



BEOORDELEN VAN ZANGONDERWIJS IN HET PRIMAIR ONDERWIJS

Amsterdamse Hogeschool voor de Kunsten
Master Kunsteducatie
1^e jaar literatuuronderzoek

Rivka Worms
02.05.2017

Inhoudsopgave

INLEIDING	3
PARAGRAAF 1: WAT VOOR VERSCHILLENDE SOORTEN BEOORDELINGSINSTRUMENTEN ZIJN ER GEPUBLICEERD VOOR HET ZANGONDERWIJS IN HET PRIMAIR ONDERWIJS?	5
PARAGRAAF 2: WELKE BEOORDELINGSCRITERIA, SPECIFIEK VAN BELANG VOOR HET ZANGONDERWIJS IN HET PRIMAIR ONDERWIJS, WORDEN IN DE INSTRUMENTEN OPGENOMEN?	9
PARAGRAAF 3: WAT IS DE KWALITEIT VAN DEZE INSTRUMENTEN IN TERMEN VAN VALIDITEIT, BETROUWBAARHEID, BRUIKBAARHEID EN AUTHENTICITEIT?	12
PARAGRAAF 4: OP WELKE MANIER KUNNEN DE BESCHREVEN INSTRUMENTEN WORDEN INGEZET OM HET LEREN VAN LEERLINGEN TE STIMULEREN?	15
PARAGRAAF 5: CONCLUSIE EN DISCUSSIE	17
BIJLAGEN	19
LITERATUURLIJST	21

Aantal woorden: 6.549

INLEIDING

Beoordelen en toetsen zijn de laatste jaren sterk aanwezig in het basisonderwijs in Nederland (Ros, 2014). Leerlingen krijgen, naast de verplichte eindtoets voor het primair onderwijs, voor steeds meer vakken toetsen. Waar in de zaakvakken, aardrijkskunde, geschiedenis en natuur & techniek, beoordelen al veel wordt ingezet, ligt dit in het muziekonderwijs een stuk gecompliceerder.

Dit komt ten eerste omdat muziekdocenten voelen dat er altijd te weinig tijd is; als de muziekdocent de leerlingen maar één keer per week ziet, hoe lukt het dan om ook nog de leerlingen te beoordelen (Dunbar, 2011)? Daarnaast is de inhoud van het muziekonderwijs op sommige punten moeilijk te meten; de leerling leert vaak dat er meerdere juiste antwoorden en interpretaties in muziek mogelijk zijn en dat er dikwijls geen "goed of fout" is. Het is dan ingewikkeld om de leerling wel op deze wijze te beoordelen (Ros, 2014).

Dunbar (2011) stelt echter in haar artikel dat het beoordelen van prestaties van leerlingen juist helemaal niet tijdrovend hoeft te zijn en dat het een goeie manier is om de vaardigheden van de leerlingen te beoordelen. Chiodo (2001) beaamt dat het ingewikkeld 'lijkt' om tijd vrij te maken voor beoordelen, zonder dat er een deel van de muziekles verloren gaat, maar dat er gelukkig genoeg instrumenten zijn ontwikkeld om effectief en efficiënt te beoordelen zonder dat de leerlingen daarnaast het plezier in de muziekles kwijt raken.

In mijn eigen muziek-lespraktijk in het basisonderwijs wordt muziek vaak als 'pret-vak' gezien; het wordt niet beoordeeld en dient enkel als ontspanning voor de leerlingen. Als muziekonderwijs wordt gezien als 'de koffiepauze voor de klasleerkracht' of als leuk 'extraatje' tussen twee reguliere vakken, dan is het moeilijk om muziekonderwijs qua waarde gelijk te stellen aan de andere vakken op school. Doordat muziek doorgaans niet wordt beoordeeld, voelt het vak niet als onderdeel van het curriculum. Door het beoordelen van muziek, zo bepleit Chiodo (2001), kunnen we het vak onderdeel maken van het kerncurriculum van de school. Mede daardoor kan het muziekonderwijs dan beter aansluiten op de landelijke richtlijnen van het basisonderwijs.

Zonder beoordeling is het daarnaast lastig om de prestaties van de leerlingen terug te koppelen naar ouders en klassenleerkracht. Door enkel de aanwezigheid en het gedrag te beoordelen, zullen ouders en docenten nooit een compleet beeld krijgen van wat de leerlingen leren tijdens de muziekles (Leonhardt, zoals geciteerd in Dunbar, 2011). Het is daarom aan te bevelen om een helder beoordelingsinstrument in te zetten; laatstgenoemde ontbreekt momenteel in mijn eigen praktijk.

Eén van de belangrijkste muzikale vaardigheden die een jong kind moet ontwikkelen is het kunnen gebruiken van de zangstem (Chiodo, 2001). In mijn eigen lespraktijk wordt zang het meest ingezet tijdens de muzieklessen. Het is makkelijk in te zetten en brengt geen extra kosten met zich mee. Daarom wordt er in dit onderzoek specifiek gekeken naar het beoordelen van zangonderwijs. In het leerplankader voor kunstzinnige oriëntatie (SLO, 2015) wordt in de leerlijn muziek dan ook veel aandacht besteed aan zingen, zowel in groepsverband als individueel. De muziekdocent moet dan ook de mogelijkheid hebben om de zangprestaties van de leerlingen te kunnen beoordelen met een goed en betrouwbaar beoordelingsinstrument.

Voor mij als muziekdocent is het belangrijk dat dit literatuuronderzoek wordt uitgevoerd, omdat ik in mijn lespraktijk wil werken met een bruikbaar beoordelingsinstrument voor het zangonderwijs. Daarnaast is het van belang voor de beroepspraktijk dat er onderzoek wordt gedaan naar het beoordelen van specifiek zangonderwijs; zang is namelijk een belangrijk onderdeel van de muzieklessen in het primair onderwijs. Een geschikt beoordelingsinstrument, specifiek voor zangonderwijs, zal dan ook een fijne toevoeging zijn voor het muziekonderwijs in het P.O.. Ten slotte kan dit literatuuronderzoek belangrijk zijn voor de onderzoekswereld. Wanneer de verschillende beoordelingsinstrumenten worden belicht, besproken en met elkaar worden vergeleken, zou dit kunnen leiden tot verder onderzoek over dit onderwerp.

De **hoofdvraag** van dit onderzoek luidt daarom als volgt:

- Welke beoordelingsinstrumenten voor het zangonderwijs in het primair onderwijs zijn er in de literatuur beschreven en hoe verhouden deze zich tot elkaar?

Om de hoofdvraag te kunnen beantwoorden zijn de volgende **deelvragen** opgesteld:

- Wat voor verschillende soorten beoordelingsinstrumenten zijn er gepubliceerd voor het zangonderwijs in het primair onderwijs?
- Welke beoordelingscriteria, specifiek van belang voor het zangonderwijs in het primair onderwijs, worden in de instrumenten opgenomen?
- Wat is de kwaliteit van deze instrumenten in termen van validiteit, betrouwbaarheid, authenticiteit en bruikbaarheid?
- Op welke manier kunnen de beschreven instrumenten worden ingezet om het leren van leerlingen te stimuleren?

BEGRIPSBEPALING

Beoordelingsinstrumenten

Er is een onderscheid tussen de taak die de leerling moet uitvoeren, bijvoorbeeld het uitvoeren van een muziekstuk, en het scoringsinstrument om deze taak vervolgens beoordelen. Samen vormen deze twee onderdelen het beoordelingsinstrument (Groenendijk, Damen, Haanstra & van Boxtel, 2015). Een beoordelingsinstrument kan worden ingezet om de resultaten van de leerlingen te meten en zo meer inzicht in hun ontwikkeling te krijgen. Aan de hand van de landelijke kerndoelen en de leerdoelen die de docent opstelt, kan een beoordelingsinstrument worden ontwikkeld. Er zijn verschillende soorten scoringsinstrumenten om de resultaten van de taak van leerlingen te meten; een observatie verslag, een logboek, een rubric, een reflectiegesprek of een rapport (LKCA, n.d.). Het scoringsinstrument kan worden gebruikt door zowel de docent als de leerling, maar ook een medeleerling kan de taak beoordelen. Het scoringsinstrument kan specifiek op een taak of product gericht zijn, maar kan ook algemener gebruikt worden bij een meer open taak of proces (Groenendijk, Damen, Haanstra & van Boxtel, 2015).

