Selfevaluation tool for learning by children in a maker space

van Eijck, Tom; Pijls, Monique; Kragten, Marco

Publication date
2019
Document Version
Submitted manuscript

Self-evaluation Tool for Learning by Children in a Maker Space

Tom van Eijck, Monique Pijls and Marco Kragten
Amsterdam University of Applied Sciences, Amsterdam, Netherlands.
t.j.w.van.eijck@hva.nl

Abstract
Maker spaces are generally regarded as a valuable innovation in comparison to traditional education, although it is largely unclear what is exactly learned. This deficiency hampers the deployment of maker spaces, particularly their embedding and integration in the existing practice in formal education. In the work presented here, we explore the possibility of having learners self-report on their learning experience. For this purpose, we developed an easy-to-use visual tool for assessing learning of 21st Century Skills in children's maker space activities, the Self-Evaluation Tool (SET). Particularly, we investigated the validation of the SET for the self-evaluation of learning activities in the maker space and how children evaluate their own performance in the various domains. The results show higher scores on learning goals in subjectification and lower scores for socialization. Future research will focus on a comparison of the different types of maker programs.

Introduction
In order to assess the development of children's 21st Century Skills as a result of maker activities (Pijls, Kragten & Van Eijck, 2018) it is important to know how children evaluate their own performance in maker spaces. Hence, a Self-Evaluation Tool (SET) was developed, in this case a graphic version of a learner report. De Groot (1974) introduced the learner report as 'an instrument to evaluate learning goals that are hard to evaluate in another way'. Over time it has been used as an instrument for evaluation of learning (Van Kesteren, 1993) and a way to stimulate the learner to express what he or she has learned. In addition, several studies show that formative (self-)assessment has a positive effect on motivation and metacognition (Castelijn & Baas, 2016; Sluysmans & Kneyber, 2016), and that self-monitoring and self-evaluation are essential for self-regulation of learning (Zimmerman, 2013).

Maker spaces are generally regarded as a valuable innovation in comparison to traditional education. The learning goals of the maker space are based on Biesta's goal domains (2014): qualification, socialization and subjectification. Comparable learning goals which were taken in account were: creation, iteration, sharing and autonomy (Cohen, 2017) and engagement, initiative/intentionality, social scaffolding and development of understanding (Bevan, Gutwill, Petrich & Wilkinson, 2014). In addition, Chu et al. (2015) state that preferred learning outcomes of making activities should be an increase in perceived ability, motivation and in taking initiative. Katterfeldt, Ditter and Schelhowe (2015) summarized three core ideas for learning environments aiming at a change of self: *begreifbarkeit* (i.e., making connections between the virtual and physical world), imagineering (i.e., creating objects with personal meaning) and self-efficacy (see also Sheridan et al., 2014).

Although it is generally accepted that learning takes place, it is still largely unclear what learners exactly learn when they are active in a maker space. This deficiency hampers the deployment of maker spaces, particularly their embedding and integration in the existing practice in formal education. In any learning-by-making environment, be it formal (school) or informal (maker space), learners and teachers need formative assessment for feedback on their

progress. For this, it is necessary to evaluate and reflect on both the making process itself and the result of this process (i.e., the product). In this study, our aim is to collect self-evaluation data at the end of a maker cycle and to compare and triangulate this data with observation and interview data. Using this approach, we aim to address two research questions:

- Validity: Is the self-evaluation instrument of learning activities in the maker space valid?
- Self-evaluation: How do learners evaluate their own learning in the various domains?

Methods

Design requirements

As part of a reorientation on its functions for the general public, the Public Library of Amsterdam is arranging maker spaces throughout the urban area, especially in neighbourhoods of low socio-economic strata. In the maker space, children and adolescents (ages 8-16 years) participate in technology-enhanced maker programs such as digital manufacturing (for example 3D-printing, laser and vinyl cutting) and programming (scratch, micro:bit). During the activities, children receive instruction and guidance by recently trained librarians: the maker space coaches.

