

# Beoordeling van beeldende producten in het primair onderwijs

F. Haanstra, M. Damen en M. van Hoorn

## Samenvatting

Onderzoek wijst uit dat Westers georiënteerde kunstenaardocenten tekeningen van 5-jarigen vaak even hoog waarderen als tekeningen van volwassen kunstenaars en de meer conventionele beeldende uitingen van oudere kinderen lager waarderen. Men spreekt van een U-vormig beoordelingspatroon. Leken op het gebied van beeldende kunst komen in het algemeen tot een ander beoordelingspatroon, namelijk een stijgende waardering bij een stijgende leeftijd van de makers. Deze gegevens vormen het referentiekader voor een onderzoek naar beoordelingspatronen van betrokkenen bij beeldend onderwijs in het primair onderwijs. Groepsleerkrachten en beroeps-kunstenaars die projecten in het basisonderwijs uitvoeren, beoordeelden tekeningen van 5-, 8-, 11- en 14-jarigen, volwassen leken en volwassen beeldende kunstenaars op de criteria expressie, techniek, mooi, interessant en goed/slecht. De beroepskunstenaars komen tot meer U-vormige oordelen dan de groepsleerkrachten, maar de groepsleerkrachten beoordeelden de tekeningen van 5-jarigen hoger dan eerder onderzochte leken zonder beeldende opleiding. De beoordelingspatronen van groepsleerkrachten en beroepskunstenaars verschillen sterk van de beoordelingspatronen van scholieren. De resultaten worden bediscussieerd in relatie tot een meer transparant beoordelingssysteem.

## 1 Inleiding

In het basisonderwijs is beeldende vorming in het curriculum een onderdeel van kunstzinnige oriëntatie. De producten die leerlingen maken zijn wel gekarakteriseerd als 'schoolkunst' (Efland, 1976): een vorm van kunst die functioneel is binnen het instituut school, maar los staat van ontwikkelingen in de kun-

sten en vaak ook los van de populaire beeldcultuur die leerlingen buiten de school omringt. Het produceren van schoolkunst mag niet te tijdrovend zijn. Het is daarom uitgevoerd met makkelijk toepasbare media (krijt, plakkaatverf), die goed en snel zijn op te ruimen. Het gaat in deze kunst eerder om motorische dan intellectuele activiteit. Het moet er enigszins spontaan uitzien (dus natekenen is niet toegestaan), maar de inhoud is conventioneel. De beelden doen vaak dienst bij vieringen, feesten, de wisseling van seizoenen en dergelijke. Schoolkunst is 'supposed to be easy and fun' (Efland, p. 41). Hoewel de analyse van Efland al een kleine vier decennia geleden is gemaakt, zijn veel van de door hem genoemde kenmerken van schoolkunst nog van toepassing (Haanstra, 2001). Volgens de Onderwijsraad en de Raad voor Cultuur (2012) ontbreekt het bij kunstzinnige oriëntatie aan opbrengstgericht werken. Er is volgens de raden een meer systematische beoordeling van opgedane kennis, vaardigheden en attitudes nodig, alsmede een beoordeling van het werk van leerlingen in relatie tot de gestelde doelen op schoolniveau. Uit de Monitor cultuureducatie (Oomen e.a., 2009) blijkt dat in de onderbouw op driekwart van de basisscholen de kunstzinnige producten van leerlingen niet worden beoordeeld, in de bovenbouw geldt dat voor 55% van de scholen. Als er wordt beoordeeld, dan blijken creativiteit en werkhouding belangrijker te zijn dan technische vaardigheid.

Verskillende redenen worden aangevoerd voor deze beoordelingsproblematiek. In de eerste plaats de lage status van kunstzinnige oriëntatie en het vrijblijvende karakter ervan. In de tweede plaats wordt gebrek aan deskundigheid van de leerkrachten genoemd. Op de meeste basisscholen is de groepsleerkracht verantwoordelijk voor de kwaliteit van de beeldende lessen, terwijl ze niet specifiek geschoold zijn op dat gebied. In veel gevallen

wordt er daarom externe ondersteuning gezocht. Op ongeveer tweederde van de 7000 basisscholen werken beeldende kunstenaars in projecten samen met leerlingen (Oomen e.a., 2009), meestal de zogeheten beroepskunstenaars in de klas (BIKkers). Een derde reden is de opvatting dat kunstzinnige processen en producten verbonden zijn met persoonlijke expressie en individuele kwaliteiten, terwijl beoordelingen juist objectiviteit en vergelijkbaarheid veronderstellen. Die opvatting heeft een lange geschiedenis. In de kindgerichte benadering die in de jaren '50 en '60 van de vorige eeuw in veel westerse landen de kunsteducatie domineerde, stond de creatieve zelfexpressie van kinderen voorop (Asselbergs, 1989). Leerplannen en formele beoordeling van de leerlingenprestaties werden afgewezen. Zo stelden Löwenfeld en Brittain (1975) in hun invloedrijke boek 'Creative and Mental Growth' dat kunst het gebied in het onderwijs vertegenwoordigt waar kinderen zich geen zorgen hoeven te maken over beoordeling en evaluatie. Maar niet beoordelen betekende niet dat er geen beïnvloeding plaats vond. Wilson (2004) laat zien dat door hun instructies, opdrachten en de keuze van materialen de docenten de vorm en inhoud van de 'vrije' kinderkunst in grote mate bepaalden.

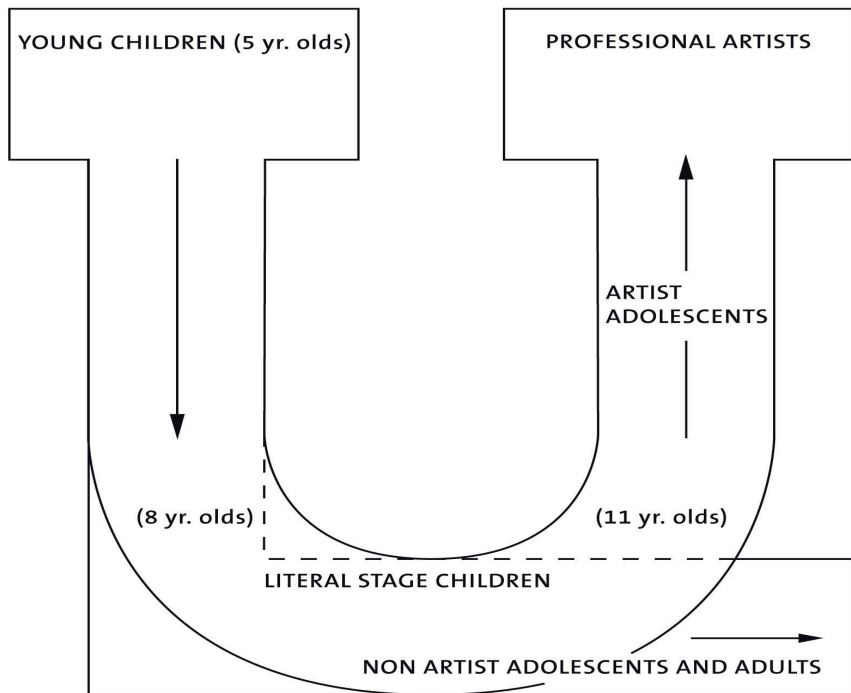
Dit artikel beschrijft een onderzoek naar het beoordelen van beeldende producten door leerkrachten en kunstenaars in het basisonderwijs in relatie tot leeftijden en ontwikkelingsniveaus van de makers. Het onderzoek past in een reeks van onderzoeken die uitgaan van een theorie over de beeldende ontwikkeling van Gardner en Winner (1982). We bespreken eerst deze theorie en het eerdere onderzoek. Vervolgens behandelen we een nieuw onderzoek, met als beoordelaars groepsleerkrachten in het basisonderwijs en beroepskunstenaars die projecten in het basisonderwijs uitvoeren.

## 2 De theorie over de U-vormige beeldende ontwikkeling

Gardner en Winner (1982) onderscheiden verschillende ontwikkelingsstadia in de beel-

dende ontwikkeling van kinderen. Zij stellen dat er geen sprake is van een progressieve ontwikkeling, maar dat de beeldende uitingen van het jonge kind van een hoger niveau zijn dan van oudere kinderen. Kinderen komen na de expressieve kleuterfase in een letterlijke of realistische fase. Ze willen of moeten zich op beeldend gebied conventioneel en realistisch uitdrukken ('net echt'), maar de technische vaardigheden om dat bevredigend te doen schieten tekort. Voor velen wordt de discrepantie tussen wat ze beeldend zouden willen uiten en waartoe ze in staat zijn te groot en ze raken teleurgesteld in het eigen beeldende kunnen. Zo blijft bij veel mensen de beeldende ontwikkeling op het niveau van een tien- of elfjarige steken en neemt de beeldende activiteit sterk af. Als we het verloop van die ontwikkeling grafisch verbeelden, zien we bij hen een L-vormige ontwikkeling. De L kan bovendien staan voor de L van Letterlijkheid. Een kleine groep adolescenten blijft beeldend actief én slaagt erin om weer vooruitgang in hun beeldende uitingen te boeken. Bij hen kan men spreken van een U-vormige ontwikkeling (zie figuur 1) en die U vorm geeft aan dat de beeldende uitingen van jonge kinderen op bepaalde punten kwalitatief vergelijkbaar zijn met die van professionele kunstenaars. In figuur 1 is de leeftijd weergegeven op de x-as en de beoordelingen op de y-as: de beeldende producten van makers op 5-jarige leeftijd worden hoger beoordeeld dan de producten van makers op 8- en 11 jarige leeftijd. Voor de niet-kunstenaar adolescenten en volwassenen blijft de beoordeling op latere leeftijd op dat lage niveau (L-vormige ontwikkeling). Voor de kunstenaar adolescenten en volwassenen stijgt de beoordeling weer op latere leeftijd (U-vormige ontwikkeling).