Beoordelingscriteria

Het SLO (2006) beschrijft dat een beoordelingsinstrument wordt ingedeeld in verschillende vaardigheden die de leerling moet beheersen. Zo een vaardigheid is gekoppeld aan een set helder geformuleerde beoordelingscriteria. Deze criteria geven weer welke aspecten

belangrijk zijn bij een bepaalde vaardigheid. Per beoordelingscriterium staat aangegeven waarop de beoordelaar moet letten. Dit kan worden omschreven in waarneembaar gedrag of resultaat. De criteria zijn van belang om de leerprestaties van de leerling inzichtelijk te maken voor zowel de docent als voor de leerling (Toetswijzer, n.d.). Belangrijke beoordelingscriteria binnen het muziekonderwijs zijn: expressie, interpretatie, technische beheersing, gehoor-klankvoorstelling, ritmische beheersing, muzikaal vocabulaire, creativiteit, samenspel en zelfreflectie (Groenendijk, Damen, Haanstra & Boxtel, 2015).

Zangonderwijs in het primair onderwijs

Het zangonderwijs dat wordt besproken in dit onderzoek, is het onderwijs dat wordt gegeven door vakleerkrachten. De beoordelingsinstrumenten die hier worden uitgezet zijn dan ook allen bedoeld voor de vakleerkracht. Hoogeveen & Oomen (2009) constateerde dat op 19% van alle scholen in Nederland een vakleerkracht of een freelancer aanwezig was voor het geven van cultuureducatie. In de meeste gevallen ging het hierbij om muzikleerkrachten. Momenteel neemt, met het huidige beleid Cultuureducatie met Kwaliteit (LKCA, n.d.), het aantal vakleerkrachten de laatste jaren meer toe. Volgens peilingen van Mocca (2014) zijn er van de 249 scholen in Amsterdam, 172 scholen die muziek als eerste leerlijn hebben gekozen en 9 scholen als tweede leerlijn als invulling van het Basispakket Amsterdam.

PARAGRAAF 1: Wat voor verschillende soorten beoordelingsinstrumenten zijn er gepubliceerd voor het zangonderwijs in het primair onderwijs?

VIER CATEGORIËN BEOORDELINGSINSTRUMENTEN

In de literatuur die er is geschreven over het beoordelen van muzikeducatie, zijn door de jaren heen veel verschillende soorten beoordelingsinstrumenten ontwikkeld en beschreven. In deze paragraaf zullen de verschillende beoordelingsinstrumenten tegen elkaar afgezet worden. Beoordelingsinstrumenten kunnen volgens Salvador (2010) worden bekeken aan de hand van twee aspecten; de *manier* van meten en de *oorsprong* van de meetinstrumenten. De twee aspecten kunnen weer opgesplitst worden in; *acoustic measurements* en *researcher-as-acoustic measurements* en in *researcher-developed measurements* en *previously published scales*.

De eerste twee categorieën zeggen iets over de manier van meten; bij *acoustic measurements* wordt gebruik gemaakt van computers en software bij het nauwkeurig meten van toonhoogte gebaseerd op de frequentie in Hertz, waar bij de *researcher-as-acoustic measurements* geen gebruik wordt gemaakt van computers en software maar van de vakbekwaamheid van mensen: de onderzoeker is als het ware het scoringsinstrument. In de derde en vierde categorie wordt naar de oorsprong van de instrumenten gekeken; de *researcher-developed measurements* zijn instrumenten die zijn ontwikkeld door onderzoekers voor het meten van de zangprestaties en de *previously published scales* maken gebruik van bestaande instrumenten die reeds ontwikkeld zijn door onderzoekers in

het verleden. Om te beoordelen waar de verschillende instrumenten in dit onderzoek onder vallen wordt zowel gekeken naar de manier van meten als naar de oorsprong van het instrument.

SOORTEN BEOORDELINGSINSTRUMENTEN

Veel onderzoekers en docenten gebruiken een ratingschaal voor het beoordelen van de leerlingen (Leighton & Lamont, 2006). De meest gebruikte scoringsinstrumenten in deze soort zijn de checklist, de numerieke schaal en de rubric (Chiodo, 2001). De checklist, rubric en numerieke schaal vallen onder de *teacher-as-acoustic measurements*, waarbij de oorsprong van de instrumenten kan variëren. Een rubric is een analytische beoordelings-schaal die bestaat uit meerdere beoordelingscriteria, hetgeen docenten in staat stelt om de leerlingen te evalueren op de kwaliteit van een uitgevoerde opdracht. (SLO, n.d.). Rubrics worden veelvuldig ingezet in het muziekonderwijs, omdat ze de complexiteit van het beoordelen van muziek goed omvatten; zowel de technische aspecten als de expressieve en prestatie aspecten worden beoordeeld (DeLuca & Bolden, 2014).

Veel van de beoordelingsinstrumenten die zijn bekeken voor dit onderzoek, vallen onder de categorie '*acoustic measurements*' van Salvador (2010), waarbij gebruikt wordt gemaakt van computermeting. Leighton en Lamont (2006) stellen dat het gebruik van computeranalyse bij individuele zangstemmen misschien wel een gedetailleerde en accurate analyse geeft, maar deze nooit écht de zangvaardigheden van de leerlingen kan beoordelen. Dit omdat zangvaardigheid over meer gaat dan alleen de nauwkeurigheid van toonhoogte. Daarnaast zijn deze onderzoeken ook moeilijk in de context van de school te plaatsen; de mogelijkheden van beoordeling die beschikbaar zijn voor de muziekdocenten in de school, waar kinderen vaak in grote groepen zingen, zijn dermate anders dan de mogelijkheden van de onderzoekers. Om deze redenen is dan ook gekozen om de *acoustic measurements* niet verder te beschrijven in dit onderzoek.

DE BEOORDELINGSINSTRUMENTEN

Voor dit onderzoek zijn verschillende keuzecriteria gebruikt voor de uiteindelijke selectie van beoordelingsinstrumenten. Twee belangrijke criteria zijn; beoordelingsinstrumenten voor het basisonderwijs en specifiek het zangonderwijs op deze scholen. In het onderzoek van Groenendijk, Damen, Haanstra & van Boxtel (2015) zijn veel bestaande instrumenten uiteengezet. Dit onderzoek bevat 98 beoordelingsinstrumenten specifiek voor muziek, waarvan 43 zijn toegespitst op het basisonderwijs. Van deze 43 instrumenten gaan er 8 expliciet over zang.

Daarnaast is een ander criterium het uitsluiten van *acoustic measurement*. Na deze selectie blijven er nog zes instrumenten over. Vervolgens wordt gekeken of het beoordelingsinstrument een bruikbaar scoringsinstrument heeft waarmee de leerlingen in de klas beoordeeld kunnen worden. Na dit criterium blijven er drie bruikbare instrumenten over. Het laatste criterium is het jaartal van uitgave. Voor een actueel onderzoek naar een onderwerp dat constant in ontwikkeling is, is het voor dit onderzoek niet actueel genoeg om onderzoeken te gebruiken van voor 1995. Na deze selectie blijven er uiteindelijk twee beoordelingsinstrumenten over. In de twee overgebleven beoordelingsinstrumenten worden twee scoringsinstrumenten gebruikt; de Singing Voice Development Measure van

Rutkowski (1998) en de Vocal Pitch Development Measure van Welch (1998). Naast het onderzoek van Groenendijk, Damen, Haanstra & van Boxtel (2015) is het aanbod van beoordelingsinstrumenten die onder al deze criteria vallen beperkt. In andere onderzoeken, verwijzingen en bronnen zijn uiteindelijk nog drie beoordelingsinstrumenten gevonden die binnen deze criteria vallen. In onderstaande tabel zijn de vijf beoordelingsinstrumenten uiteengezet.