The programs studied for the research reported here were either attended after school (small groups, averaging in 7 children of age 8-11 years, weekly 3-hour sessions) or during school hours (school classes, 22 children of 8-10 years, single 3-hour sessions). In both these contexts, the children should be able to use the SET at the end of a making session, preferably without help of the coaches.

The learning goals of these maker spaces are based on Biesta's goal domains (2014) and operationalized into eight categories (Table 1).

Table 1. Operationalization of learning goals.

Domain	Category	Item
Qualification	Creativity	A. 'I invented something new'
	Maker skills	B. 'I have learned something'
Socialization	Social scaffolding	C. 'I have helped another child'
	Social scaffolding	D. 'Another child has helped me'
	Sharing	E. 'I show what I make'
Subjectification	Intrinsic motivation	F. 'I like it'
	Persistence	G. 'I did persist'
	Self-efficacy	H. 'I dare to'

The graphic design of the learner report is based on a tool for self-evaluation from an e-learning application for primary education that stimulates children to design and investigate (Fraij & Zegers, 2018). The following aspects of this tool were incorporated in the design requirements:

1. It is a colouring sheet that can be filled out autonomously, without much explanation;
2. The symbols used are 'hollow' and therefore seduce to colouring;
3. The symbolic language is playful but unambiguous, and coherent with accompanying texts;

4. The symbols form distinctive series that grow larger towards the periphery.

An additional requirement was that every sequence consisted of a quantifiable gradation, from 1 (small, in the center of the sheet) to 5 (large, in the periphery).

The backside of the form was used to collect additional data, such as age, date, and the completion of three statements: 'I worked at…', 'I am proud of…', and 'Next time I will…'.

Design process

The design process consisted of 5 iterations (as shown in Table 2). After the first prototype was tested with a small group of children and evaluated by representatives of the partner organizations, a second version was validated by comparison with observation and interview data from seven children (Van Eijck, 2018). This led to the third prototype (shown in Figure 1), which was tested with a large group (N = 90).

Table 2. Overview of the design process

#	Prototype	Test group	Data	Data processing
1	Initial test + evaluation 24-11-2017	Children 7-10 yrs. (n=5) Representatives of partner organizations (n=11)	Filled-out forms 1st version Observation data (Handwritten) remarks	Recording of variance and (handwritten) remarks
2	Validation 30-1-2018	Children 9 yrs. (n=7)	Filled-out forms 2nd version Observation data Interview data	Recording of variance Comparison with observation data and interview data
3	Validation and data collection October 2018 – January 2019	Children 8-11 yrs. (n=90)	Filled-out forms 3rd version Observation data Interview data School test data	Adding to database Comparison of data Statistical analysis
4	Validation, data collection 28-1-19, evaluation 6-3-2019	Children 9-10 yrs. (n=23) Learning community (n=12)	Filled-out forms 4th version Interview data School test data	Adding to database Comparison of data
5	Validation and data collection March 2019-June 2019	Children 8-11 yrs.	Filled-out forms 5th version Observation data Interview data School test data	Adding to database Comparison of data Statistical analysis

Findings of the third test resulted in a fourth prototype, which was tested with a smaller group and re-evaluated by the Library Maker space Learning Community. A few minor adaptations resulted in the final version, which was used for data collection in March-June 2019.