Gardner en Winner onderbouwen dit met wat Goodman (1984, p.135 e.v.) de 'esthetische symptomen' van kunst noemt. Kunstwerken hebben een symbolische functie, maar dat hebben ze gemeen met veel andere objecten, zoals verkeersborden. Er zijn echter bepaalde eigenschappen die van een alledaags object een kunstwerk maken. Eén van die esthetische eigenschappen is expressie. Dit houdt in



*Figuur 1.* Het U- en L-vormige model van artistieke ontwikkeling (Davis, 1991).

dat door bepaald gebruik van kleuren, lijnen, compositie en abstractie een beeldend werk op een metaforische manier een emotie of gedachte kan uitdrukken. Naast expressie onderscheidt Goodman denotatie: de symbolische verwijzing naar de werkelijkheid door het afbeelden van personen, objecten, gebeurtenissen enzovoort. Beeldende kunst kan zowel van expressie als van denotatie gebruikmaken, maar expressie maakt een beeldende product eerder tot een esthetisch object (kunst) dan wanneer uitsluitend sprake is van denotatie. De tekeningen van kinderen in de letterlijke fase maken vrijwel alleen gebruik van denotatie, terwijl de tekeningen van jongere kinderen (en van professionele kunstenaars) ook worden gekenmerkt door expressie.

Van verschillende kanten werd al snel kritiek geuit op deze theorie (Wilson, 1981; Duncum, 1986). Volgens critici is de geconstateerde overeenkomst tussen het werk van jonge kinderen en beroepskunstenaars tijdsgebonden. Ze geldt voor het modernisme in

de westerse kunst, met een sterke nadruk op formele en expressieve elementen. Ze geldt in het bijzonder voor het abstract expressionisme (bijvoorbeeld de Cobra kunstenaars), maar niet of nauwelijks voor de kunst daarvoor en ook in steeds mindere mate voor de kunst daarna. Oordelen vanuit een andere culturele achtergrond of andere stijloppvattingen zouden tot andere conclusies over de (ontwikkeling van) beeldende producten leiden.

### **2.1 Empirische bevestiging**

De U-vormige ontwikkeling werd voor het eerst empirisch getoetst in het dissertatieonderzoek van Davis (1991). Zij liet verschillende makers drie tekeningen maken die de emoties woede, blijdschap en verdriet moesten uitdrukken. Davis onderscheidde de makers in zeven leeftijds- en expertisegroepen, namelijk vijf-, acht-, elf- en veertienjarigen en volwassenen, waarbij de veertienjarigen en de volwassenen waren verdeeld in degenen die zichzelf als kunstenaars beschouwden en zij die dat niet deden

(‘leken’). Per groep werden twintig personen gekozen, dus in totaal 140 personen. In totaal 420 tekeningen werden beoordeeld door twee experts op vier ‘esthetische dimensies’.

De uitkomsten van het onderzoek bevestigden de belangrijkste hypothesen: tekeningen van jonge kinderen, artistieke adolescenten en professionele kunstenaars werden gemiddeld significant hoger gewaardeerd dan tekeningen van acht- en elfjarigen en van adolescenten en volwassenen zonder beeldende activiteiten. Vooral de algehele expressiviteit van de tekeningen (‘overall expression’) van de vijfjarigen kwam het sterkst overeen met die van de kunstenaars. In bijna alle gevallen kregen de tekeningen van de professionele kunstenaars de hoogste scores. Ook als de beoordelingen van de verschillende leeftijdsgroepen op de vier dimensies worden samengenomen, ontstaat een U-curve met aan de uiteinden de hoogste gemiddelde scores. De achteruitgang in de esthetisch-expressieve kwaliteit van beeldende producten, maar ook de frequentie van beeldende activiteiten (kinderen gaan minder tekenen en schilderen) na de kleuterleeftijd verklaart Davis ten dele door de invloed van de school die prioriteiten gaat stellen (namelijk taal en rekenen).

## **2.2 Empirische verwerping**

Pariser en Van den Berg (1997) repliceerden Davis’ onderzoek, maar betrokken daarbij ook beoordelaars met een niet-westerse achtergrond om de eerder geuite kritiek op het vermeende universele karakter van de U-bocht te toetsen. Ze lieten 165 tekeningen van 55 personen van verschillende leeftijdsgroepen beoordelen door twee personen van Chinese herkomst en twee Noord-Amerikaanse beoordelaars. Het ging weer om dezelfde onderwerpen (verdriet, blijdschap en woede), dezelfde zeven leeftijds- en expertisegroepen en dezelfde vier beoordelingscriteria. Daarbij gaven ze de beoordelaars een minder bewerkelijke beoordelingstaak. Deze moesten de tekeningen over drie stapels verdelen: goede, redelijke en slechte tekeningen. Het gemiddelde resultaat over de vier esthetische beoordelingscriteria week weinig af van het resultaat van deze driedeling in tekeningen.

Zij vonden voor de twee Noord-Amerikaanse experts dezelfde resultaten als Davis (1991). Bij de Chinese expertoordelen vonden ze geen bevestiging van de U-vormige ontwikkeling. Globaal genomen beoordeelden deze de tekeningen van de vijfjarigen als laagste en steeg hun waardering enigszins bij de hogere leeftijdsgroepen. Hun waardering voor het werk van de volwassenen (zowel leken als professionals) was veruit het hoogst. De onderzoekers vermoeden dat deze verschillen in expertoordelen komen doordat de Chinese beeldende kunst niet de modernistische, deels abstract expressionistische traditie kent en ambachtelijkheid en realisme hier in hoger aanzien staan dan in de westerse moderne kunst.

In een vervolgstudie (Pariser & Van den Berg, 2001) lieten de onderzoekers de tekeningen uit het eerdere onderzoek door andere beoordelaars beoordelen op techniek. Vervolgens werden de samenhangen berekend van deze oordelen met de scores voor esthetische waardering van de Noord-Amerikaanse en Chinese experts. De esthetische oordelen van de Chinese experts bleken inderdaad veel sterker samen te hangen met de onafhankelijk verkregen oordelen over technische vaardigheid.

In het daaropvolgende onderzoek naar de U-vormige ontwikkeling onderzocht Kindler (2000) tekeningen van Taiwanese kinderen, adolescenten en volwassenen, zoals beoordeeld door Taiwanese beoordelaars. Het onderzoek was geen exacte replicatie van het onderzoek van Davis (1991), want de tekenopdrachten waren anders (namelijk een cartoon en een portret) en het leek Kindler van belang voor het beeldende onderwijs om oordelen van experts (docenten beeldende vorming) te vergelijken met de oordelen van de direct belanghebbenden: de leerlingen. Vandaar dat ook 8- 11- en 14-jarigen als beoordelaars aan het onderzoek deelnamen. De in totaal acht beoordelaars moesten de tekeningen in drie stapels leggen: goed, redelijk en slecht. De uitkomst hiervan was dat de meeste beoordelaars geen terugval in kwaliteit zagen na de kleuterleeftijd, maar juist een steeds stijgende kwaliteit. Kinderen en volwassenen gaven daarvoor wel verschillende

redenen aan. Verder gaven de kinderen de voorkeur aan de cartoons, terwijl de volwassen beoordelaars die juist lager beoordeelden.

In de discussie van hun replicatie onderzoek stellen Pariser en Van den Berg (1997) zichzelf de vraag wat nu eigenlijk het onderwerp van hun studie is. Er zijn immers twee afhankelijke variabelen: (1) de ontwikkeling van de esthetische kwaliteit van beeldende producten en (2) de esthetische beoordeling door beoordelaars met verschillende culturele achtergronden. Door hun resultaten is de nadruk in het onderzoek op de tweede variabele komen te liggen. Dit soort onderzoek kan daarom niet volstaan met slechts enkele beoordelaars die geacht worden de kwaliteit van de tekeningen vast te stellen. Als verschillen tussen oordelen van beoordelaars zelf mede onderwerp van onderzoek is, moeten ook grotere aantallen beoordelaars aan het onderzoek deelnemen, om toevallige verschillen tussen beoordelaars zoveel mogelijk te neutraliseren.

In de studie van Pariser, Kindler, Van den Berg, Dias & Chen Liu (2007) is voor het eerst een opzet gekozen met grote steekproeven beoordelaars. En conform de veranderde focus stellen ze: 'Our findings shed light on the judges' assumptions rather than on the 'true nature' of graphic development' (p. 79). De onderzoekers lieten tekeningen maken door 360 kinderen en volwassenen in zes groepen: de leeftijdsgroepen vijf, acht, elf en veertien jaar, volwassen leken en professionele kunstenaars. Alle tekenaars maakten eerst een tekening van een zelf gekozen onderwerp. Daarna maakte de ene helft de klassieke opdrachten van Davis: teken verdriet, blijdschap en woede. De andere helft maakte tekeningen met als opdrachten: de route van huis naar school of werk, een favoriete cartoonfiguur en een tekening die een verhaal vertelt waarin mensen voorkomen. De tekeningen werden beoordeeld door 192 beoordelaars. De beoordelaars waren andere personen dan de makers en waren verdeeld over achtjarigen, veertienjarigen, volwassen leken en volwassen experts op beeldend gebied. De belangrijkste uitkomst is dat een ruime meerderheid van alle beoordelaars een 'traditionalistisch' beoordelingspatroon heeft.

Voor hen geldt: oefening baart kunst en die kunst wordt vooral beoordeeld op technisch kunnen. Bij de jongere beoordelaars en de volwassen leken komt de modernistische U-curve, met hoogste waardering voor tekeningen van kleuters en kunstenaars nauwelijks voor. Bij de expertbeoordelaars is dat wel het geval, maar dan vooral bij de Canadese experts.

Onderzoek in Nederland van Haanstra, Van Hoorn en Damen (2011) is een combinatie van de oorspronkelijke aanpak van Davis en de latere aanpak van onderzoekers als Kindler en Pariser. De tekeningen van vijf-, acht-, elf- en veertienjarigen en volwassenen (leken en beeldende professionals) betroffen twee tekenopdrachten: zelfportret en verdriet. De 287 beoordelaars bestonden uit experts (beeldend kunstenaars en docenten beeldende kunsten), volwassen leken op beeldend gebied en kinderen. Ze hadden een Nederlandse, Turkse of Marokkaanse achtergrond. Uit de resultaten blijkt dat het U-vormige beoordelingspatroon zich niet of nauwelijks voordoet wanneer kinderen en volwassen leken de tekeningen beoordelen. De culturele achtergrond van de beoordelaars maakt hierbij geen verschil. Het U-vormige beoordelingspatroon doet zich wel voor bij de expertbeoordelaars. Culturele achtergrond maakt hierbij een verschil. De U-vorm is niet aanwezig in de oordelen van experts met een Marokkaanse achtergrond, maar wel bij de beoordelaars met een Nederlandse of Turkse achtergrond. De experts hebben de tekeningen niet alleen algemeen beoordeeld op goed, middelmatig of slecht, maar ook op techniek en expressiviteit. Bij technische vaardigheid deed zich geen U-vormig beoordelingspatroon voor, maar vonden de experts (ongeacht culturele achtergrond) dat die vaardigheid enigszins toenam in de leeftijd 5 tot 14-jarigen en ze werd pas bij de professionele kunstenaars echt hoog gewaardeerd. Het criterium expressie vertoont bij de beoordelaars met Turkse en Nederlandse achtergrond een duidelijke U-curve. Dat geldt niet voor de beoordelaars met een Marokkaanse achtergrond; daar verschillen de gemiddeld lage oordelen van de vijf-, acht- en elfjarigen niet significant van elkaar.