Beoordelingsinstrumenten	Auteur	Jaartal	Categorie instrument	Soort Instrument	Taak beoordeling	Product/ Proces beoordeling	Vakleerkracht/ Leerkracht/ Leerling/ Peers
Singing Voice Development Measure	J. Rutkowski	1998	researcher-developed measurement & researcher-as-acoustic-measurement	Rubric	Uiting in het artistieke medium	Product en Proces	Vakleerkracht
The Singing Voice Rubric	P. Chiodo	2001	previously published scale & researcher-as-acoustic-measurement	Rubric	Uiting in het artistieke medium	Product	Vakleerkracht
Vocal Pitch-Matching Development	G. Welch	1998	researcher-developed measurement & researcher-as-acoustic-measurement	Rubric	Uiting in het artistieke medium	Product en Proces	Vakleerkracht
Singing Self Assessment	P. E. Riley	2010	researcher-developed measurement & researcher-as-acoustic-measurement	zelf rapportage en rapportage expert	Uiting in het artistieke medium	Product	Vakleerkracht en leerling
Boardman Scale	J. Boardman	1997	researcher-developed measurement & researcher-as-acoustic-measurement	Rubric en zelf rapportage	Combinatie taak	Product	Vakleerkracht en Leerling

1. SINGING VOICE DEVELOPMENT MEASURE

(researcher-developed measurement & researcher-as-acoustic measurement)

Joanne Rutkowski begon in de jaren 90' met haar onderzoek naar de zangstem van kinderen. Rutkowski (2003) stelde dat zang altijd een belangrijk onderdeel is geweest van de muziekles en er vele manieren zijn bedacht om kinderen beter te leren zingen, maar dat sommige kinderen gewoonweg niet in staat zijn om te zingen. Het is misschien onrealistisch om te denken dat alle kinderen kunnen leren zingen door middel van wekelijkse muzieklessen van 30 of 40 minuten. Om dit te onderzoeken, en om het proces van de leerlingen te volgen, ontwikkelde zij de *Singing Voice Development Measure* (SVDM). Met dit instrument wordt bekeken of leerlingen vooruit zijn gegaan ten opzichte van het jaar daarvoor. Het instrument is bedoeld om de ontwikkeling van de zangstem bij kinderen te meten. De SVDM (Tabel 1, zie bijlage) is een 5-punten scoringsinstrument dat is ontwikkeld voor gebruik in de klas. Niveau 1 is het laagst haalbaar, niveau 3,5 het hoogst. Door dit instrument kunnen de verschillende fases in de ontwikkeling van de kinder-zangstem in kaart worden gebracht en kan het gebruik van de zangstem worden gemeten

(Levinovitch, 1998). In Rutkowski's onderzoek werd een bekend liedje in harmonisch mineur als uitgangspunt genomen. Zij koos voor een lied in mineur, omdat meerdere onderzoeken hadden aangetoond dat het voor kinderen makkelijker is om in mineur, dan in majeur, te zingen. Daarnaast maakte zij gebruik van een bekend lied voor de leerlingen, omdat kinderen waarschijnlijk beter presteren als zij een bekend lied moeten zingen.

2. THE SINGING VOICE RUBRIC

(previously published scale & researcher-as-acoustic measurement)

De *Singing Voice Rubric* van Chiodo (2001), is afgeleid van de *Singing Voice Development Measure* die is ontwikkeld door Joanne Rutkowski (Rutkowski & Runfola, zoals geciteerd in Chiodo, 2001). De muziekdocent moet volgens Chiodo de mogelijkheid krijgen om de ontwikkeling van de individuele zangstem van de leerlingen te beoordelen. Hiervoor is het belangrijk dat de muziekdocent de leerlingen ook solo hoort zingen. Aangezien kinderen soms niet alleen durven te zingen, zijn zangspelletjes hier een geschikt middel voor. Aan de hand van deze zangspelletjes worden de leerlingen beoordeeld op het correct gebruik van de zangstem.

Het zangspel bij de *Singing Voice Rubric* is als volgt omschreven (Chiodo, 2001); één leerling zit met zijn rug naar de klas, deze leerling is de 'rader'. Deze leerling kan de rest van de klas niet zien. De klas zingt het lied en de docent wijst één leerling aan die het lied verder solo zingt. De rader probeert de identiteit van de solist te raden door naar het geluid van de zangstem van deze leerling te luisteren. Nadat de rader de juiste leerling heeft geraden, mag de solist deze rol overnemen. Wanneer er telkens een nieuwe leerling aangewezen wordt om solo te zingen, kan de docent dit moment gebruiken om het niveau van de leerling aan de hand van de *Singing Voice Rubric* (Tabel 2, zie bijlage) te noteren. Het beoordelen gaat onopvallend terwijl de docent het spel leidt. Afhankelijk van de lengte van het lied, kan de gehele klas in ongeveer tien minuten worden beoordeeld.

3. VOCAL PITCH-MATCHING DEVELOPMENT

(researcher-developed measurement & researcher-as-acoustic measurement)

De *Vocal Pitch-Matching Development* (Tabel 3, zie bijlage) werd in 1998 ontwikkeld door Welch (2011). Dit beoordelingsinstrument meet de *vocal pitch-matching*; het met de stem reproduceren van de toonhoogte van een toon die klinkt. Hierbij gaat het niet enkel over het meten van een goede zangstem, maar over de vraag of de leerling in staat is de toonhoogte of frequentie van de toon juist te zingen, zonder dat deze te hoog of te laag is. De rubric is zo ingedeeld dat fase 1 het laagste niveau is en fase 4 het hoogst haalbare.

4. SINGING SELF-ASSESSMENT

(researcher-developed measurement & researcher-as-acoustic measurement)

De *Singing Self-assessment*, zelfbeoordeling van leerlingen, is ontwikkeld door Riley (2010). Het beoordelingsinstrument bestaat uit twee onderdelen (Tabel 4, zie bijlage); de zelfbeoordeling van de leerlingen en de beoordeling van drie docenten aan de hand van dezelfde vragenlijst. De leerlingen ontwikkelen zelf criteria voor hun beoordeling; correcte liedtekst, correct gebruik van de zangstem, correct zingen van toonhoogte en het in tempo

zingen. De leerlingen beoordelen zichzelf aan de hand van een checklist, waar bij één onderdeel is gekozen voor een keuzevraag met drie opties waarvan de leerlingen één keuze omcirkelen.

5. BOARDMAN SCALE

(researcher-developed measurement & researcher-as-acoustic measurement)

De *Boardman scale* (Boardman,1997) is een beoordelingsinstrument waarbij gebruik gemaakt wordt van drie taken; een kaartspel, een 'geluiden pot' en een zangspel, activiteiten die makkelijk in te zetten zijn door de docent. Het doel van deze activiteiten is om te kijken of vocaal zwakkere leerlingen uit de lagere groepen door middel van deze technieken beter gaan zingen en daarnaast hun eigen zangprestaties kunnen beoordelen. De activiteiten duren telkens vijftien minuten en worden enkele weken herhaald. Het scoringsinstrument dat hierbij gebruikt wordt is de *Boardman Scale*; deze bestaat uit 7 niveaus, waarbij niveau 7 het hoogst haalbare is. Naast de *Boardman scale* (Tabel 5, zie bijlage) voor de docent, wordt de leerlingen gevraagd om een zelfbeoordeling uit te voeren. De leerling wordt gevraagd of ze de taak correct (+), niet-correct (-) of geen mening (o) hadden uitgevoerd. Als dit overeenkomt met de beoordeling van de docent heeft de leerling dit met goed gevolg behaald.

De deelvraag die in deze paragraaf werd behandeld was: wat voor verschillende soorten beoordelingsinstrumenten zijn er gepubliceerd voor het zangonderwijs in het P.O.? Uit meer dan twintig jaar onderzoek over beoordeling van muziekonderwijs kan worden opgemaakt dat met vijf beoordelingsinstrumenten, er weinig beoordelingsinstrumenten zijn ontwikkeld voor het zangonderwijs. De instrumenten die zijn ontwikkeld, zijn allen rubrics met uitzondering van de *Singing Self Assessment* die gebruikt maakt van een checklist. Ondanks dit beperkte aanbod kan er geconstateerd worden dat de gekozen instrumenten allemaal als belangrijkste eigenschap het beoordelen van de zangstem hebben. Door het beoordelen van verschillende aspecten van de zangstem kunnen ze redelijk goed vergeleken worden met elkaar. De verschillende instrumenten geven een beeld van de beoordelingsinstrumenten die door de jaren heen ontwikkeld zijn voor het zangonderwijs op de basisscholen.

