstijlen, terwijl andere instrumenten juist ontwikkeld waren om elke dans ongeacht stijl te kunnen beoordelen.

Uitvoering van dans werd in de meeste instrumenten beoordeeld. Creatie was minder vaak object van beoordeling. Dit was ook terug te zien in de beoordelingscriteria. Veel daarvan betroffen danstechniek. Improvisatie en creativiteit werden minder vaak beoordeeld. Ook samenwerking kwam lang niet in alle instrumenten voor.

Bij de meeste instrumenten wordt wel iets geschreven over de kwaliteit ervan. Dit varieerde van het noemen van een pilottest of een theoretische basis tot grondig empirisch onderzoek naar validiteit en betrouwbaarheid van het instrument. Omdat in de meeste gevallen de kwaliteit heel verschillend onderzocht is, is deze niet onderling vergelijkbaar. Dit is ook niet zinvol, omdat de instrumenten vaak verschillen in doelstelling of doelgroep. In vrijwel alle gevallen komt men wel tot een positieve conclusie over de kwaliteit, maar vaak zouden de instrumenten ook nog uitgebreider onderzocht kunnen worden.

Twee voorbeelden lieten zien welke keuzes ontwikkelaars maakten voor het omgaan met diversiteit in stijl, techniek, achtergrond en aanvangsniveau. Uiteraard zijn dit niet de enige manieren. Oliver en Sprague (2007) beschreven een geheel andere manier om die diversiteit ruimte te geven: leerlingen kiezen hier twee van de drie dimensies 'uitvoeren, creëren of receptie' en moeten in minimaal twee zelfgekozen dansstijlen vloeiendheid laten zien.

Hoewel er relatief veel gepubliceerd is over beoordeling in dans – vermoedelijk meer dan in de decennia voor 1990 – waren niet alle publicaties vernieuwend of verrassend. Waar bijvoorbeeld bij beoordelingsinstrumenten voor beeldende kunst veel aandacht is voor het proces, lijkt dit bij dans minder het geval. Ross (1994, p. 13) zegt hierover echter: 'The process of understanding an assessment task or a student sampling trial and error solution may be very useful process for recording and assessing.' Het gebruik van video beschrijven auteurs nu voornamelijk als problematisch in beoordelingssituaties, maar dit kan leerlingen ook veel mogelijkheden om zelf of met klasgenoten formatief te beoordelen of te reflecteren op hun eigen dans. Dit kwam in de aangetroffen publicaties nog niet aan bod. Enkel Harding (2009) en Harding & Haven (2012) beschrijven peer learning en peer assessment uitgebreid als belangrijke leerprocessen bij dans.

Ondanks de hierboven genoemde beperkingen kunnen de aangetroffen beoordelingsinstrumenten behulpzaam zijn voor de praktijk. Ze kunnen in enkele gevallen direct gebruikt worden, maar ook dienen ter inspiratie om zelf instrumenten te ontwikkelen die passend zijn in een bepaalde context. Het overzicht van criteria biedt zicht op mogelijke beoordelingsdimensies bij dans die in nieuwe instrumenten een rol kunnen spelen. Zoals geconstateerd is niet elk instrument grondig onderzocht. Meer onderzoek zou zich kunnen richten op het testen in de praktijk van beoordelingsinstrumenten en evaluatie op validiteit en betrouwbaarheid.

De problematiek bij beoordeling in dans is vergelijkbaar met die in de andere kunsten. De beschreven instrumenten zijn daarom ook interessant voor de andere kunsten. Veel van de beschreven publicaties zijn via het LKCA toegankelijk en meer gegevens over de instrumenten zijn via de website inzichtelijk: [www.lkca.nl/publicaties/informatiebank](http://www.lkca.nl/publicaties/informatiebank)

## Literatuur

- Bonbright, J., & Mcgreevy-Nichols, S.(1999). NAEP and dance: On contextual data, process, and problems in dance assessments, and recommendations for the field. *Arts Education Policy Review*, 100(6), 27-32. doi:10.1080/10632919909605995
- Bonbright, J., & Mcgreevy-nichols, S.(2012). National Dance Education Association: Building a future for dance education in the arts. *Arts Education Policy Review*, 113 (4), pp. 147-151. doi: 10.1080/10632913.2012.719430
- Englebright, K., & Mahoney, M. R. (2012). Assessment in elementary dance education. *Journal of Dance Education*, 12(3), 87-92. doi:10.1080/15290824.2012.701176
- Ericson, G. (1996). *Assessment of schoolchildren's performance in dance*. Uppsala University. Dissertation.
- Giguere, M. (2012). Self- reflective journaling: A tool for assessment. *Journal of Dance Education*, 12(3), 99-103. doi:10.1080/15290824.2012.701168
- Groenendijk, T, Damen, M.L., Haanstra, F., & Van Boxtel, C. (ingediend). *Beoordelingsinstrumenten in de kunstvakken – een review*.
- Gullikers, J. (2011). Authentiek beoordelen: kansen voor het kunstonderwijs. *Cultuur+Educatie* 31, pp. 38- 61. Utrecht: Cultuurnetwerk.
- 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
- Hoeven, M. van der et al. (2014). *Cultuur in de Spiegel in de praktijk. Een leerplankader voor cultuuronderwijs*. Enschede: SLO.
- Kranicke, M., & Pruitt, L. (2012). Are they getting it?: Creating dance assessments that honor the creative process. *Journal of Dance Education*, 12(3), 113-117. doi:10.1080/15290824.2012.701171
- Oliver, W., & Sprague, M. (2007). Dance proficiency in Rhode Island: Opportunities and challenges. *Journal of Dance Education*, 7(3), 72-79. doi: 10.1080/15290824.2007.10387341
- 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.
- 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.
- Ross, J.(1994). The right moves: Challenges of dance assessment. *Arts Education Policy Review*, 96(1), 11.
- 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

Slettum, B.M., Fox, C., Looney, M.A., & Jay, D.M. (2001). Validity and reliability of a folk dance performance checklist for children. *Measurement in Physical Education and Exercise Science*, 5(1), 35-55. doi: 10.1207/s15327841MPPEE0501\_3

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.

Warburton, E. (2000).The dance on paper: The effect of notation use on learning and development in dance. *Research in Dance Education*, 1(2), 193-213. doi: 10.1080/713694267

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

Wong, A. et al. (2012). Questionnaire of executive function for dancers: An ecological approach. *Assessment* 19(3), 383-387. doi: 10.1177/1073191110397483