Tenslotte is een onderzoek gedaan waarbij de oordelen van 117 Amerikaanse psychologie studenten en 38 kunststudenten zijn vergeleken (Haanstra, Damen, & Van Hoorn, 2013). De oordelen betroffen weer zelfportretten en tekeningen over verdriet die gemaakt zijn door vijf-, acht-, elf- en veertienjarigen en volwassenen (leken en beeldende professionals). Naast de driedeling in goed, middelmatig en slecht werd een oordeel op twee andere criteria gevraagd: interessant en mooi. Bij de kunststudenten laten de gemiddelde beoordelingspatronen bij de criteria interessant en mooi een U-vormig patroon zien. Bij het criterium goed is het patroon minder eenduidig. Bij de psychologiestudenten is bij geen van de criteria gemiddeld een U-vormig patroon zichtbaar.

Samenvattend kan worden gesteld dat alleen het empirisch onderzoek van Davis de theorie van de U- of L-vormige ontwikkeling in beeldende productie ondersteunt, maar dat er meer onderzoeken zijn die de oorspronkelijke theorie niet ondersteunen en laten zien dat de U-curve een cultureel bepaalde, modernistische voorkeur is van experts op beeldend gebied (kunstenaars, kunstenaardocenten en kunststudenten). De latere onderzoeken waren geen exacte replicatieonderzoeken van het oorspronkelijke onderzoek van Davis, want onderzoekers brachten bewust wijzigingen aan in de tekenopdrachten, in de gehanteerde beoordelingscriteria en in de leeftijd, de expertise en culturele achtergrond van beoordelaars en makers. In de woorden van Pariser (2006) vertegenwoordigt de U-vormige beoordelingscurve 'a specialized and somewhat rarified view of aesthetic development, one associated exclusively with training in the arts.' (p.268).

### 2.3 Onderzoeksvraag

Gezien het beleidsstreven naar een meer systematische beoordelingssystematiek van kunstzinnige oriëntatie is ook nader onderzoek naar beoordeling van beeldende producten door leerkrachten in het primair onderwijs gewenst. Die problematiek valt op zeer uiteenlopende manieren te onderzoeken. In het voorliggende onderzoek hebben we het model van de U-vormige beeldende ontwikkeling

als uitgangspunt genomen. Door het onderzoek dat hiernaar al is verricht, heeft het huidige onderzoek een duidelijk vergelijkingskader. Het dient daarmee twee doelen: ten eerste om bij te dragen aan de internationale discussie over de theorie van de U-vormige beeldende ontwikkeling en ten tweede om bij te dragen aan de discussie over beoordeling van beeldende producten op Nederlandse basisscholen. Meer inzicht in dat laatste kan worden verkregen door beeldende producten te laten beoordelen door groepsleerkrachten en deze oordelen te vergelijken met die van zogeheten BIKkers, Beroepskunstenaars In de Klas. De beoordelingspatronen van BIKkers zijn naar verwachting vergelijkbaar met die van de beeldende experts uit voorgaand onderzoek. De vraag is of de beoordelingspatronen van de groepsleerkrachten vergelijkbaar zijn met die van volwassen leken of meer met die van beeldende experts. Bovendien bouwt het onderzoek voort op resultaten over het gebruik van verschillende beoordelingscriteria, zoals expressie versus techniek.

De hoofdvraag van het onderzoek luidt:

Versillen groepsleerkrachten en beroepskunstenaars in hun beoordeling van tekeningen van makers van verschillende leeftijden op de vijf criteria expressie, techniek, mooi, interessant en goed/slecht?

## 3 Opzet onderzoek

### 3.1 De beoordelaars

De beoordelaars zijn 73 groepsleerkrachten in het primair onderwijs en 71 kunstenaars die een opleiding Beroepskunstenaar In de Klas hebben gevolgd. Deze post-hbo-opleiding is bedoeld voor professionele kunstenaars die projecten willen ontwikkelen en uitvoeren in het basisonderwijs. In de Appendix is de steekproeftrekking beschreven. Van de groepsleerkrachten is 90% vrouw en de gemiddelde leeftijd is 45. Van de BIKkers is 80% vrouw en de gemiddelde leeftijd 41. De verschillen in sekseverdeling zijn niet significant (Cramer's  $V = .135$ ). Het verschil in gemiddelde leeftijd tussen groepsleerkrachten en beroepskunstenaars is wel significant

(T-waarde = 2.6). Er is echter geen reden om aan te nemen dat de beoordelingen verschillend zullen zijn voor jongere of oudere beoordelaars.

De deelname aan kunst en cultuur is gemeten door te vragen naar bezoekfrequentie van de afgelopen twaalf maanden aan negen culturele activiteiten, zoals toneel- en dansvoorstellingen, klassieke en popconcerten, museum, bioscoop etc. (Damen, 2010). De beroepskunstenaars scoorden significant hoger op cultuurdeelname dan de groepsleerkrachten (T-waarde = 2.5 (df 118),  $p < .05$ ). Dit betekent naast beroepsopleiding een extra indicatie voor het verschil in expertise op het gebied van de kunsten.

De respondenten is gevraagd naar de frequentie en de manieren van beoordelen van het beeldende werk van leerlingen. Van de groepsleerkrachten beoordeelt 39% het beeldende werk vaak, 54% soms en 7% nooit. Van de BIKkers zegt 31% vaak te beoordelen, 54% soms en 15% nooit. De BIKkers die beoordelen, doen dat vaak met een groeps gesprek en soms met een individueel gesprek. Geschreven beoordelingen en cijfers geven ze niet tot nauwelijks. Dat komt mogelijk omdat ze alleen kortlopende projecten verzorgen en niet voor langere tijd aan een groep zijn verbonden. Groepsleerkrachten houden wat vaker individuele gesprekken, een derde geeft soms geschreven beoordelingen en cijfers en 16 respectievelijk 23% doet dat vaak. Maar bijna de helft van de groepsleerkrachten doet dat helemaal niet.

De respondenten is ook gevraagd naar het relatieve belang dat ze hechten aan enkele beoordelingscriteria, zoals creativiteit, aantrekkelijkheid eindproduct, technische kwaliteit en criteria die op het maakproces slaan zoals werkhouding en reflectievermogen. Er blijkt weinig verschil van mening tussen groepsleerkrachten en BIKkers. Creativiteit en werkproces scoren het hoogst: 80% van de respondenten vond die van veel belang. Techniek scoort het laagst: 8% van de BIKkers en 18% van de groepsleerkrachten acht dat van veel belang. Die rangorde komt goed overeen met eerdergenoemde resultaten van de monitor cultuureducatie (Oomen e.a., 2009).

### 3.2 De tekeningen

Het onderzoek gebruikt dezelfde stimuli als in de eerdere Nederlandse onderzoeken (Haanstra, Van Hoorn & Damen, 2011; Haanstra, Damen & Van Hoorn, 2013): een collectie tekeningen afkomstig van zes groepen, namelijk van vijf-, acht-, elf- en veertienjarigen en volwassenen. De volwassenen zijn onderverdeeld in leken en professionals. Professionals zijn personen die een beeldende opleiding hebben (autonoom, toegepast of docentschap), dan wel werkzaam zijn als beeldend kunstenaar of docent beeldende vorming. Elke deelnemer maakte twee tekeningen. De eerste opdracht luidde: *'Tekening verdrietig. Het gaat er dus om een tekening te maken van verdriet'*. Deze opdracht is gekozen omdat hij overeenkomt met een opdracht uit het oorspronkelijke onderzoek van Davis (1991) en ook in replicatie onderzoeken is gebruikt. Omdat de materiële restricties hierbij groot zijn (zwarte viltstift op A4 formaat) is daarnaast gekozen voor een opdracht met eigen materiaal- en kleurkeuze, maar wel een die vaak in het onderwijs wordt gegeven, namelijk een zelfportret: *'Tekening of schilder een zelfportret: een portret van jezelf zoals jij vindt dat je eruit ziet. Je kunt kiezen voor alleen je hoofd, een deel van je lichaam erbij of jezelf helemaal ten voeten uit. Je kunt gebruik maken van een spiegel.'*

In totaal hebben 162 personen 324 tekeningen gemaakt. Uit eerder onderzoek (Haanstra, Van Hoorn & Damen, 2011) bleek dat dit aantal voor een individuele beoordelaar te groot is om in een onderzoek te beoordelen. Daarom is een gestratificeerde steekproef getrokken van 48 tekeningen: 8 tekeningen (vier 'verdriet tekeningen' en vier zelfportretten) voor alle zes leeftijds/expertise groepen.

### 3.3 Beoordelingsprocedure en -criteria

De tekeningen zijn gescand en op een website gezet. In vergelijking met de originelen is het verlies aan beeldkwaliteit van de digitale afbeeldingen zeer gering. Het bestand met afbeeldingen was bereikbaar via een toegangscode. Beoordelaars konden zo op een door hen gewenste plek en tijd op een computer de beoordelingsopdrachten uitvoeren. Ze kregen een korte uitleg van de beoorde-

lingstaak, gevolgd door een aantal te verwachten vragen met antwoorden ('FAQ's'), namelijk welke instructie de tekenaars hadden gekregen, hoeveel tijd de beoordeling ongeveer zou duren en of men beoordelingen tussentijds nog kon wijzigen. Dat laatste was inderdaad het geval. Over de verschillen in expertise van de makers van de tekeningen werd niets gezegd, alleen dat de leeftijden van de makers sterk verschilden.

De beoordelaars werd gevraagd de tekeningen volgens vijf criteria te beoordelen, namelijk expressie, techniek, mooi, interessant en goed/slecht. Het ging om vijfpuntschalen (aangegeven met --, -, ±, +, ++). Het oordeel over de mate van expressie is ontleend aan Davis' (1991) criterium 'overall expression'. Zij baseerde zich daarbij op Gardner en Winner (1982), die expressie relevant achten als criterium dat de esthetische waarde van beeldende producten bepaalt. Dit criterium is direct verbonden met het modernistische, abstract expressionistische idioom en zou daarom bij westerse beoordelaars het best de U-vormige ontwikkeling moeten genereren.