PARAGRAAF 2: Welke beoordelingscriteria, specifiek van belang voor het zangonderwijs in het primair onderwijs, worden in de instrumenten opgenomen?

Om leerlingen accuraat te beoordelen is het belangrijk dat een beoordelingsinstrument is opgebouwd uit verschillende beoordelingscriteria. Deze beoordelingscriteria specificeren dat waar de leerlingen op worden beoordeeld (Chiodo, 2001). De Stichting Leerplanontwikkeling (2015) laat in hun opgestelde leerlijn muziek, aan de hand van competenties en beoordelingscriteria, zien wat er in Nederland van de leerlingen van groep 1 t/m 8 wordt verwacht. De opgestelde competenties gaan onder andere over zelfstandig-, in groepsverband- en in canon zingen, over zuiver gearticuleerd en met expressie zingen en het kunnen uitvoeren en presenteren van een muziekstukje. Het onderzoek van

Groenendijk, Damen, Haanstra & van Bortel (2015) noemt daarnaast negen domeinen waar verschillende beoordelingscriteria onder vallen. De beoordelingscriteria van het SLO kunnen worden onderverdeeld in enkele van deze domeinen: expressie en interpretatie, technische beheersing, gehoor- en klankvoorstelling, ritmische beheersing, creativiteit, attitude en samenspel.

Aan de hand van het onderzoek van Groenendijk, Damen, Haanstra & van Bortel (2015) is onderstaande tabel (Tabel 2) ontwikkeld. Deze laat zien welke beoordelingscriteria, specifiek voor het zangonderwijs, onder de verschillende domeinen vallen.

Tabel 2.

Domeinen	Beoordelingscriteria
Expressie/Interpretatie	articulatie, dynamiek, frasering, dictie, expressiviteit, individuele toon, muzikale vloeiendheid, gevoel voor timing, melodische- en ritmische ontwikkeling, muzikale interpretatie
Technische beheersing	adem controle, ademhaling, ademondersteuning, articulatie, noot nauwkeurigheid en intonatie, ritme- en nootnauwkeurigheid, tempo, stembereik, zangkwaliteit
Gehoer- en klankvoorstelling	aandachtig luisteren, correctheid van toonhoogte, klankmenging, klankrijk, toonkwaliteit, zuiverheid, toon-intonatie, nauwkeurigheid en precieze toonhoogte
Ritmische beheersing	metronomisch tempo, nauwkeurigheid ritme, gevoel voor timing, ritmische nauwkeurigheid, stabiel tempo
Attitude	doorzettingsvermogen, enthousiasme, houding, nauwkeurigheid van uitvoering, overtuiging, zekerheid
Creativiteit	originaliteit, innovatief gebruik van ritmische elementen
Samenspel	balans, harmonie, klankmenging

In Tabel 3 staat beschreven welke van de beoordelingscriteria, specifiek voor het zangonderwijs, uit het onderzoek van Groenendijk, Damen, Haanstra & van Bortel (2015) in de beoordelingsinstrumenten van dit onderzoek zijn opgenomen.

Tabel 3.

Criteria Expressie & Interpretatie	Instrumenten	Criteria Technische beheersing	Instrumenten	Criteria Gehoor- en klankvoorstelling	Instrumenten
Articulatie	4	Ademcontrole	x	Aandachtig luisteren	x
Dynamiek	x	Ademhaling	x	Correctheid van toonhoogte	1, 4, 5
Frasering	3	Ademondersteuning	x	Klankmenging	x
Dictie	4	Articulatie	x	Klankrijk	1, 3
Expressiviteit	x	Noot nauwkeurigheid en intonatie	1, 3, 4, 5	Toonkwaliteit	1, 2, 3, 4, 5
Individuele toon	x	Ritme en nootnauwkeurigheid	x	Zuiverheid	4
Muzikale vloeiendheid	x	Tempo	x	Toonintonatie	x
Gevoel voor timing	4	Stembereik	2, 3, 4	Nauwkeurigheid en precieze toonhoogte	1, 2, 3, 4, 5
Melodische en ritmische ontwikkeling	3	Zangkwaliteit	1, 3, 4, 5		
Muzikale interpretatie	x				

Criteria	Instru	Criteria	Instru	Criteria	Instrument	Criteria Samenspel	Instrum
Ritmische beheersing	ment en	Attitude	ment en	Creativiteit	en		enten
Metronomisch tempo	x	Doorzettingsvermogen	x	Originaliteit	x	Balans	x
Nauwkeurigheid ritme	x	Enthousiasme	x	Innovatief gebruik van ritmische elementen	x	Harmonie	x
Gevoel voor timing	4	Houding	x		x	Klankmenging	x
Ritmische nauwkeurigheid	x						
Stabiel tempo	x						

1= Singing Voice Development Measure

2= The Singing Voice Rubric

3= Vocal Pitch Matching Development

4= Singing Self-Assessment

5= The Boardman Scale

In deze paragraaf werd antwoord gezocht op de vraag: welke beoordelingscriteria, specifiek van belang voor het zangonderwijs in het primair onderwijs, worden in de instrumenten opgenomen? In Tabel 3 is te zien welke beoordelingscriteria uit het onderzoek van Groenendijk, Damen, Haanstra & van Boxtel (2015) in de instrumenten worden opgenomen. Wat opvalt is dat de domeinen attitude, creativiteit en samenspel in het geheel niet zijn opgenomen in de beoordelingsinstrumenten. Daarentegen zijn technische beheersing en gehoor- en klankvoorstelling zeer overheersend. Expressie en interpretatie komen voor in twee instrumenten; de *Vocal Pitch Matching Development* en de *Singing Self Assessment*. Ritmische beheersing komt met één criterium, gevoel voor timing, enkel voor in de *Singing Self Assessment*. Wordt er specifiek naar de beoordelingscriteria gekeken, dan zijn stembereik, zangkwaliteit, nootnauwkeurig en intonatie, correctheid van toonhoogte, klankrijk en toonkwaliteit de meest voorkomende beoordelingscriteria in de gekozen instrumenten.

Echter, wanneer er wordt gekeken naar de competenties van het SLO (2015), worden alle verschillende domeinen uit het onderzoek van Groenendijk, Damen, Haanstra & van Boxtel (2015), en de daaraan gekoppelde beoordelingscriteria, hierin wel opgenomen. Er zou voorzichtig kunnen geconcludeerd worden dat de gebruikte beoordelingsinstrumenten in dit onderzoek niet volledig aansluiten bij de richtlijnen voor zangonderwijs van het SLO. Doordat sommige belangrijke beoordelingscriteria, zoals dynamiek, expressiviteit en houding ontbreken in de beoordelingsinstrumenten, kunnen deze niet volledig toetsen wat er belangrijk is in het zangonderwijs op de basisschool. Om die reden zou ter discussie kunnen worden gesteld, of deze instrumenten voldoende geschikt zijn voor het zangonderwijs in het primair onderwijs.

PARAGRAAF 3: Wat is de kwaliteit van deze instrumenten in termen van validiteit, betrouwbaarheid, bruikbaarheid en authenticiteit?

Bij het meten van de kwaliteit van beoordelingsinstrumenten, is het belangrijk te bekijken of deze betrouwbaar en valide is (Salvador, 2010). Volgens Salvador refereert de *betrouwbaarheid* van het instrument naar de consistentie van de resultaten; een instrument is betrouwbaar wanneer een leerling bijna hetzelfde zou scoren indien de toets nog een keer wordt afgenomen, of wanneer de uitvoering van de taak door iemand anders zou worden beoordeeld. Betrouwbaarheid is noodzakelijk om een instrument als valide te beschouwen; dit verwijst naar nauwkeurigheid en consistentie. Om de beoordeling van een zangprestatie als valide te beschouwen, moet het ook daadwerkelijk het zangvermogen van het kind meten. Daarnaast moet de toetsvorm aansluiten bij de inhoud van de taak; wanneer een leerling wordt getoetst op zijn zangprestaties is het logisch dat dit met een praktische toets wordt beoordeeld en niet met een schriftelijke toets, dit is de *validiteit* van de toets. Tevens is er de eis van praktische *bruikbaarheid*, dat wil zeggen dat de toetsing bij voorkeur haalbaar en efficiënt is en dat het voor de docent mogelijk is om de toets af te nemen in het gegeven tijdsbestek (van der Schaaf & Wessels, 2014).

