



The construction of an assessment instrument based on the competence model

Talita Groenendijk & Folkert Haanstra

1. Introduction

1.1 Forms of assessment

Assessing students is considered a complex issue in art education. For many art teachers the artistic process remains associated with unpredictability and individuality, whereas assessment stands for predictability and comparison. However, when a curriculum is described and organized in the form of competences, the need to operationalize what behaviour and products are expected from students, becomes necessary.

Assessment can have a summative function which means that the assessment is used to make decisions about what was learned after an educational process has finished (Brookhart, 2001). Formative assessment, on the other hand, refers to assessment which aims at improving student performance (Brookhart, 2001). It takes place in the course of the learning process rather than at the end of the process. Black and Wiliam (2009) describe three key processes involved in formative assessment: 1) the students should understand the learning goals, 2) the students evaluate their current performance and recognize the gap between this performance and the desired goals and 3) the students should know how to bridge this gap and act accordingly. Formative and summative assessment can be combined and are not necessarily opposites (Sluijsmans et al., 2013).

Assessment can be carried out by teachers, but also by students themselves or peers. The advantage of having students assess their own performance is that the assessment itself can become a tool for learning as it stimulates self-regulation and metacognition (Nicol & MacFarlane-Dick, 2006). When students learn to assess themselves they can monitor and regulate their own learning processes. Self-regulation can be considered an important goal of art education as it is required for learning after school when external feedback is not always present (Boud, 2015). Self-assessment is to some extent always present in engaged task execution, but the teachers should develop this capacity more by offering "structured opportunities" (Nicol & MacFarlane-Dick, 2006, p207).

Much of the literature on self-assessment concerns higher education. Lucas, Claxton & Spencer (2012) developed a self-assessment tool for assessing creativity in primary and secondary education and they found that engaging in self-assessment is difficult for students. Sluijsmans et al. (2013) suggest training in self-assessment is important and they list four important components in training students to assess themselves (2013, p50, after Brookhart, 2007): 1) involve students in the definition of criteria, 2) instruct them how to apply the criteria, 3) provide feedback after the self-assessment, 4) help students planning their activities based of the assessment results.

It is important then to use self- assessment in combination with teacher or peer assessment. "Feedback from teachers is a source against which students can evaluate progress, and check out their own internal constructions of goals, criteria and standards" (Nicol & MacFarlane-Dick, 2006,

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

p208). Peer assessment is often connected to self-assessment, but in general is used less frequently (Sluijsmans et al., 2013). Peer feedback may stimulate better understanding of the criteria as students express the criteria in their own words while communicating with their peers (Sadler, 1998). In art education informal peer assessment is not uncommon, as many teachers end an assignment with having the students present the visual products to each other and discussing them.

Taking into account the importance of formative assessment as well as self-assessment in art education, it was decided to construct an instrument that was aimed at these assessment tasks. However the instrument should also be suited for teacher assessment and peer assessment and potentially could be used for summative assessment. The next section describes the choice of an assessment form that can cover these different functions.

1.2 Assessment instruments

Haanstra, Damen, Groenendijk & Van Boxtel, (2015) conducted a study to review the scientific literature on art assessment instruments. The outcomes demonstrate that assessment instruments exist in many different forms and functions. The assessment task of most instruments collected in the review consists of an arts performance. Others tasks are paper and pencil tests, portfolio's and logbooks. A checklist of criteria is the most used scoring tool, but the review shows that also the rubric is currently a very popular instrument for assessment in art education in primary and secondary education. A rubric also consists of criteria, but they are stated in several different levels of competence (for instance from beginner to proficient). Rubrics are transparent for students and therefore enable students to assessment themselves and regulate their own learning. On the other hand, rubrics enable the assessment of complex tasks, such as in art education, with many dimensions and open answers which are not just true or false. Sluijsmans et al (2013) describe that rubrics are useful for summative and formative testing as they provide early and informative feedback.

Rubrics can be used for scoring and explaining scores, as an instrument to have students understand criteria, as an instrument to support self- and peer assessment and as an instrument to define the contents of education (Sluijsmans et al., 2013, p.53). Rubrics can be either general or task specific and holistic or analytic in nature. Task specific rubrics are often easy for teachers to score as indicators are very explicitly related to the task. But, on the other hand, general rubrics enable student to monitor their progress over time, over different tasks. The analytic rubric is very detailed and transparent, but the holistic rubric may provide more reliable scoring as real student performance is usually not so easily caught in very strict rubrics.

Disadvantage of the rubric in general is that it is time consuming to construct such a rubric. Another disadvantage can be the textual character of the rubric. Oskar Maarleveld and Hannie Kortland from the Amsterdam University of the Arts developed and tested a more visual rubric for secondary art education (Maarleveld & Kortland, 2013). In this instrument, pictures represent the criteria and students can colour these indicating their level. Results from a first field test revealed that art teachers and students in The Netherlands were enthusiastic about this visual rubric.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

Therefore a European variation of this rubric was developed and tested in the current Comenius project.