Het criterium techniek of technische vaardigheid is een traditioneler en conventioneel criterium en verwijst naar beheersing van het medium. In het genoemde onderzoek van Pariser en Van den Berg (1997) bleek dat het criterium technische vaardigheid sterk samenhangt met de oordelen van de Chinese experts.

De criteria mooi en interessant verwijzen naar een belangrijk onderscheid in de literatuur over esthetische beoordeling (Berlyne, 1974; Hekkert, 1995; Leder, Belke, Oeberst & Augustin, 2004), namelijk tussen affectieve voorkeur enerzijds en beredeneerd oordeel anderzijds. Aantrekkelijkheid of schoonheid zou de onbereflecteerde persoonlijke voorkeur weergeven en interessant de weloverwogen beoordeling van de kwaliteit.

Het criterium goed/slecht tenslotte wil een algemeen kwaliteitsoordeel van de beeldende producten meten. Het is gebruikt in de onderzoeken van Kindler (2000), Pariser e.a. (2007) en in de vorige onderzoeken van Haanstra, Van Hoorn en Damen (2011) en Haanstra, Damen en Van Hoorn (2013). Het oordeel op dit criterium wordt daarom

gebruikt om de resultaten van dit onderzoek te vergelijken met eerdere onderzoeken.

In eerder onderzoek (Haanstra, Damen & Van Hoorn, 2013) is vergeleken of het beoordelen van tekeningen in afzonderlijke rondes per criterium tot andere resultaten leidt dan een beoordeling per tekening op meerdere criteria tegelijkertijd. Dat bleek niet het geval. Het huidige onderzoek is een tussenvorm: de beoordelaars moesten eerst 24 verdriettekeningen en daarna 24 zelfportretten tegelijkertijd beoordelen op de criteria expressie, techniek, mooi en interessant. Daarna moesten ze dezelfde tekeningen beoordelen op goed/slecht. De volgorde van tekeningen per beoordelingsronde is gerandomiseerd.

## 4 Resultaten

### 4.1 De beoordelingen van groepsleerkrachten en kunstenaars vergeleken

De gemiddelde beoordelingen van de groepsleerkrachten en van de beroepskunstenaars van tekeningen van verschillende leeftijdsgroepen op vijf criteria leiden tot tien beoordelingspatronen (zie figuren 2 tot en met 6). Op het oog zijn de meeste beoordelingspatronen van beide beoordelingsgroepen U-vormig, dat wil zeggen dat de gemiddelde oordelen over de tekeningen van de vijfjarigen veelal hoger zijn dan van de andere leeftijdsgroepen, behalve de oordelen over de tekeningen van de beroepskunstenaars. Een uitzondering vormen de beoordelingspatronen op het criterium techniek. Die zijn eerder lineair stijgend te noemen, dat wil zeggen dat met het stijgen van de leeftijd de technische vaardigheid toeneemt en bij de beroepskunstenaars veruit het grootst is. Een vergelijkbaar patroon bij het oordeel over de techniek van tekeningen werd gevonden in het onderzoek van Haanstra, Van Hoorn en Damen (2011)<sup>1</sup>.

Om de vraag naar verschillen in beoordeling tussen groepsleerkrachten en beroepskunstenaars te beantwoorden zijn de beoordelingspatronen per criterium (expressie, techniek, mooi, interessant en goed/slecht) geanalyseerd met mixed model ANOVA's. Omdat de 48 tekeningen van de zes verschil-



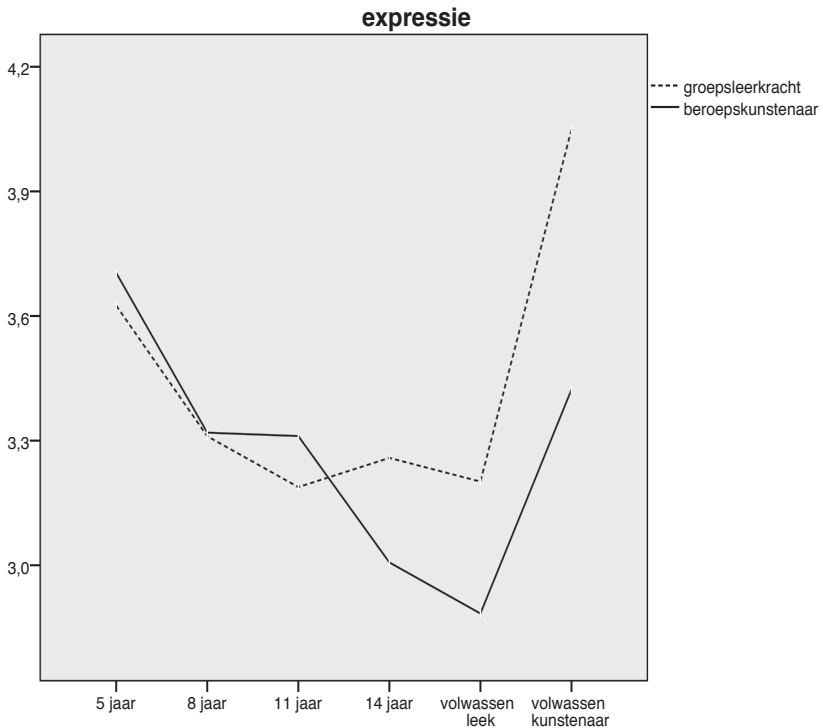
lende leeftijdsgroepen steeds door alle beoordelaars zijn beoordeeld is sprake van herhaalde metingen. De factor leeftijd van de tekenaars is derhalve een binnen-proefpersoon variabele met zes niveaus. De factor expertise is de tussen-proefpersoon variabele en heeft twee niveaus: leken (groepsleerkrachten) en experts (beroepskunstenaars).

Mauchly's test gaf steeds aan dat niet werd voldaan aan de aanname van sphericity (gelijkheid van variantie van de verschillen tussen de condities) voor de hoofdeffecten van de leeftijd van de tekenaars. Daarom werd de Greenhouse-Geisser correctie toegepast. Er waren significante hoofdeffecten van de leeftijd van de tekenaars voor alle vijf criteria: expressie,  $F(3.46, 501.75) = 77.00, p = .000$ ; techniek,  $F(2.20, 317.79) = 211.66, p = .000$ ; mooi,  $F(2.81, 407.69) = 53.63, p = .000$ ; interessant,  $F(3.38, 491.20) = 63.86, p = .000$  en goed/slecht  $F(2.27, 309.22) = 69.57, p = .000$ . Dit betekent dat er op alle

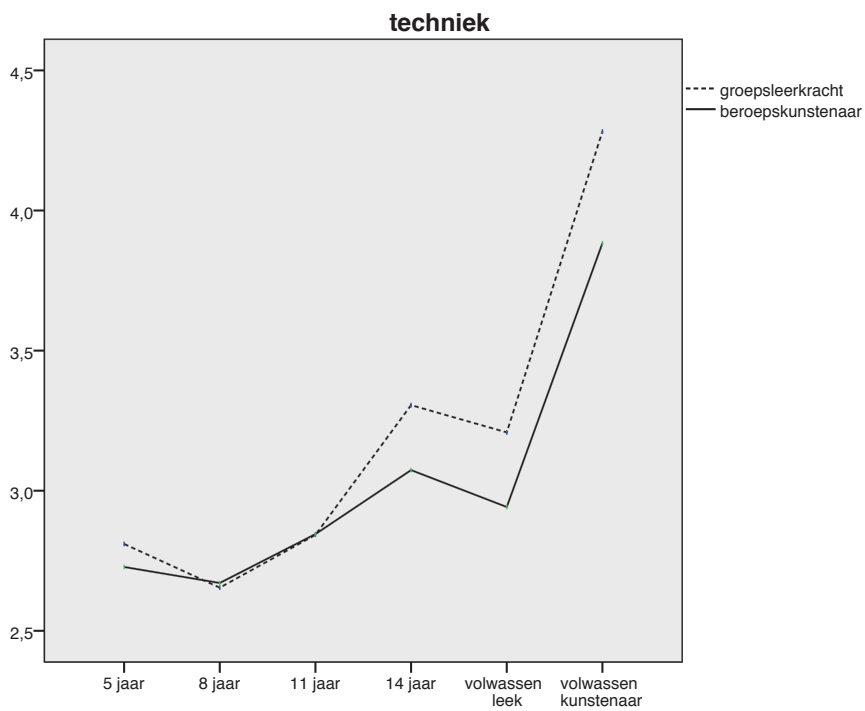
criteria significante verschillen bestaan tussen de oordelen over tekeningen van de 5-, 8-, 11- en 14-jarigen, de volwassen leken en de volwassen kunstenaars.

Er waren geen significante hoofdeffecten voor de factor expertise van de beoordelaars voor de criteria mooi  $F(1, 145) = 0, p = .990$ , interessant,  $F(1, 145) = 2.42, p = .122$  en goed/slecht  $F(1, 136) = .69, p = .406$ . Er zijn voor deze criteria dus geen verschillen in het gemiddelde oordeel van de groepsleerkrachten en de kunstenaar docenten. Die zijn er wel voor de criteria expressie  $F(1, 145) = 3.96, p = .048$  en techniek  $F(1, 144) = 4.11, p = .045$ .