De ontwikkeling van betrouwbare en valide instrumenten voor kunstvakken kent daarbij ook vele uitdagingen (Oreck, Owen & Baum, 2003). Zo is er het verschil onder leerlingen wat betreft de toegang tot artistieke bronnen en ervaringen, het onderlinge verschil in verkregen kunstonderwijs en de verschillende manieren waarmee leerlingen met presenteren en beoordelen omgaan. Verder maken de culturele verschillen, de verschillende leefomstandigheden en populaties het ingewikkeld om beoordeling valide te maken. Om deze problemen te voorkomen is het belangrijk dat de instrumenten worden ontwikkeld voor een grote verscheidenheid aan leefomgevingen en leerling populaties. Voor een *authentieke beoordeling* moeten er in het proces dan ook authentieke kunstervaringen worden aangeboden in een omgeving die de artistieke gedragingen stimuleert (Oreck, Owen & Baum, 2003). Daarnaast vindt muziekonderwijs meestal plaats in groepen; om het zangonderwijs authentiek te kunnen beoordelen is het dus van belang dat de instrumenten ook beoordelen op samenzang (Leighton, 2006).

Of een instrument authentiek is, kan bekeken worden aan de hand van vijf facetten (Gulikers, 2007);

1. De taak: de opdracht die moet worden uitgevoerd.
2. De fysieke context: de omgeving waarin de taak wordt uitgevoerd.
3. De sociale context: samenwerkingsvormen en de rol van docent/leerling tijdens de taakuitvoering.
4. De vorm: de beoordelingsmethodes, los van de inhoud.
5. De criteria: de punten waarop de leerlingen worden beoordeeld.

Onderstaande tekst laat zien in hoeverre de gekozen instrumenten, aan de hand van reviews van verscheidene auteurs, voldoen aan de termen validiteit, betrouwbaarheid, bruikbaarheid en authenticiteit.

1. THE SINGING VOICE DEVELOPMENT MEASURE (SVDM)

Validiteit: Salvador (2010) stelt dat de validiteit is vastgesteld in verschillende onderzoeken, maar gaat verder niet in op deze onderzoeken. Volgens haar laat het duidelijk diagnostische informatie zien die aantoont wat een leerling moet leren om de zangprestaties te laten toenemen.

Betrouwbaarheid: Welch et al. (2011) gebruikte de SVDM bij het Nationale zangprogramma *Sing Up* in Engeland. Ze gebruikte hier zowel SVDM als de *Vocal Pitch Matching Development*. Om beide instrumenten meer betrouwbaar te maken begonnen de onderzoekers eerst met een pilot: twee leden van het onderzoeksteam gingen naar een school en maakten geluidsopnames van verschillende leerlingen van verschillende leeftijden. Anoniem beoordeelde zij, en andere leden van het team, de leerlingen aan de hand van de twee scoringsinstrumenten. Vervolgens werden de uitkomsten met elkaar vergeleken en bleken deze niet veel van elkaar te verschillen. Hieruit kan verondersteld worden dat dit instrument betrouwbaar is. In 1998 gebruikte Levinowitz bij zijn onderzoek de SVDM. Met zes verschillende beoordelaars die het instrument gebruikten lagen de resultaten zeer dicht bij elkaar. Door deze uitkomst kan het instrument als betrouwbaar worden gezien.

Bruikbaarheid: Volgens Salvador (2010) is SVDM praktisch en helder. Daarnaast is SVDM oorspronkelijk ontworpen voor een specifieke set tonaal materiaal, maar deze kan ook gebruikt worden voor het beoordelen van zangprestaties bij algemene liedjes. Het instrument kan derhalve goed gehanteerd worden in de klas door een vakleerkracht en is hiermee bruikbaar (Salvador, 2010).

Authenticiteit: Het beoordelingsinstrument is niet authentiek op het gebied van de sociale- en fysieke context (Gulikers, 2007); leerlingen worden individueel beoordeeld in een stille ruimte, of met een klein groepje (2-4 leerlingen) buiten het klaslokaal (Salvador, 2010).

2. THE SINGING VOICE RUBRIC

Er is geen literatuur gevonden die de *Singing Voice Rubric* bekijkt aan de hand van validiteit, betrouwbaarheid, authenticiteit en bruikbaarheid. Onderstaande bevindingen zijn hierdoor niet volledig.

Validiteit: De taak van dit beoordelingsinstrument is 'solo zingen in de klas', waarbij gemeten wordt hoe de zangkwaliteit van de solozanger is. De rubric meet de zangkwaliteit van de zanger. Er wordt getoetst wat er getoetst moet worden, daarmee is dit instrument valide.

Betrouwbaarheid: Doordat er geen reviews zijn te vinden over het gebruik van deze rubric, en daarmee dit instrument niet aantoonbaar door verschillende beoordelaars is gebruikt, is het moeilijk vast te stellen of het instrument betrouwbaar is.

Bruikbaarheid: Het beoordelen kan in korte tijd plaatsvinden, aangezien de beoordelaar kan beoordelen tijdens het uitvoeren van de taak. Hiermee maakt dit het instrument bruikbaar.

Authenticiteit: De taak wordt afgenomen in het klaslokaal, waarbij individuele leerlingen worden beoordeeld terwijl de rest van de klas ook deelneemt. Door deze authentieke situatie, op het gebied van de sociale- en fysieke context (Gulikers, 2007), met de gehele klas zingen in het klaslokaal, kan dit instrument op dit gebied authentiek worden genoemd.

3. VOCAL PITCH MATCHING DEVELOPMENT (VPMD)

Validiteit: De taak van dit instrument is het zingen van een bekend lied. Vervolgens worden de leerlingen in het scoringsinstrument beoordeeld op de woorden van het lied, de melodie, toonhoogte en melodische zinnen. Al deze vier aspecten zijn gericht op het zingen van een lied, daarmee kan dit instrument als valide worden beschouwd (Welch et al., 2011).

Betrouwbaarheid: De betrouwbaarheid van dit instrument is vastgesteld tijdens de pilot fase van het Nationale zangprogramma *Sing Up* in Engeland, hierboven al eerder beschreven bij de *Singing Voice Development Measure* (Welch et al., 2011).

Bruikbaarheid: Om de bruikbaarheid onder docenten te vergroten, werden in het onderzoek van Welch et al. (2011) de docenten eerst getraind en mochten ze het beoordelingsinstrument uittesten, alvorens ze het onderzoek het daadwerkelijk uitvoerden.

Authenticiteit: Dit instrument beoordeelt leerlingen individueel op zang. Deze context verschilt van het klaslokaal, de fysieke- en sociale context (Gulikers, 2007), en van de beoordelingsmogelijkheden van muziekdocenten op school, waar leerlingen in grote groepen zingen. Hierdoor is dit instrument op dit punt weinig authentiek (Leighton, 2006). Aan de andere kant wordt voor een lied gekozen uit de 'lokale cultuur', waardoor de leerlingen iets zingen dat zij al kennen, de taak is hiermee authentiek (Welch et al., 2011).

4. SINGING SELF-ASSESSMENT

Validiteit: De leerlingen ontwikkelen zelf criteria voor de beoordeling (Riley, 2010). Hierdoor weten de leerlingen precies waar ze op beoordeeld worden en is het beoordelingsinstrument valide te noemen.

Betrouwbaarheid: De resultaten van de beoordeling van docenten en die van de leerlingen lagen zeer dicht bij elkaar (Riley, 2010). Door de consistentie van deze resultaten kan dit instrument als betrouwbaar worden beschouwd.

Bruikbaarheid: Het beoordelingsinstrument bevat een korte vragenlijst en is snel in te vullen door zowel docent als leerling, hierdoor is dit instrument bruikbaar (Riley, 2010).