2. Instrument construction

2.1. Expert consultation

The first step concerned the development of the assessment instrument by the two researchers. The instrument was based on the original version of the instrument by Maarleveld & Kortland (2013) and in the first round the researchers adapted the instrument to fit the CEFR_VL competence model. Criteria for both production and reception were formulated. They consisted of both a visual rubric (drawings made by Oskar Maarleveld) and a text rubric.

Then seven experts from different countries were requested to respond to this first version of the instrument by providing extensive feedback on the structure, the criteria, the text rubric, teacher instruction and the visuals used in the instrument. The experts came up with suggestions for adding and removing criteria, for alternative formulations, and for drawings. Also the number of levels of the rubric was discussed.

Subsequently, the researchers developed a second version of the instrument which was based on the feedback received from the experts. Of course, some of the feedback was contradicting and the researchers needed to develop a compromise which was based on a majority of opinions. This second version of the instrument was returned to the same experts for another feedback round. Based on feedback in the second round, the final version of the instrument was developed.

2.2 Criteria

The instrument consists of 11 criteria for art production: Collecting sources of inspiration; Experimenting; Making personal work; Communicating visually; Planning; Collaborating; Persevering; Using media and techniques; Using visual elements; Applying functions, styles and genres; Presenting. The instrument consists of 8 criteria for art reception: Recognizing and describing media and techniques; Recognizing and describing visual elements; Recognizing and describing functions, styles and genres; Analysing and interpreting; Using perspectives; Having an opinion and supporting it; Having insight in your own preferences; Being curious and open.

For all criteria both visual rubrics and text rubrics were developed. Students and teachers only had to rate the visual rubric, containing all the criteria. The text rubric was meant as background information and explanation. Figure 1 shows an example for production and figure 2 shows an example for reception.

Figure 1: visual rubric and text rubric for criterion Experimenting

Experimentieren

This project
reflects the
made of



support from the European Commission. This publication [communication]
and the Commission cannot be held responsible for any use which may be
rein.

<i>Niveaus</i>			
1	2	3	4
Du hast sofort begonnen dein Endprodukt zu gestalten. Du hast keine neuen Dinge ausprobiert.	Du hast wenig Zeit damit verbracht zu experimentieren. Du hast wenige neue Medien, Techniken, Methoden oder Anwendungen ausprobiert.	Du hast einige Zeit damit verbracht zu experimentieren. Du hast einige neue Medien, Techniken, Methoden oder Anwendungen ausprobiert.	Du hast viel Zeit damit verbracht zu experimentieren. Du hast viele neue Medien, Techniken, Methoden oder Anwendungen ausprobiert. Deine Experimente haben zu Entdeckungen geführt die im Endprodukt angewendet wurden.

Figure 2: visual rubric and text rubric for criterion Analysing and interpreting



<i>Niveaus</i>			
1	2	3	4

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



<p>Du hast kaum Verbindungen zwischen Form, Inhalt, Funktion und Kontext der visuell gestalteten Produkte gezogen. Die Verbindungen, die du gezogen hast sind nicht ausgereift.</p>	<p>Du hast wenige Verbindungen zwischen Form, Inhalt, Funktion und Kontext der visuell gestalteten Produkte gezogen. Die Verbindungen, die du gezogen hast sind teilweise ausgereift.</p>	<p>Du hast einige Verbindungen zwischen Form, Inhalt, Funktion und Kontext der visuell gestalteten Produkte gezogen. Die Verbindungen, die du gezogen hast sind ausgereift und verständlich.</p>	<p>Du hast viele Verbindungen zwischen Form, Inhalt, Funktion und Kontext der visuell gestalteten Produkte gezogen. Die Verbindungen, die du gezogen hast sind sehr ausgereift und verständlich (umfassend erläutert und/oder illustriert unter Angabe von Quellen).</p>
---	---	--	--

2.3 Levels

The CEFR_VL model distinguishes three levels of competence, but it was decided that the rubric would distinguish four levels. The reason is that in contrast to the model (that indicates general levels of competence) the rubric is directed toward the assessment of specific assignments in school practice. The researchers and the majority of the experts were of the opinion that for the assessment of assignments three levels that are positively stated, are not sufficient. A fourth level should be added that indicates that a student has failed the minimum level of one or several of the assessment criteria. Another argument was that with three levels, the second, intermediate level would be an attractive 'way out', whereas with four levels one is forced to decide for the lower of the upper half of the levels. Not all experts agreed and some insisted that three positively formulated levels should be applied in the rubric as well.