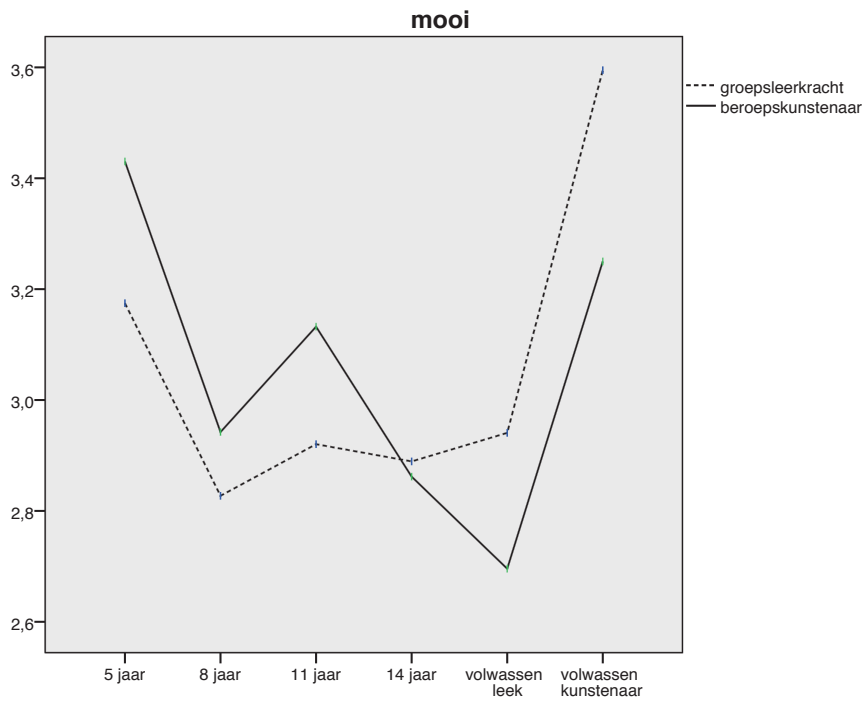
Tenslotte waren er significante interactie effecten tussen de factoren expertise van de beoordelaars en de leeftijd van de tekenaars voor alle vijf beoordelingscriteria: expressie  $F(3.46, 501.77) = 19.77, p = .000$ ; techniek,  $F(2.20, 317.79) = 5.07, p = .005$ ; mooi  $F(2.8, 407.69) = 12.09, p = .000$ ; interessant  $F(3.38,$



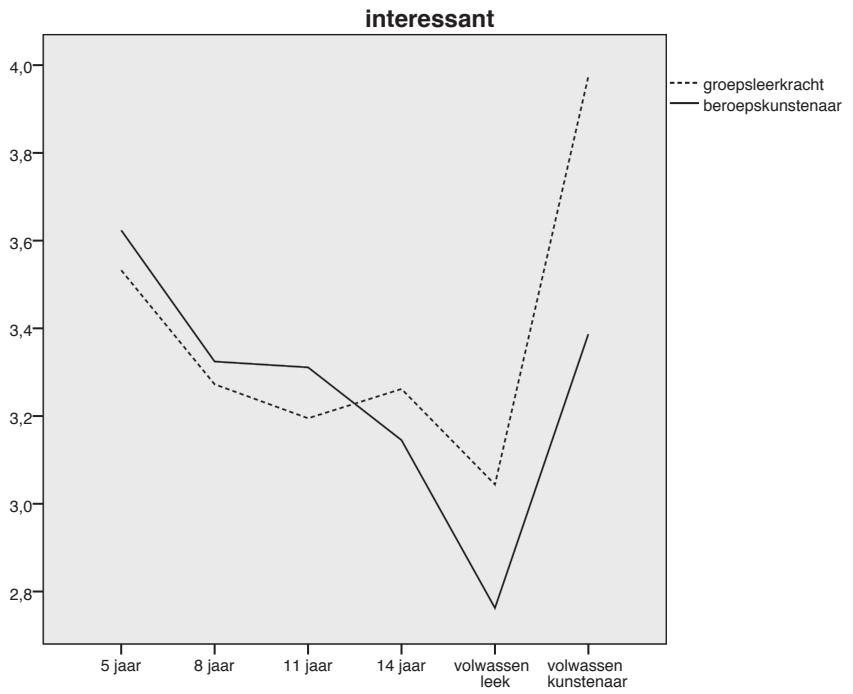
Figuur 2. Beoordelingspatronen van groepsleerkrachten en kunstenaar docenten voor het criterium expressie.



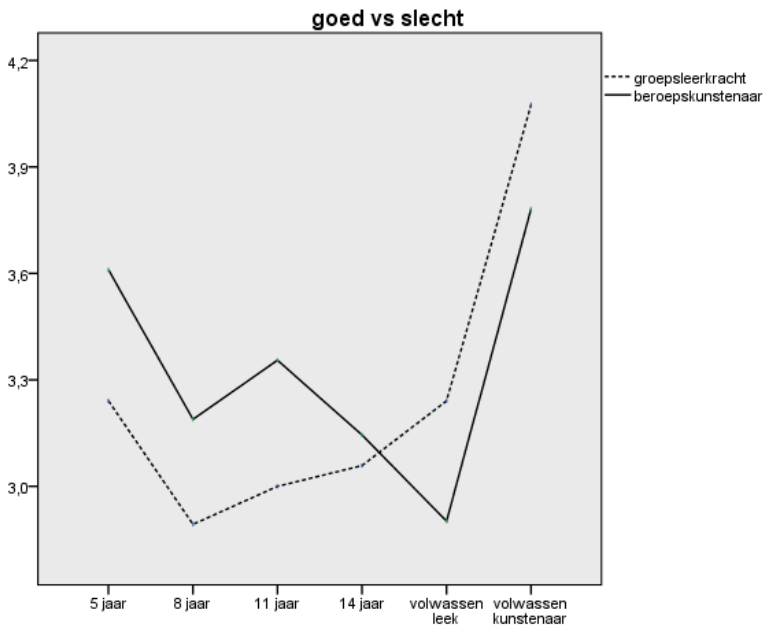
Figuur 3. Beoordelingspatronen van groepsleerkrachten en kunstenaar docenten voor het criterium techniek.



Figuur 4. Beoordelingspatronen van groepsleerkrachten en kunstenaar docenten voor het criterium mooi.



Figuur 5. Beoordelingspatronen van groepsleerkrachten en kunstenaar docenten voor het criterium interessant.



Figuur 6. Beoordelingspatronen van groepsleerkrachten en kunstenaar docenten voor het criterium goed/slecht.

491.20) = 15.11,  $p = .000$ ; goed/slecht  $F(2.27, 309.22) = 12.91, p = .000$ . Dit betekent dat de gemiddelde beoordelingen van de zes verschillende leeftijdsgroepen op de vijf criteria afhankelijk zijn van de mate van expertise van de beoordelaar: groepsleerkrachten of beroepskunstenaars.

#### **4.2 Is er sprake van U-vormige beoordelingspatronen?**

De vraag is nu in welke gevallen er sprake is van de door Davis voorspelde U-vormige patronen. Geldt die voorspelling, zoals het later verrichte onderzoek aangeeft (o.a. Pariser e.a., 2007) alleen voor de expert beoordelaars: de beroepskunstenaars? En geldt dit dan voor alle vijf criteria die in dit onderzoek zijn gebruikt?

We toetsen dit bij de tien beoordelingspatronen door middel van simpele contrasten van de within subjects factor (leeftijd/expertise makers) binnen de between subjects factor (groepsleerkrachten versus beroepskunstenaars). De oordelen over de tekeningen van vijfjarige makers zijn de referentiegroep (tabel 1). We beschouwen een beoordelingspatroon als een 'echte' U-vorm, wanneer de oordelen over tekeningen van vijfjarigen en van volwassen kunstenaars niet significant van elkaar verschillen, maar wanneer de oordelen over de tekeningen van de vijfjarigen wel significant hoger worden beoordeeld dan de oordelen over tekeningen van de andere leeftijdsgroepen: de 8-, 11-, en 14-jarigen en de volwassen leken.

Bij de beoordelingen van de beroepskunstenaars in de klas zijn - op de criteria mooi en goed/slecht - de gemiddelde scores voor de vijfjarigen significant hoger dan de gemiddelde scores voor alle andere leeftijdsgroepen, behalve voor de volwassen kunstenaar-makers. Deze beoordelingspatronen voldoen dus aan een belangrijk kenmerk van de U-vorm. Op de criteria expressie en interessant zijn de scores op de tekeningen van vijfjarigen ook significant hoger dan die van de andere leeftijden, maar bovendien ook significant hoger dan de scores van de tekeningen van de volwassen kunstenaars.

Bij de groepsleerkrachten zijn de gemid-

delde scores van de tekeningen van vijfjarigen significant hoger dan die van de 8-, 11- en 14-jarigen en de volwassen leken op de criteria mooi, expressie en interessant. Maar vergeleken met de scores van tekeningen van volwassen kunstenaars zijn ze significant lager. Bij het criterium goed/slecht is dit minder eenduidig want de scores voor de vijfjarigen verschillen niet significant van die voor 14-jarigen en de volwassen leken.

Voor het criterium techniek volgen de beoordelingen een ander patroon. Er zijn tussen de scores van de vijfjarigen en die van de 8- en 11-jarigen geen significante verschillen, maar de scores van de 14-jarigen, volwassen leken en volwassen kunstenaars zijn significant hoger. De tekeningen van de volwassen kunstenaars worden door beide groepen beoordelaars als veruit het meest technisch bekwaam beschouwd.

We concluderen dat de beroepskunstenaars in de klas op de criteria mooi, expressie, interessant en goed meer tot U-vormige patronen komen dan de groepsleerkrachten. Op de criteria expressie en interessant zijn hun oordelen bovendien extremer dan Gardner en Winner het in hun theorie hadden verwoord: zelfs de tekeningen van de volwassen kunstenaars komen in de ogen van hun collega-beoordelaars niet meer toe aan het niveau van de vijfjarigen.