Authenticiteit: De leerlingen ontwikkelen hun eigen criteria (Riley, 2010), hiermee is het instrument op het gebied van criteria (Gulikers, 2007) authentiek. Daarnaast beoordelen de leerlingen zichzelf, waardoor de vorm van het instrument ook authentiek kan worden genoemd. Tijdens de taak werken de leerlingen in tweetallen, waarbij zij elkaar sturen en tips geven. Door dit groepswork is dit instrument op het gebied van sociale context authentiek. Een deel van de taak wordt uitgevoerd in het klaslokaal waardoor de fysieke context hiermee enigszins authentiek is.

5. THE BOARDMAN SCALE

Validiteit: Salvador (2010) stelt dat het instrument niet zo betrouwbaar is als de computergestuurde akoestische meting van toonhoogte. Maar dat het een meer valide meting van de zangprestaties is, omdat het diagnostische informatie toont die gebaseerd is op een authentiek zangstukje. Dit helpt de docent om erachter te komen wat een leerling moet leren om een betere zanger te worden. Afgezien van deze bevinding is het volgens Salvador aan te raden voor vakleerkrachten, om andere scoringsinstrumenten te gebruiken met heldere criteria en een hogere validiteit.

Betrouwbaarheid: Bij een onderzoek van Mathias (1997) werden verschillende leerlingen aan de hand van de Boardman Scale door drie beoordelaars beoordeeld; in 48% van de gevallen waren de beoordeelaars het compleet met elkaar eens, in 80% maar op één punt.

Deze uitkomst is waarschijnlijk te wijten aan de verwarring die is ontstaan bij de beschrijving van de punten; hoe beoordeel je een leerling die exact het patroon uitvoert (punt 5) maar ondertussen in toonsoorten schommelt (punt 6)? En hoeveel moet een leerling variëren in toonsoorten om van punt 7 naar punt 6 af te glijden? Hiermee maakt dit de Boardman Scale niet geheel betrouwbaar (Salvador, 2010).

Bruikbaarheid: Er zijn drie verschillende taken die elk met een apart scoringsinstrument moeten worden beoordeeld (Boardman, 1997), dit is tijdrovend voor de docent. Wanneer het scoringsinstrument alleen kan worden gebruikt in combinatie met deze taken, is het wellicht niet bruikbaar genoeg voor in het klaslokaal.

Authenticiteit: Bij de Boardman Scale (Boardman, 1997) wordt individueel met de leerlingen gewerkt, qua sociale context (Gulikers, 2007) is dit niet zeer authentiek.

In deze deelvraag werd gekeken naar de kwaliteit van de instrumenten in termen van validiteit, betrouwbaarheid, authenticiteit en bruikbaarheid. Wat opvalt is dat het begrip authenticiteit in de literatuur weinig wordt genoemd in relatie tot de gekozen instrumenten. Oreck et al. (2003) benadrukken het belang van authenticiteit, maar vrijwel geen van de reviews die zijn gevonden voor dit onderzoek bekijken de authenticiteit van een instrument. De vraag rest dan ook of authenticiteit überhaupt als criterium kan worden opgenomen in relatie tot deze beoordelingsinstrumenten. Daarnaast valt op dat auteurs de termen validiteit en betrouwbaarheid vaak door elkaar heen gebruiken, waardoor het soms lastig is om de instrumenten op deze afzonderlijke punten te bekijken. Aan de hand van de gevonden literatuur kan geconcludeerd worden dat *The Singing Voice Development Measure* en de *Vocal Pitch Matching Development* het beste scoren op het gebied van validiteit, betrouwbaarheid en bruikbaarheid. De *Singing Self-Assessment* is het meest authentieke instrument. Daarnaast zijn de *Singing Voice Development Measure* en de *Vocal Pitch Matching Development* het meest gebruikt in vervolg onderzoeken, waardoor de kwaliteit van deze instrumenten het sterkst onderbouwd is. Door het ontbreken van literatuur bij The Singing Voice Rubric scoort dit instrument het laagst op bovenstaande punten.

PARAGRAAF 4: Op welke manier kunnen de beschreven instrumenten worden ingezet om het leren van leerlingen te stimuleren?

Beoordeling van het muziekcurriculum is zowel een noodzakelijk als krachtig middel binnen het onderwijs (Shuler, 2011). Veel van de leerling activiteiten vinden plaats tijdens de muziekles, maar het moment van beoordelen laat zien wat er geleerd is en hoe het leren heeft plaatsgevonden (McGuire, 2002).

Volgens Scott (2012) zijn er drie soorten manieren van beoordelen; beoordelen van het geleerde, beoordelen voor het leren en beoordelen *als* leren. *Beoordelen van het geleerde* is een traditionele manier van beoordelen waarbij de leerling een passief lerende is die een taak moet voltooien die door de docent wordt geleverd. De student wordt beoordeeld op het begrijpen en uitvoeren van de instructies van de leerkracht. Bij het *beoordelen voor het leren* zijn studenten actief lerenden waarbij beoordeling hun verder

helpt in het leerproces, de student staat centraal. De student gebruikt de feedback van de docent om zijn/haar huidige niveau te verbeteren. Bij het *beoordelen als leren* vindt beoordelen plaats aan de hand van zelfreflectie door de student. Studenten monitoren hun eigen leerproces en reflecteren op dat wat ze bereikt hebben. Als actief lerende zijn ze betrokken in het beoordelingsproces.

Scott (2012) stelt dat, *beoordelen voor het leren* en *beoordelen als leren* beide bijdragen aan het actief betrekken van leerlingen bij hun onderwijs. *Beoordelen van het geleerde* zit volgens Scott diep geïntegreerd in onze onderwijs tradities, maar betreft de leerlingen minder bij het leerproces. Om juist de minder gebruikelijke manieren van beoordelen te implementeren in het onderwijssysteem is er verandering nodig. De manier waarop scoringsinstrumenten worden gebruikt, kan hier ook aan bijdragen. Wanneer gebruik gemaakt wordt van rubrics, numerieke schalen of checklists bij het beoordelen van prestaties, kan dit als *beoordelen van het geleerde* worden ingezet wanneer de uitkomst van deze instrumenten wordt gebruikt als eindcijfer voor de leerling. Wanneer de instrumenten worden ingezet om bovenal de leerlingen feedback te geven om hun prestaties te verbeteren, wordt er ingezet op *beoordelen voor het leren*. Daarnaast kunnen de leerlingen hierbij ook meedenken in de ontwikkeling van de rubric. *Beoordelen als leren* wordt bereikt wanneer de leerlingen het instrument mogen gebruiken als hulpmiddel voor het proces van zelfreflectie en ontwikkeling. Uiteraard hebben leerlingen nog altijd sturing nodig in het leren te beoordelen en zal een docent minder van een kleuter kunnen verwachten dan van een leerling uit een hogere groep.

De scoringsinstrumenten die in dit onderzoek uiteen worden gezet, zijn allen rubrics of checklists. Om het leren van de leerlingen te stimuleren bij het gebruik van deze instrumenten, is het derhalve aan te raden om ze in te zetten bij *beoordelen voor het leren* en *beoordelen als leren*. Om dit te behalen kunnen de instrumenten allereerst worden ingezet als een feedbackmoment in plaats van een eindoordeel door de docent. Met andere woorden, er wordt dan formatief en niet summatief getoetst. Daarnaast is er de mogelijkheid om de leerlingen mee te laten denken over eventuele toevoegingen of veranderingen aan de bestaande scoringsinstrumenten, dit om leerlingen actiever te betrekken bij hun onderwijs. Bij de oudere leerlingen zouden de instrumenten kunnen worden gebruikt bij zelfreflectie en ontwikkeling, waarbij de docent geen overheersende rol speelt in de beoordeling; de student staat zo centraal en wordt verder geholpen in zijn leerproces.

PARAGRAAF 5: Conclusie en Discussie

Paragraaf 5.1: Conclusie

In dit onderzoek is geprobeerd om antwoord te vinden op de vraag: welke beoordelingsinstrumenten voor het zangonderwijs in het primair onderwijs zijn er in de literatuur beschreven en hoe verhouden deze zich tot elkaar?

In de eerste deelvraag is gekeken naar de verschillende beoordelingsinstrumenten zoals beschreven in de literatuur. Daar zijn in totaal vijf instrumenten gevonden, waarvan de scoringsinstrumenten viermaal een rubrics zijn en eenmaal een checklist. De instrumenten beoordelen vrijwel allemaal dezelfde aspecten van zangonderwijs; zangstem versus spreekstem, het correct zingen van toonhoogte, het stembereik en het zingen van de correcte toonsoort. Hierdoor is er veel overlap tussen de verschillende instrumenten.