2.4 Teacher instruction

In the instruction for teachers it is stated that it is essential that the teacher and the students study and discuss the rubrics before the start of the production and reception tasks. It is also possible to discuss in class the criteria for good work in production or reception based on examples. When is an art work good? What do concepts such as experiment exactly mean? It is crucial that students understand the textual rubrics and agree to a certain extent with the contents of the rubrics.

After the discussion the teacher chooses (if possible: together with the students) which aspects are relevant considering the upcoming assignment. The teacher clarifies beforehand which of the aspects/ criteria will be assessed, for not always all aspects are relevant. It should also be clear for the students if the assessment based on the instrument is used in a formative or summative way. The teacher explains how the visual reflection forms will be used: the student assesses his/her performance by either colouring the visuals, or by ticking or marking the level. The teacher or a peer may also assess the student. The teacher has several options here.

The rubrics may be used during several class hours. Through repeated use the student and the teacher will acquire overview of the progress in the process. By the end of the lesson series, the

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



students and the teacher complete again the forms. They may become part of a (digital) portfolio. Again, the teacher chooses what is suitable for the specific class and assignment.

3 Relationship between rubric and CEFR_VL model

Table 1 shows the 11 criteria that were chosen for the art production rubric and it shows their relationship with the competencies of the CEFR_VL model. At the time of the instrument construction, the structure model CEFR_VL had not yet found its final version, so there is no one on one relationship of competencies mentioned in the model and criteria used in the instrument. However there is sufficient overlap as is shown in the table.

Table 1 Art production criteria and relationship with CEFR_VL model competencies

Visual Rubric: Produktion	CEFR_VL model: Teildimensionen Imaginieren, Entwerfen; Gestalten; Transformieren; Realisieren
-	<i>Imaginieren:</i> Innere, oft bildliche Vorstellungen erzeugen, die nicht einer aktuellen Sinneswahrnehmung entsprechen
Inspirationsquellen sammeln	<i>Entwerfen:</i> Ideen finden; relevante Kenntnisse recherchieren <i>Verwenden:</i> Bilder die als Werkzeuge fungieren, etwa als vorbereitende Skizze
Experimentieren	<i>Experimentieren:</i> risiko –und forschungsbereit sein; ungewohne Möglichkeiten ausprobieren
Persönliche/individuelle Arbeiten gestalten	<i>Gestalten:</i> gestalten jenseits erlernten Regeln erfordert Selbstbewusstsein
Planen	<i>Methodenkompetenz</i> Planungs und Strukturierungsprozesse als Methodenkompetenz zu verstehen
Zusammenarbeit	Als <i>Sozialkompetenz</i> zu verstehen
Ausdauer/ Durchhaltevermögen	<i>Realisieren:</i> in gestalterische Prozessen sich anstrengen; Widerstände überwinden; ausdauern
Medien und Techniken verwenden	<i>Gestalten/Realisieren:</i> Techniken, Medien und Materialien absichtsvoll einsetzen
Gestaltungselemente verwenden	<i>Gestalten/Realisieren:</i> angemessene Gestaltungsmitteln einsetzen

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



	(Ausschnitt, Farbe, Komposition)
Funktionen, Stile und Genres verwenden	<i>Gestalten/Realisieren</i> : Regeln (Codes, Konventionen von Genres) Berücksichtigen; Bildideen in einem bestimmten Genre ausarbeiten
	Teildimensionen Verwenden-Kommunizieren-Präsentieren
-	<i>Verwenden</i> : Bilder zielgerichtet einsetzen
Visuell kommunizieren	<i>Kommunizieren</i> : mitteilen, sich verständigen
Präsentieren	<i>Präsentieren</i> : für andere wahrnehmbar machen

Table 2 shows the 8 criteria that were chosen for the art reception rubric and it shows their relationship with the competencies of the CEFR_VL .

Table 2 Art reception criteria and relationship with CEFR_VL model competencies

Visual Rubric: Rezeption	CEFR_VL model : Teildimensionen Wahrnehmen- Untersuchen- Deuten- Beurteilen- Wertschätzen
-	<i>Wahrnehmen</i> : Gegenstände, Formen Prozesse in adäquaten Modi beobachten
Medien und Techniken erkennen und beschreiben	<i>Beschreiben</i> : die sprachliche Formulierung der bewusst gemachten sinnlichen Wahrnehmung <i>Untersuchen</i> : Analyseaspekte sind Material, Technik, Medium
Gestaltungselemente erkennen und beschreiben	<i>Beschreiben</i> : die sprachliche Formulierung der bewusst gemachten sinnlichen Wahrnehmung (z.B. Form-, Raum- und Farbkomposition) <i>Untersuchen</i> : Analyseaspekte sind bildnerische Mittel (Formen, Farben)
Funktionen, Stile und Genres erkennen und beschreiben	<i>Untersuchen</i> : Genres und Typen erkennen; Mitteilungskategorien identifizieren
Analysieren und interpretieren	<i>Untersuchen</i> : Analyseaspekte sind die Beziehung von Form-Inhalt, Gestaltung-Thema; <i>Deuten</i> : Merkmale von Bildern werden miteinander und mit relevanten Kontexten in Verbindung gesetzt
Perspektiven anwenden	<i>Deuten</i> : multiperspektivisch interpretieren <i>Wertschätzen</i> : andere Sichtweisen und deren Ausdruck in Bildern als wertvoll bzw. Wichtig einstufen
Eine Meinung haben und diese	<i>Beurteilen</i> : eine begründete