#### **4.3 Zijn de oordelen van groepsleerkrachten vergelijkbaar met de oordelen van andere 'leken'?**

Evenals in eerder onderzoek blijkt dat de expertbeoordelaars, in dit geval de BIKkers, meer tot U-vormige oordelen komen dan de 'leken', in dit geval de groepsleerkrachten. Zoals aangegeven wordt expertise bepaald door het al of niet hebben van een kunstopleiding en door de kunstzinnige activiteiten die worden ontplooid. Toch worden ook door de groepsleerkrachten de tekeningen van vijfjarigen op de meeste criteria hoger beoordeeld dan van de 8-, 11- en 14-jarigen en ook hoger dan van de volwassen leken. Dat patroon wijkt af van het gemiddelde beoordelingspatroon van de eerdere onderzochte groepen leken. In die onderzoeken (o.a. Kindler, 2000; Pariser e.a., 2007; Haanstra, Van Hoorn, &

Tabel 1

*Enkelvoudige contrasten: gemiddelde kwadraten en F-toetsen voor 5 beoordelingscriteria. De beoordelingen van groepsleerkrachten en beroepskunstenaars over tekeningen van 8-jarigen, 11-jarigen, 14-jarigen, volwassen leken en volwassen kunstenaars vergeleken met de beoordelingen over tekeningen van 5-jarigen.*

	Mooi		Ex- pressie		Interes- sant	
	M <sup>2</sup>	F	M <sup>2</sup>	F	M <sup>2</sup>	F
<i>Groepsleerkracht</i>						
8-jarigen	9.529	23.495***	7.646	24.782***	5.010	15.409***
11-jarigen	4.697	10.717**	14.106	46.371***	8.331	23.450***
14-jarigen	5.969	10.222**	9.350	19.524***	5.142	10.110**
Volwassen leken	4.195	7.026*	12.848	27.604***	17.142	29.855***
Volwassen kunstenaars	13.058	15.268***	13.164	22.356***	14.581	20.893***
<i>Beroepskunstenaar in de klas</i>						
8-jarigen	18.251	67.448***	11.846	58.552***	6.885	24.926***
11-jarigen	6.675	18.076***	11.996	38.093***	6.918	18.317***
14-jarigen	23.154	46.534***	34.733	91.972***	15.956	36.790***
Volwassen leken	39.301	77.906***	50.009	120.462***	53.473	111.276***
Volwassen kunstenaars	2.086	2.402 ns	5.757	9.818**	3.683	5.153*

Vervolg Tabel 1

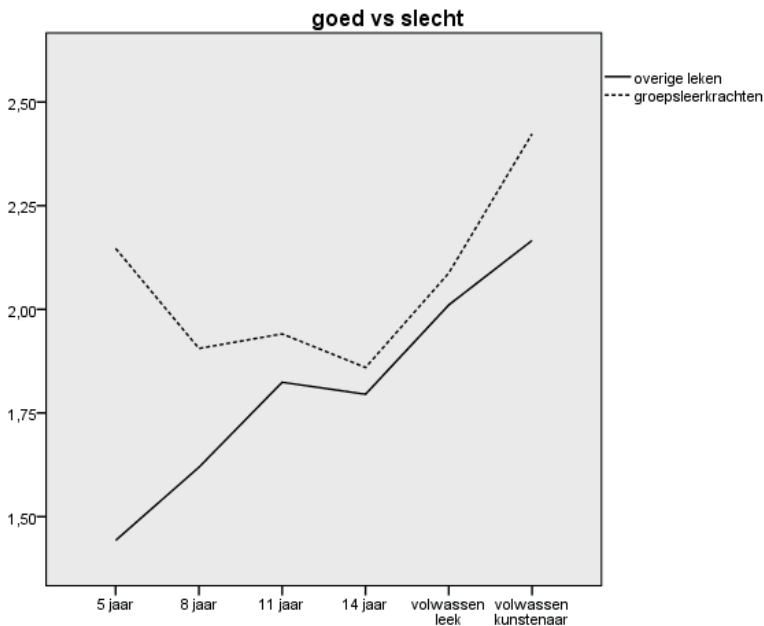
	Techniek		Goed /slecht	
	M <sup>2</sup>	F	M <sup>2</sup>	F
<i>Groepsleerkracht</i>				
8-jarigen	1.932	5.608*	11.100	26.346***
11-jarigen	0.077	0.157ns	4.629	10.387**
14-jarigen	16.781	24.832***	1.429	2.167ns
Volwassen leken	11.421	16.623***	1.690	2.236ns
Volwassen kunstenaars	157.937	181.352***	44.004	42.804***
<i>Beroepskunstenaar in de klas</i>				
8-jarigen	0.398	1.377ns	12.798	39.567***
11-jarigen	0.804	1.504ns	5.736	13.063**
14-jarigen	8.181	11.761**	12.050	18.914***
Volwassen leken	3.183	3.859*	26.255	31.017***
Volwassen kunstenaars	97.371	94.920***	0.574	0.454ns

Damen, 2011; Haanstra, Damen, & Van Hoorn, 2013) zijn de beoordelingspatronen van de leken steeds overwegend lineair stijgend, dus hoe ouder de leeftijd van de tekenaar des te hoger de waardering.

We kunnen op basis van de ons beschikbare data toetsen of de oordelen van de groepsleerkrachten inderdaad afwijken van andere leken-oordelen. In het huidige en in de twee voorgaande onderzoeken is een bestand van 36 tekeningen steeds door alle groepen beoordelaars op het criterium goed/slecht beoordeeld. We hebben vervolgens met een mixed model ANOVA getoetst of het gemiddelde beoordelingspatroon van de groepsleerkrachten afweek van het gemiddelde beoordelingspatroon van de twee lekengroepen uit de vorige onderzoeken: in Nederland wonende volwassenen en psychologiestudenten uit Boston. Uit eerdere analyses is gebleken dat deze twee groepen niet significant verschillen in hun oordelen over de verschillende leeftijdsgroepen.

De factor beoordelaars is de tussen-proefpersoon variabele en heeft twee niveaus: groepsleerkrachten basisonderwijs en overi-

ge leken. De factor leeftijd van de tekenaars is weer een binnen-proefpersoon variabele met zes niveaus. Mauchley's test gaf aan dat opnieuw niet werd voldaan aan de aanname van sphericity voor de hoofdeffecten van de leeftijd van de tekenaars. Daarom werd ook hier de Greenhouse-Geisser correctie toegepast. Figuur 7 laat de twee beoordelingspatronen zien. Er was wederom een significant hoofdeffect van leeftijd van de tekenaars  $F(3.48, 776.15) = 65.77, p = .000$ . Dus de tekeningen van de verschillende leeftijdsgroepen worden verschillend beoordeeld. Er was tevens een significant hoofdeffect van soort beoordelaar  $F(1,223) = 34.16, p = .000$  en tenslotte was er een significant interactie-effect tussen de factoren type beoordelaar en de leeftijd van de tekenaars  $F(3.48, 776.15) = 23.09, p = .000$ . De conclusie is dat het oordeel van de groepsleerkrachten inderdaad afwijkt van dat van de andere leken. De oordelen lijken een tussenpositie in te nemen tussen de meer lineair stijgende patronen van de andere leken groepen en de 'echte' U-vormige patronen van experts.



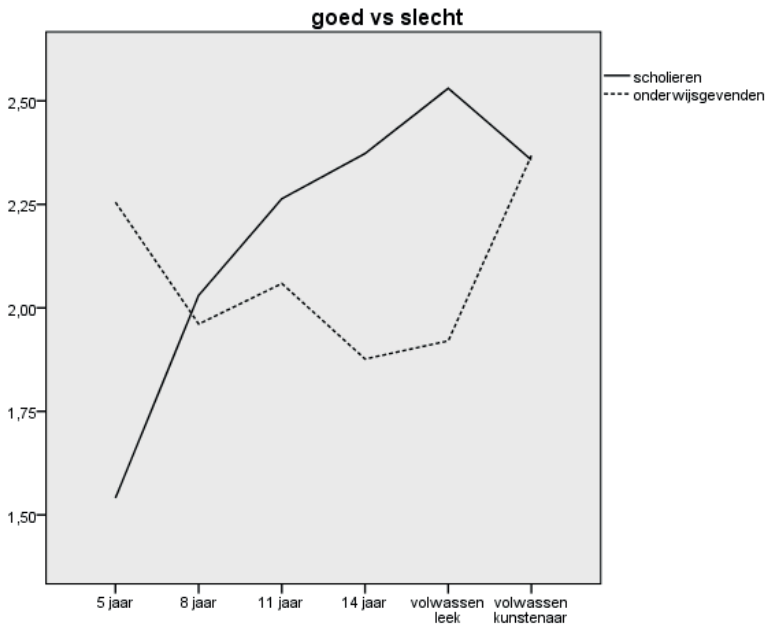
Figuur 7. Beoordelingspatronen van groepsleerkrachten en overige groepen leken voor het criterium goed/slecht.

#### 4.4 Zijn de oordelen van de onderwijsgevenden vergelijkbaar met die van leerlingen?

De oordelen van de groepsleerkrachten en de BIKkers verschillen minder van elkaar dan die tussen kunstprofessionals en andere volwassen leken. Een vervolgvraag is nu of de oordelen van de onderwijsgevenden (groepsleerkrachten en BIKkers) overeenkomen of afwijken van de oordelen die leerlingen geven aan de tekeningen van makers van verschillende leeftijden en expertise. Kindler (2000) wees in haar onderzoek al op het belang van de oordelen van de leerlingen voor de beoordelingspraktijk in het onderwijs. De oordelen over de tekeningen door leerlingen en docenten bleken in Taiwan niet van elkaar te verschillen, al gaven leerlingen en docenten wel andere redenen aan voor het hoger of lager beoordelen van een tekening. Bovendien hadden de leerlingen een sterkere voorkeur voor een ander type tekening (een cartoon) dan de docenten, die een voorkeur hadden voor de portrettekeningen. Het onderzoek van Kindler (2000) vond echter plaats in een andere cultu-

rele context (Taiwan) en betrof bovendien slechts 8 beoordelaars. In ons onderzoek beschikken we over gegevens van 55 basisschoolleerlingen tot en met twaalf jaar en 166 onderwijsgevenden in het basisonderwijs (groepsleerkrachten en BIKkers). Deze beoordelaars beoordeelden 36 dezelfde tekeningen van makers van verschillende leeftijden en expertise op het criterium goed/slecht met een driepuntsschaal.

We hebben met een mixed model ANOVA getoetst of het gemiddelde beoordelingspatroon van de onderwijsgevenden (groepsleerkrachten en beroepskunstenaars samen) afweek van het gemiddelde beoordelingspatroon van leerlingen tot en met 12 jaar. De factor onderwijsgevende versus scholieren is de tussen-proefpersoon variabele en heeft twee niveaus. De factor leeftijd van de tekenaars is de binnen-proefpersoon variabele met zes niveaus. Mauchley's test gaf aan dat niet werd voldaan aan de aanname van sphericity voor de hoofdeffecten van de leeftijd van de tekenaars. Daarom werd ook hier de Greenhouse-Geisser correctie toegepast. Er



Figuur 8. Beoordelingspatronen van onderwijsgevenden (groepsleerkrachten en beroepskunstenaars) en van scholieren t/m 12 jaar voor het criterium goed/slecht.

was een significant hoofdeffect van onderwijsgevende versus scholier  $F(1, 200) = 4.35, p = .000$ . Het beoordelingspatroon van de onderwijsgevendenden wijkt dus af van dat van de scholieren (zie figuur 8). Er was ook weer een significant hoofdeffect van leeftijd van de tekenaars  $F(3.63, 726.27) = 34.14, p = .000$ . Er was tevens een significant interactie-effect tussen de factoren onderwijsgevende/scholier en de leeftijd van de tekenaars:  $F(3.63, 726.27) = 69.20, p = .000$ .