Deelvraag twee bekeek de verschillende beoordelingscriteria van de beoordelingsinstrumenten en hoe deze zich verhouden tot de competenties die zijn vastgesteld door het SLO (2015) voor het zangonderwijs in het primair onderwijs. Daaruit kan geconcludeerd worden dat de instrumenten niet goed aansluiten bij deze competenties en dat er veel aspecten van het beoordelen van zangonderwijs ontbreken; aspecten als expressie, houding, creativiteit en samenspel komen niet aan bod in de gevonden beoordelingsinstrumenten.

In deelvraag 3 werd gekeken naar de kwaliteit van de beoordelingsinstrumenten in termen van validiteit, betrouwbaarheid, bruikbaarheid en authenticiteit. In de literatuur is het meest geschreven over de *Singing Voice Development Measurement* (Rutkowski, 1998) en de *Vocal Pitch Matching Development* (Welch, 1998) en worden deze twee instrumenten op bovenstaande punten, met uitzondering van authenticiteit, ook het meest geprezen. Weinig van de gevonden instrumenten zijn geschikt voor in het klaslokaal en de leerlingen worden vrijwel altijd individueel en buiten het klaslokaal beoordeeld. Hiermee ontbreekt het vaak aan authenticiteit. Het instrument dat hoog scoort op authenticiteit is de *Singing Self Assessment* (Riley, 2010) waar leerlingen zelf een deel van de beoordelingscriteria opstellen en waar een groot deel van de beoordeling plaatsvindt in het klaslokaal.

In de laatste deelvraag werd gekeken naar hoe de beschreven instrumenten kunnen worden ingezet om het leren van leerlingen te stimuleren. Scott (2012) maakt een onderscheid tussen drie soorten manieren van beoordelen: beoordelen van het geleerde, beoordelen voor het leren en beoordelen als leren. Concluderend zijn beoordelen voor- en als leren het meest efficiënt om het leren van leerlingen te stimuleren. Wanneer hier ruimte voor wordt geboden in het beoordeelproces, en wordt ingezet in combinatie met de beoordelingsinstrumenten, is dit de meest effectieve manier om het leren te stimuleren.

Een verbinding van bovenstaande deelvragen leidt tot de volgende conclusie. Vastgesteld kan worden dat de aspecten zangstem versus spreekstem, het correct zingen van toonhoogte, het stembereik en het zingen van de correcte toonsoort, belangrijke beoordelingscriteria zijn in alle ontwikkelde instrumenten. De gevonden instrumenten beoordelen veelal dezelfde aspecten van het zangonderwijs. Daarnaast kan geconcludeerd worden dat er in de afgelopen twintig jaar weinig bruikbare instrumenten zijn ontwikkeld en dat de ontwikkelde instrumenten weinig variatie bevatten. In recent mailcontact met

Professor Welch (persoonlijke communicatie, 29 maart 2017) bevestigt hij het gebrek aan bruikbare instrumenten voor het zangonderwijs. In een door hem recent uitgevoerd onderzoek in Australië maakt hij gebruik van zijn eigen beoordelingsinstrument dat hij ontwikkelde in 1998. Wanneer er twintig jaar na dato nog steeds gebruik wordt gemaakt van dit instrument, is het de vraag of de ontwikkelde instrumenten nog accuraat genoeg zijn in het huidige leerklimaat en of het geen tijd is voor een nieuw instrument. Daarnaast zijn veel van de beoordelingsinstrumenten veelvuldig ingezet in onderzoek naar het zangonderwijs, maar zijn de beoordelingsinstrumenten die bruikbaar zijn voor het klaslokaal beperkt.

Paragraaf 5.2: Discussie

Concreet kan deze conclusie beschouwd worden als een aanmoediging voor onderzoekers en muziekdocenten om beoordelingsinstrumenten te gaan ontwikkelen waarin de ontbrekende beoordelingscriteria, in combinatie met de reeds gebruikte criteria, ook aan bod komen. Zoals reeds in de inleiding is beschreven is het belangrijk om het beoordelen meer in te zetten in het klaslokaal, omdat met geschikte beoordelingsinstrumenten muziek meer onderdeel kan worden van het kerncurriculum van de school. Daarom is het aan te bevelen de beoordelingsinstrumenten te ontwikkelen voor gebruik in het klaslokaal.

Deze conclusie laat ook ruimte voor discussie; een aantal belangrijke punten worden hier genoemd. Als er wordt gekeken naar de beoordelingscriteria in de verschillende instrumenten valt op dat er veel woorden gebruikt worden als 'af en toe', 'enige', 'voornamelijk' en 'goede zanger'. Maar wanneer is iemand een 'goede zanger' en wat is 'af en toe'? Dit zijn termen die door verschillende beoordelaars op verschillende manieren geïnterpreteerd kunnen worden, waardoor de instrumenten op het gebied van validiteit en betrouwbaarheid minder hoog zouden kunnen scoren. Zijn er woorden die dit probleem uitsluiten of zal dit altijd een knelpunt blijven bij beoordeling? Het is daarom raadzaam om de instrumenten in een eventueel vervolgonderzoek op deze punten nogmaals aandachtig te bestuderen en eventueel aan te passen.

Een ander punt is de vraag wat de volgende stap zou kunnen zijn na dit onderzoek. Wellicht is het belangrijk om te bekijken hoe we deze instrumenten kunnen integreren in het onderwijs en hoe ervoor gezorgd kan worden dat ze niet blijven liggen maar écht onderdeel van het curriculum worden. Het is daarom aan te bevelen om onderzoek te doen naar de inzetbaarheid van de beschreven beoordelingsinstrumenten in het primair onderwijs.

Wanneer de verschillende beoordelingsinstrumenten naast elkaar worden gezet, rest de vraag welke van deze instrumenten nu daadwerkelijk geschikt zouden zijn voor het primair onderwijs. Zoals reeds in de inleiding is geschreven is het belangrijk dat een beoordelingsinstrument niet tijdrovend is. Naar mijn mening zijn de *Singing Voice Rubric* van Chiodo (2010) en de *Singing Self Assessment* van Riley (2010) het minst tijdrovend en daarmee het meest bruikbaar in de klas. Maar er spelen nog veel meer factoren mee in de keuze van een beoordelingsinstrument; wie beoordelen er, wordt het proces of product beoordeeld, welk type taakinstrument wordt gebruikt? Het is daarom ingewikkeld om één beoordelingsinstrument te kiezen, vandaar dat ik aanraad meerdere van de genoemde beoordelingsinstrumenten naast elkaar te gebruiken.

BIJLAGEN

Tabel 1. The Singing Voice Development Measure by J. Rutkowski(1998)

1	"Pre-zanger": zingt niet maar spreekt de tekst
1.5	Inconsistente "spraakstem-omvang zanger": zingt af en toe, behoudt af en toe tonen, verklankt af en toe toonhoogte, maar blijft voornamelijk in de spraakstem omvang (meestal A2 tot C3)
2	"Sprakstem-omvang zanger": behoudt tonen en vertoont enig gevoel voor toonhoogte, maar blijft voornamelijk in de spraakstem omvang (meestal A2 tot C3)
2.5	Inconsistente "gelimiteerde-omvang zanger": varieert tussen spreek- en zangstem en gebruikt een gelimiteerde omvang bij de zangstem (meestal tot F3)
3	"Gelimiteerde-omvang zanger": vertoont het gebruik van een gelimiteerde zang-omvang
3.5	Inconsistente "beginnend juiste omvang- zanger": vertoont soms alleen het gebruik van een beperkte zangomvang, maar op andere momenten vertoont hij/zij het gebruik van een beginnend juiste omvang van de zangstem (meestal D3 tot A3)
4	"Beginnend juiste omvang-zanger": vertoont het gebruik van een beginnend juiste omvang van de zangstem (meestal D3 tot A3)
4.5	Inconsistente "zanger": vertoont soms alleen het gebruik van een beginnend juiste-omvang van de zangstem, maar op andere momenten vertoont hij/zij het gebruik van een uitgebreide omvang van de zangstem (van B3b en daarboven)
5	"Zanger": vertoont het gebruik van een uitgebreide omvang van de zangstem (zingt B3(b) en daarboven)

Aangepast van "Assessing a cast of thousands," door P. Chiodo, 2001, *Music Educators Journal*, 87(6), p. 23.