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



begründen	Einschätzung formulieren <i>Wertschätzen</i> : Bilder einen Wert zumessen und dies kommunizieren
Sich über eigene Vorlieben bewusst sein	<i>Deuten</i> : Bereitschaft zur kritischen Reflexion der eigenen Reaktion
	Teildimensionen Ästhetisch Erfahren-Einfühlen
-	<i>Ästhetisch erfahren</i> : spezifische Form des sinnlich-emotionalen Wahrnehmens
Offen und neugierig sein	<i>Einfühlen</i> : Gefühle anderer nachvollziehen; Gedanken und Assoziationen anderer tolerieren

Not all of the dimensions and competencies of the CEFR_VL model are represented in the two rubrics. 'Imagining', 'perceiving' and 'aesthetic experience' are missing, for the same reason they are missing in the Niveaumodel. They are not observable (and thus not assessable) as such, but can be assessed through other competencies such as 'experimenting' and 'describing'. 'Aesthetic experience' is a very complex competence, and is assessed through related competencies such as to value, empathy, to interpret and to judge. 'Using images' is too broad a category, whereas 'communicating visually' is more specific.

In the model it is explained that often a competence is relevant for both reception and production. To make the assessment manageable, the competencies in the rubrics are either related to reception or to production. We realize this is a simplification. The CEFR_VL model also refers to meta cognition as an overarching competence. Metacognition is thinking about (reflecting) the processes of production and reception both during and after the process. Self-assessment with the rubrics is in itself is a form of metacognition. Therefore metacognition is not one of the criteria, as reflecting about your reflection would make things too complicated.

4. Conclusion

Regular and structured forms of formative self-assessment are important in art education. Therefore an assessment instrument was constructed that can be used by students, but that is also suitable for teacher assessment and peer assessment. The form of the instrument was inspired by literature and previous research on assessment in art education in schools. The content of the instrument was based on the competence model of CEFR_VLI. Experts advised on the resulting assessment instruments: a visual and a text rubric for production and for reception of art. At face value the instrument seems to be a sufficient operationalisation of the CEFR_VL model, although some simplifications and reductions are inevitable. In order to test the instrument for feasibility and validity, empirical research was conducted in schools in Austria, Germany, Holland and Hungary. This evaluation is reported in the next chapter.

This project has been funded with support from the European Commission. This publication [communication] reflects the views only of the author, and the Commission cannot be held responsible for any use which may be made of the information contained therein.



5. References

- Black, P., & William, D. (2009). Developing a theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21, 5-31. doi: 10.1007/s11092-008-9068-5
- Boud, D. & Soler, R. (2015). Sustainable Assessment revisited. *Assessment and Evaluation in Higher Education*, doi: 10.1080/02602938.2015.1018133
- Brookhart, S.M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education; Principles, Policy and Practice*, 8(2), 153-169. doi: 10.1080/09695940123775
- Haanstra, F., Damen, M., Groenendijk, T., & van Boxtel, C. (2015). A review of assessment instruments in arts education. In S. Schönmann, *The wisdom of the many. Key issues in arts education. International Yearbook for Research in Arts Education* (p. 413- 418)New York : Waxmann.
- Nichol, D.J., & MacFarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, 31(2), 199-218. doi: 10.1080/03075070600572090
- Lucas, B., Claxton, G., & Spencer, E. (2012). Progression in creativity: Developing new forms of assessment. Background paper for the OECD conference "educating for innovative societies".
- Maarleveld, O., & Kortland, H. (2013). Beeldend reflecteren: een nieuwe beoordelingsmethode voor praktisch beeldend werk in het voortgezet onderwijs.
http://www.ahk.nl/fileadmin/download/ahk/Lectoraten/Kunst-en_cultuureducatie/beeldend_reflecteren_-_o_maarleveld_en_h_kortland.pdf
- Sadler, R.D. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Sluismans, D., Joosten-ten Brinke, D., & Van der Vleuten, C. (2013). *Toetsen met leerwaarde. Een reviewstudie naar de effectieve kenmerken van formatief toetsen*. NWO-PROO rapport.