Zoals we in dit artikel eerder lieten zien hebben onderwijsgevendenden een U-vormig beoordelingspatroon: de tekeningen van 5-jarigen worden hoger gewaardeerd dan die van de overige leeftijdsgroepen (8-, 11-, 14-jarigen en volwassen leken), maar niet hoger dan die van de volwassen kunstenaars. Scholieren hebben echter een geheel ander beoordelingspatroon: er is sprake van een stijgende lijn van de 5-jarigen tot en met de volwassen leken, waarna de tekeningen van de volwassen kunstenaars vervolgens lager gewaardeerd worden dan die van de volwassen leken (ongeveer vergelijkbaar met die van de 14-jarigen). Met uitzondering van deze laatste knik lijkt het beoordelingspatroon van de leerlingen op het patroon van de overige leken, maar tevens op het patroon dat ook onderwijsgevendenden laten zien bij het criterium techniek. Samenvattend kunnen we stellen dat de oordelen van scholieren op het criterium goed/slecht sterk afwijken van die van de onderwijsgevendenden, anders dan eerder in het onderzoek van Kindler (2000) naar voren kwam.

## 5 Conclusies en discussie

In de moderne westerse beeldende kunst werd conventionaliteit en het navolgen of herhalen van voorgangers of bestaande tradities niet hoog gewaardeerd. Een kunstenaar moest bestaande grenzen doorbreken en zijn unieke, persoonlijke visie uitdrukken. Het streven naar experiment en oorspronkelijkheid is kenmerkend voor het modernisme, dat zich vanaf circa 1880 ontwikkelde en in het abstract expressionisme van de jaren vijftig zijn hoogtepunt vond (Danto, 1997). Die laatstgenoem-

de periode in de professionele kunst ging samen met een kindgerichte benadering van het beeldende onderwijs op school, waarbij het streven naar creativiteit en persoonlijke expressie een belangrijke plaats in nam. De expressionistische en enigszins geabstraheerde uitingen van jonge kinderen stonden daarbij in hoog aanzien. Hoewel de professionele kunst hele andere richtingen uitging bleef die waardering voor de expressieve kindertekening bestaan. Belangrijke kenmerken ervan zijn door Efland (1976) de schoolkunststijl genoemd en dit soort schoolkunst bloeit ook nu nog. De overeenkomst tussen beeldende kwaliteiten van kindertekeningen en werk van professionele kunstenaars (en het ontbreken van die kenmerken in de conventionele uitingen van oudere kinderen en volwassenen) is door Gardner en Winner (1982) verwoord in de theorie van de U-vormige beeldende ontwikkeling. Deze theorie vormt de achtergrond van ons onderzoek. Het is een gedeeltelijke replicatie van eerdere onderzoeken naar deze theorie. Daaruit was al duidelijk geworden dat het universele karakter van de U-vormige beeldende ontwikkeling moet worden gerelativeerd. Afhankelijk van expertise, leeftijd en culturele achtergrond van de beoordelaars en van de gehanteerde beoordelingscriteria kunnen de beoordelingspatronen verschillen. Alleen bij westers georiënteerde experts (kunstenaars en docenten beeldende kunst) is het U-vormige beoordelingspatroon duidelijk aanwezig. Het is voor het beeldende onderwijs een relevant gegeven dat modernistische criteria - en dan vooral expressiviteit - nog steeds een belangrijk uitgangspunt voor de waardering van tekeningen en schilderijen van leerlingen lijkt te vormen. In het primair onderwijs zijn het in meerderheid echter niet vakdocenten met een beeldende opleiding, maar de groepsleerkrachten met een Pabo-opleiding die de kunstzinnige oriëntatie (en daarmee het beeldende onderwijs) verzorgen. Vaak worden daarnaast wel kortlopende projecten uitgevoerd door kunstenaardocenten, de zogeheten BIKkers (beroepskunstenaars in de klas). Dit roept de vraag op of de groepsleerkrachten en de beroepskunstenaars in de klas tot verschillende oordelen komen.

Groepsleerkrachten en BIKkers vormden



de twee groepen respondenten in het huidige onderzoek. Ze beoordeelden tekeningen van 5-, 8-, 11- en 14-jarigen, volwassen leken en volwassen beeldende kunstenaars op de criteria expressie, techniek, interessant, mooi en goed/slecht. Evenals in eerder onderzoek blijkt dat de expertbeoordelaars, in dit geval de BIKkers, meer tot U-vormige oordelen komen dan de 'leken', de groepsleerkrachten. Op de criteria expressie en interessant zijn ze zelfs 'plus royaliste que le roi', want daar krijgen de tekeningen van 5-jarigen een significant hogere waardering dan die van de professionele kunstenaars. De groepsleerkrachten beoordelen de tekeningen van 5-jarigen op de meeste criteria weliswaar niet zo hoog als de BIKkers, maar wel hoger dan de 8-, 11- en 14-jarigen en ook hoger dan de volwassen leken. Dat beoordelingspatroon wijkt af van het gemiddelde beoordelingspatroon van de eerdere onderzochte leken, dus respondenten zonder beeldende opleiding. In die onderzoeken zijn de beoordelingspatronen van de leken steeds overwegend lineair stijgend, dus hoe ouder de leeftijd van de tekenaar des te hoger de waardering. Groepsleerkrachten zullen beter dan de andere onderzochte leken groepen in staat zijn de leeftijden van de tekenaars in te schatten. Maar dat is op zich geen verklaring voor het afwijkende beoordelingspatroon. Het betekent dat ze heel bewust de tekeningen van jonge kinderen hoger waarderen dan van oudere kinderen en jongeren. Wel is het waarschijnlijk dat de gemiddelde groepsleerkracht meer deelneemt aan kunst dan de gemiddelde leek en in die zin dichter bij de professionele kunstenaars staat. Recente cijfers ontbreken, maar in 2003 toonde Kalmijn op basis van cultuurdeelnamecijfers aan dat onderwijsgevenden tot een culturele elite behoren (Kalmijn, 2003).

Een relativering van de resultaten is dat het onderzoek opdrachten betrof die het expressieve, modernistische idioom veel ruimte lieten. Men kan verwachten dat opdrachten waarbij ambachtelijke vaardigheden en technische kennis een belangrijke rol spelen (zoals bij toegepaste vormen van kunst of fotografie) tot andere beoordelingspatronen zullen leiden. De beoordelingen op het

criterium techniek wijken af van de andere criteria (expressie, mooi, interessant en goed). Dat criterium laat bij de groepsleerkrachten en BIKkers ook geen U-vormig patroon zien, maar een enigszins stijgend patroon per leeftijd. Het is een meer traditioneel criterium dat echter volgens de respondenten zelf in hun dagelijkse beoordelingspraktijk van minder belang wordt geacht dan creativiteit. Dat past weer in het beeld van een modernistisch beoordelingspatroon. Ondanks bovengenoemde kanttekening kan men concluderen dat bij degenen die het beeldend onderwijs in het primair onderwijs verzorgen de modernistische criteria nog sterk aanwezig zijn. Dit strookt met de dominante vormen van schoolkunst, die los staan van actuele postmoderne stromingen in de kunst, zoals de ontkenning van de formalistische esthetiek, het citeren van verschillende beeldende tradities en populaire beeldcultuur. Verder blijkt uit het onderzoek dat de oordelen van groepsleerkrachten en BIKkers sterk verschillen van die van schoolkinderen. Ook de onderzoeken van Kindler (2000), Pariser e.a. (2007) en Haanstra (2008) wijzen op het verschil tussen de beeldende voorkeuren van leerlingen en die van veel beeldende docenten. Leerlingen hechten meer aan technische vaardigheden en streven veel minder naar oorspronkelijkheid. Het goed kunnen natekenen van populaire beeldcultuur, zoals figuren uit strips of games, wordt door hen juist positief beoordeeld. Kindler (2000) stelt dat als de esthetische criteria en kunstzinnige aspiraties van leerlingen zo verschillen van die van hun docenten dit een probleem kan vormen voor beeldend onderwijs. Volgens Kindler beschikt een kind in potentie over een breed beeldend repertoire, dat wil zeggen dat het kan kiezen tussen verschillende soorten beeldende uitingen met verschillende functies. Maar in het onderwijs 'belonen' docenten bewust of onbewust vooral de beeldende uitingen die in de richting gaan van de modernistische westerse kunst. Door slechts bepaalde uitingen te stimuleren en waarderen beperken we volgens Kindler de beeldende ontwikkeling van kinderen. Diverse auteurs hebben niet alleen gepleit voor beeldend onderwijs dat meer verbinding heeft met voorkeuren en interesses van leer-

lingen, maar dat tevens een toegang wil verschaffen tot actuele ontwikkelingen in het vakgebied (o.a. Haanstra, 2011; Heijnen, 2011). Dit kan zowel de persoonlijke als de maatschappelijke relevantie van het beeldenonderwijs versterken. De conclusie is echter niet dat de docenten hun oordelen moeten laten samenvallen met die van hun leerlingen of met die van de meest actuele stromingen in de beeldende kunst. Van belang is wel dat er voor de beeldende vorming in het primair onderwijs een beoordelingssysteem komt, dat transparant is, dat wil zeggen dat er expliciete criteria worden gehanteerd waarover betrokkenen van mening kunnen verschillen.

Inventariserende onderzoeken naar kunstzinnige oriëntatie in het basisonderwijs laten zien dat het ontbreekt aan een transparante en systematische beoordeling van leerlingprestaties op dit gebied. In het huidige onderzoek onder groepsleerkrachten en BIKkers wordt dit gegeven bevestigd. Gevraagd naar hun beoordelingspraktijk blijkt dat de meesten terughoudend zijn met het summatief beoordelen van tekeningen van hun leerlingen. De vaak informele vormen van formatieve beoordeling kunnen ertoe leiden dat de criteria en voorkeuren van de onderwijsgeevenden (die er blijkens het onderzoek wel zijn) impliciet blijven en dat dit de bovenbeschreven beperkende invloed kan hebben.