Tabel 2. The Singing Voice Rubric by P. Chiodo (2001)

5: zanger	De leerling tilt de zangstem boven het hogere bereik (boven Bb) en maakt gebruik van het hoogste register van de stem (de kopstem)
4: correcte zangstem-omvang zanger	De leerling gebruik het midden register van zijn/haar stem en zingt het lied in de bedoelde toonomvang (D-A). De leerling haalt alleen het hogere bereik (boven Bb) nog niet.
3: gelimiteerde zangstem-omvang zanger	De leerling gebruik het midden register (lichte kwaliteit) maar zingt in een vrij smalle omvang (meestal D-F#)
2: spraak-omvang zanger	De leerling zingt in het lage (borst) register, van een lage A tot een middel C, de omvang van de spreekstem
1: 'pre'-zanger	De leerling spreekt de woorden uit het lied in plaats van deze te zingen

Aangepast van "How can elementary teachers measure singing voice achievement? A critical review of assessment, 1994-2009," door K. Salvador, 2010, *The National Association for Music Education*, 29(1), p. 44.

Tabel 3. The Vocal Pitch Matching Development by G. Welch (1998)

Fases	
1	De woorden uit het lied blijken aanvankelijk belangrijk voor de leerling te zijn dan de melodie, het zingen kan worden omschreven als 'gesproken', beperkt stembereik en beperkte muzikale zinnen. Infantiele uitbreiding van het stembereik, neerdalende patronen overheersen.
2	Er is bewustwording bij de leerling dat stembereik een continu proces is en dat de veranderingen in stembereik te controleren zijn. Gezongen melodieën beginnen langzaam de algemene contouren van de opgegeven melodie of toonsoort te volgen. Zelfverzonnen en 'schematische' liedjes vertonen overeenkomsten met de muzikale belevingswereld van het kind. Het stembereik, gebruikt in de 'liedjes' verbreed zich.
3	Melodieën en intervallen zijn overwegend foutloos gezongen, maar af en toe kan er iets in de tonaliteit veranderen, misschien heeft dit te maken met het niet goed gebruiken van de verschillende stemregisters. Over het geheel genomen, is het aantal verschillende gebruikte toonsoorten afgenomen.
4	Geen significante melodieuze of tonale fouten in relatie tot de relatief simpele liedjes uit de muzikale belevingswereld van het kind.

Aangepast van "Researching the first year of the National Singing Programme *Sing Up* in England: An initial impact evaluation," door Welch et al., 2011, *Psychomusicology: Music, Mind & Brain*, 21(1), p. 88.

Tabel 4. Singing Self Assessment by P.E. Riley (2010)

Naam:	
Groep:	
Zelfbeoordeling Zang ✓ = check	
1. Ik zong de correcte woorden	
2. Ik gebruikte mijn zangstem	
3. Ik zong de juiste toonhoogte of ik klonk als het voorbeeld	Omcirkel: Te hoog Goeie hoogte Te laag
4. Ik zong in het juiste tempo	

Aangepast van "Singing Self-Assessment Accuracy of Elementary School Children" door P.E. Riley, 2010, *The practice of assessment in music education; frameworks, models and design*, p. 315.

Tabel 5. The Boardman Scale by J. Boardman (1997)

Niveau	
7	Accuraat zingen van alle tonen van het patroon zonder aarzeling
6	De leerling zingt het patroon in één of meerdere toonsoorten
5	De leerling voert een exacte transpositie van het patroon uit
4	De leerling behield een algemene omvang van het patroon, maar zong foute intervallen
3	De leerling behield een algemene richting in het patroon, maar niet in de exacte omvang
2	De leerling reageerde niet correct op de patronen door het negeren van de omvang van de toonhoogte
1	De leerling sprak meer dan dat hij/zij zong of reageerde helemaal niet

Aangepast van "How can elementary teachers measure singing voice achievement? A critical review of assessment, 1994-2009," door K. Salvador, 2010, *The National Association for Music Education*, 29(1), p. 43

Literatuurlijst

- Chiodo, P. (2001). Assessing a cast of thousand. *Music Educators Journal*, 87(6), 17-23. doi: 10.2307/3399687
- DeLuca, C. & Bolden, B. (2014). Music performance assessment: Exploring three approaches for quality rubric construction. *Music Educators Journal*, 101(1), 70-76. doi: 10.1177/0027432114540336
- Dunbar, L. (2012). Performance assessment of the masses in 30 seconds or less. *General Music Today*, 25(2), 31-35. doi: 10.1177/1048371311406266
- Gulikers, J. (2007). Authentiek beoordelen in een curriculum. *Onderwijsinnovatie*, 9(2), 11-14. Geraadpleegd op https://www.ou.nl/Docs/TijdschriftOI/OI_juni_2007.pdf
- Groenendijk, T., Damen, M., Haanstra, F., & van Boxtel, C. (2015). Beoordelingsinstrumenten in de kunstvakken - een review. Geraadpleegd op <http://www.lkca.nl/kennisdossiers/beoordelen/reviewstudie-assessment-in-kunsteducatie>
- Leighton, G., & Lamont, A. (2006). Exploring children's singing development: do experiences in early schooling help or hinder? *Music Education Research*, 8(3), 311-330. doi: 10.1080/14613800600957461
- Landelijk Kenniscentrum Amateurkunsten. (n.d.) Modellen en instrumenten [artikel]. Geraadpleegd op <http://www.lkca.nl/kennisdossiers/beoordelen/beoordelingsinstrumenten>
- McGuire, K. (2002). Doing it all: Linking the elements of music and the standards. *Music Educators Journal*, 69(1), 49-52. doi: 10.2307/3399885
- 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. doi: 10.1177/016235320302700105
- Riley, P. E. (2010). Singing Self-Assessment Accuracy of Elementary School Children. In T. S. Brophy (Ed.), *The practice of assessment in music education; frameworks, models and design* (307-321). Chicago, IL: GIA Publications
- Salvador, K. (2010). How can elementary teachers measure singing voice achievement? A critical review of assessments, 1994-2009. *The National Association for Music Education*. 29(1), 40-47. doi: 10.1177/8755123310378454

- Scott, S. J. (2001). A task-centered approach to performance-based assessment. *General Music Today Spring*, 14(3), 10-14. doi: 10.1177/104837130101400304
- Scott, S. J. (2012). Rethinking the roles of assessment in music education. *Music Educators Journal*, 98(3), 31-35. doi: 10.1177/0027432111434742
- Shuler, S. C. (2011). Music education for life: Music assessment, part 1: What and why. *Music Educators Journal*, 98(2), 10-13. doi: 10.1177/0027432111427651
- Stichting Leerplanontwikkeling (2006). Rubrics als beoordelingsinstrument voor vaardigheden. Geraadpleegd op <http://www.slo.nl/downloads/rubricsals.pdf>
- Stichting Leerplanontwikkeling (2015). Leerplankader kunstzinnige oriëntatie. Geraadpleegd op <http://kunstzinnigeorientatie.slo.nl/leerlijnen/kunstzinnige-vakdisciplines-en-cultureel-erfgoed/muziek>
- Stichting Leerplanontwikkeling (n.d.). Rubrics. Geraadpleegd op <http://www.slo.nl/voortgezet/vmbo/themas/Rubrics/>
- Taalunieversum. (n.d.). Geraadpleegd op <http://taalunieversum.org/onderwijs-/termen/term/411/zaakvak/>
- Toetswijzer. (n.d.). Geraadpleegd op <http://www.toetswijzer.nl/html/competentiegerichtleren-VO/beoordelingsinstrument%202.htm>
- Welch, G. F., Himonides, E., Saunders, J., Papageorgi, I., Rinta, T., Preti, C., . . . Hill, J. (2011). Researching the first year of the national singing programme *Sing Up* in England: an initial impact evaluation. *Psychomusicology: Music, Mind & Brain*, 21(1-2), 83-97. doi:10.1037/h0094006