Onderwijsbeleid en -praktijk vragen om een minder vrijblijvende invulling van de kunstvakken (o.a. Onderwijsraad & Raad voor Cultuur, 2012). Gesproken wordt over doorlopende leerlijnen, referentiekaders, competentieniveaus e.d. om gewenste kennis, vaardigheden en attitudes voor bepaalde leeftijden in kaart te brengen. Daarbij hoort een instrumentarium voor het volgen en beoordelen van leerlingprestaties. Vervolgonderzoek naar het beoordelen van identieke opdrachten uitgevoerd door verschillende leeftijdsgroepen, zoals in het huidige onderzoek, kan bijdragen aan de empirische onderbouwing van zowel het vaststellen van competentieniveaus als van werkbare beoordelingsinstrumenten. De opdrachten in vervolgonderzoek zouden een breder spectrum moeten omvatten dan de meer traditionele tekenopdrachten uit het huidige onderzoek. Ook voor andere disciplines

dan de beeldende kunst zoals muziek, dans, theater en literair schrijven is zulk onderzoek relevant. Het kan inzicht verschaffen in de interactie van competentiegerichte opdrachten, verschillende beoordelingscriteria en kenmerken en voorkeuren van beoordelaars.

## Noot

1. Dezelfde patronen vonden we bij jongere en oudere beoordelaars, voor de portret en verdriettekeningen afzonderlijk en bij vrouwelijke en mannelijke beoordelaars. Vrouwelijke beoordelaars beoordelen de tekeningen op elke leeftijd hoger dan mannelijke beoordelaars, maar de patronen bij de verschillende beoordelingscriteria zijn hetzelfde voor mannelijke en vrouwelijke beoordelaars.

## Literatuur

- Asselbergs, V. (1989). *Kind, kunst en opvoeding. De Nederlandse beweging voor beeldende expressie (1876-1968)*. Academisch proefschrift. Amsterdam: Universiteit van Amsterdam.
- Berlyne, D. E. (Ed.). (1974). *Studies in the new experimental aesthetics*. Washington, DC: Hemisphere.
- Damen, M. L. (2010). *Cultuurdeelname en CKV. Studies naar de effecten van kunsteducatie op de cultuurdeelname van leerlingen tijdens en na het voortgezet onderwijs*. Academisch Proefschrift. Universiteit Utrecht.
- Danto, A. (1997). *After the end of art: Contemporary art and the pale of history*. Princeton: Princeton University Press.
- Davis, J. (1991). *Artistry lost: U-shaped development in graphic symbolization*. Unpublished Dissertation Harvard University, Graduate School of Education.
- Duncum, P. (1986). Breaking down the U curve of artistic development. *Visual Arts Research*, 12(1), 43-54.
- Efland, A. (1976). The school art style: A functional analysis. *Studies in Art Education*, 17(2), 37-44.
- Gardner, H. & Winner, E. (1982). First imitations of artistry. In S. Strauss (Ed.), *U-shaped behavi-*

- oral growth* (pp. 147-168). New York: Academic Press.
- Goodman, N. (1984). *Of mind and other matters*. Cambridge: Harvard University Press.
- Haanstra, F. (2001). *De Hollandse schoolkunst: mogelijkheden en beperkingen van authentieke kunsteducatie*. Utrecht: Cultuurnetwerk Nederland.
- Haanstra, F. (2008). *De thuiskunst van scholieren*. Amsterdam: Lectoraat Kunst- en cultuureducatie, Amsterdamse Hogeschool voor de Kunsten.
- Haanstra, F. (2011). Authentieke kunsteducatie: Een stand van zaken. *Cultuur + Educatie*, 31, 8-36.
- Haanstra, F., Hoorn, M. van, & Damen, M. L. (2011). The U-curve going Dutch: Cultural differences in judgements of artwork from different age and expertise groups. *International Journal of Education through Art*, 7(2), 153-169.
- Haanstra, F., Damen, M. L., & Hoorn, M. van (2013). Interestingness and pleasingness of drawings from different age and expertise groups. *Empirical Studies of the Arts*. Advance online publication. doi: 10.2190/EM.EOV.2.
- Heijnen., E (2011). Informele visuele netwerken. *Cultuur + Educatie*, 30, 32-53.
- Hekkert, P. (1995). *Artful judgements. A psychological inquiry into aesthetic preference for visual patterns*. Dissertation. Delft: Technische Universiteit Delft.
- Kalmijn, M. (2003). Trends in de culturele leefstijl van onderwijsgevenden. Nog steeds een elite? *Cultuur + Educatie*, 7, 50-70.
- Kindler, A. (2000). From the U-Curve to dragons: Culture and understanding of artistic development. *Visual Arts Research*, 26(2), 15-28.
- Leder, H., Belke, B., Oeberst, A., & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgments. *British Journal of Psychology*, 95, 489-508.
- Löwenfeld, V. & Brittain, W. L. (1975). *Creative and Mental Growth*. New York: MacMillan Publishing Co.
- Onderwijsraad & Raad voor Cultuur (2012). *Cultuureducatie: leren, creëren, inspireren!* Den Haag: Onderwijsraad.
- Oomen, C., Visser, I., Donker, A., Beekhoven, S., Hoogeveen, K., & Haanstra, F. (2009). *Cultuureducatie in het primair en voortgezet onderwijs. Monitor 2008-2009*. Utrecht: Oberon/Sardes.
- Pariser, D. (2006). Considering the U-curve. In J. A. Schaler (Ed.), *Howard Gardner under fire: The rebel psychologist faces his critics* (pp. 255-275). Chicago: Open Court Publishing Co.
- Pariser, D., & Berg, A. van den (1997). The mind of the beholder: some provisional doubts about the U-curved aesthetic development thesis. *Studies in Art Education*, 38(3), 158-178.
- Pariser, D., & Berg, A. van den (2001). Teaching art versus teaching taste: what art teachers can learn from looking at a cross cultural evaluation of children's art. *Poetics*, 29(6), 331-350.
- Pariser, D., Kindler, A., Berg, A. van den, Dias, B. & Chen Liu, W. (2007). Does practise make perfect? Children's and adults' constructions of graphic merit and development: a cross-cultural study. *Visual Arts Research*, 33 (65), 96-114.
- Wilson, B. (1981). Response to Winner and Gardner. *Review of research in art education*, (7)14, 32-33.
- Wilson, B. (2004). Child art after modernism. In E.W. Eisner & M.D. Day (Eds.), *Handbook of research and policy in art education* (pp. 299-328). New Jersey: Lawrence Erlbaum Publishers.

Manuscript aanvaard op: 27-09-2013

## Auteurs

**Folkert Haanstra** is verbonden aan de Amsterdamse Hogeschool voor de Kunsten / Universiteit Utrecht, **Marie-Louise Damen** aan de Universiteit van Amsterdam en **Marjo van Hoorn** werkt bij het Landelijk Kennisinstituut Cultuureducatie en Amateurkunst (LKCA)

Correspondentieadres: folkert.haanstra@ahk.nl

## Abstract

### Assessment of visual arts in primary education

Teachers in primary education with different levels of art expertise (general teachers and artist teachers) judged a collection of drawings produced by people from different age and expertise groups (5, 8, 11 and 14-year-old children, and adult artists and non-artists). Five assessment criteria were used, namely expression, technical quality, pleasingness, interestingness and 'overall quality' (good-poor). Previous studies had de-

monstrated that unlike laymen, art experts value drawings by the youngest children and artists more highly than drawings from the other groups. In other words, they produce U-curve appreciation patterns. This study confirms the mean judgment pattern of artist teachers for all assessment criteria except technical quality is a U-curve. Technical quality improves with age, but it is not considered an important criterion by both artist teachers and general teachers. Pupils hold an opposite view. The results are discussed in relation to more transparent assessment in visual arts education in primary schools.

## Appendix: Steekproeftrekking beoordelaars

De steekproef van de beoordelaars is als volgt in zijn werk gegaan. De groepsleerkrachten zijn geselecteerd via de basisschool waar ze werkzaam zijn. We zijn uitgegaan van de scholenlijst van de Centrale Financiën Instellingen (CFI). In 2010 stonden daar 7199 primair onderwijs scholen (inclusief het speciaal onderwijs) geregistreerd. Uit deze lijst is een getrapte steekproef getrokken van 400 scholen, rekening houdend met een evenredige vertegenwoordiging over de provincies. Vervolgens is door een random startpunt te kiezen bepaald welke leerkracht (groep 1 t/m 8) van de school benaderd zou worden. Hier is gestreefd naar een evenredige vertegenwoordiging van onder-, midden-, en bovenbouwleerkrachten. De scholen zijn gebeld om de namen en - indien beschikbaar - e-mailadressen van de betreffende leerkrachten te achterhalen. Als er meerdere leerkrachten voor een groep stonden, is er met toeval bepaald welke leerkracht benaderd zou worden.

Wanneer het e-mailadres niet beschikbaar was, is de leerkracht met een brief op het adres van de school benaderd. Na een eerste ronde, was de respons iets meer dan 10% (40 volledig ingevulde beoordelingen en enkele onvolledige). Daarop is besloten uit de resterende groep van 6799 scholen opnieuw een steekproef te trekken, ditmaal van 200 scholen. Vervolgens is op gelijke wijze bepaald welke groepsleerkracht per school benaderd moest worden. Eveneens dezelfde procedure is gevolgd om de namen en e-mailadressen van de leerkrachten te achterhalen. De uiteindelijke respons was 12%, oftewel 73 groepsleerkrachten die de opdracht volledig hebben beoordeeld. Het is onbekend of deze relatief lage respons geleid heeft tot een bias naar culturele interesse in de steekproef, maar dit is vanzelfsprekend niet uit te sluiten. Groepsleerkrachten zijn bovendien in het algemeen wellicht cultureel actiever dan de gemiddelde Nederlander (Kalmijn, 2003), maar ze zijn gemiddeld nog wel minder cultureel actief dan de beroepskunstenaars in de steekproef. Dit is naast de beroepsopleiding een indicatie voor hun verschil in expertise in de kunsten.

De beroepskunstenaars zijn aangeschreven via de afstudeerlijsten van de verschillende BIK-opleidingen. De opleidingen van de AHK in Amsterdam, HKU in Utrecht, Artez in Zwolle, Minerva in Groningen en Willem de Kooning in de Rotterdam zijn gevraagd om de lijsten met afgestudeerden en hun discipline te leveren over de periode 2006-2010. Niet alle opleidingen konden deze informatie over alle jaren leveren en niet alle opleidingen hadden de e-mailadressen van hun afgestudeerden. De lijsten zijn zo volledig mogelijk gemaakt door de ontbrekende informatie aan te vullen met gegevens uit de database van beroepskunstenaarsindex.nl, die online beschikbaar is. Per email zijn 336 beroepskunstenaars aangeschreven met de vraag of zij mee wilden werken aan het onderzoek, 71 BIKkers hebben een volledige beoordeling geleverd. De respons is derhalve 21%.