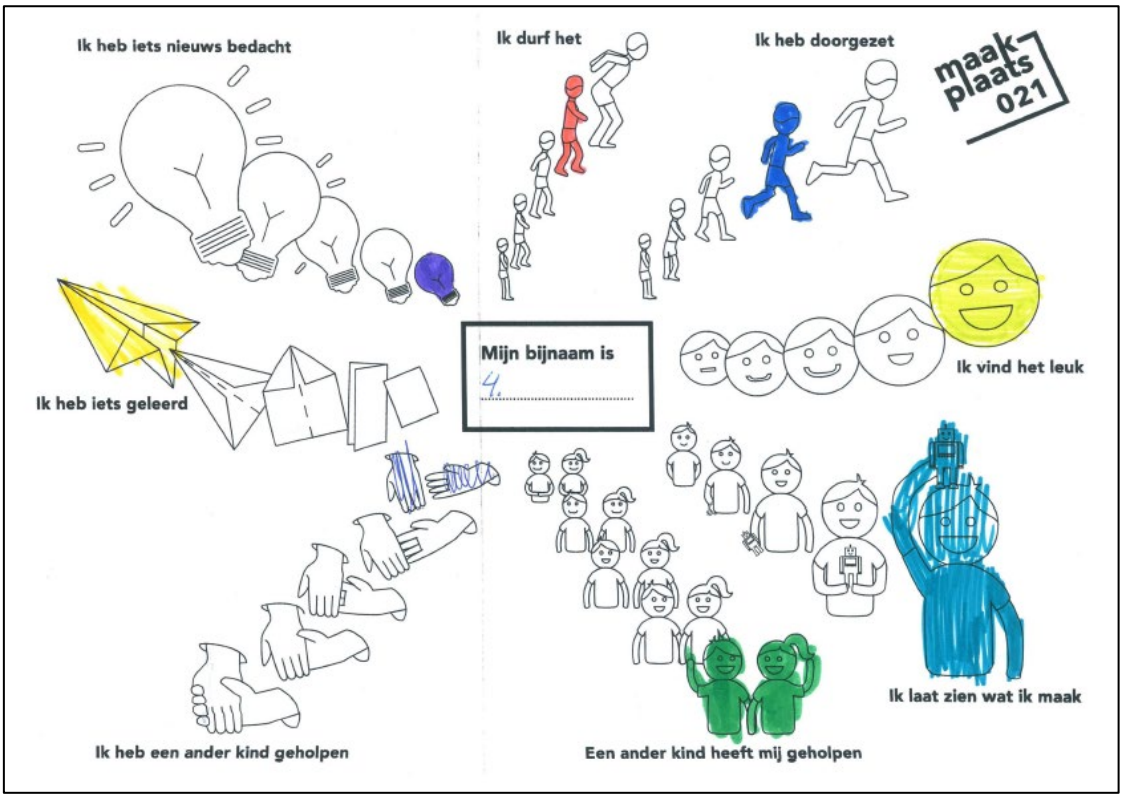

Figure 1. Self-evaluation Tool (SET), 3rd prototype.

During October 2018 - January 2019, the SET 3rd prototype was filled out by 90 children (age 8-11), during last session of their after-school programs and single sessions of school programs (Table 3). Completion of the form took less than 10 minutes per child.

In addition to the SET-scores, data from the maker programs followed during the school hours could also be compared with standardized test data from 43 learners, provided by the school. This data included test results for maths and reading comprehension, as well as psycho-social data from a Strengths and Difficulties Questionnaire (SDQ).

Table 3. Overview of data collection, SET 3rd prototype (N=90).

Maker space	Program	Date	Size
South	Science Club	29-10-2018	9
South	Fabschool	7-12-2018	3
South	School Program	10-12-2018	24
West	CodeTeam Junior	11-12-2018	7
South	CodeTeam Junior	12-12-2018	9
North	Fabschool	14-12-2018	7
South	Fabschool	14-12-2018	7
North	Fabschool	18-12-2018	5
South	School Program	14-01-2019	19

Completion of the 3rd prototype revealed some minor linguistic misunderstandings of the accompanying texts, leading to a 4th prototype, in which also the graduation of the symbols

was diminished. Testing of this prototype, however, led back to the spatial design of the 3rd prototype. Combined with the aforementioned linguistic adjustments, this resulted in the final design (shown in Figure 2).

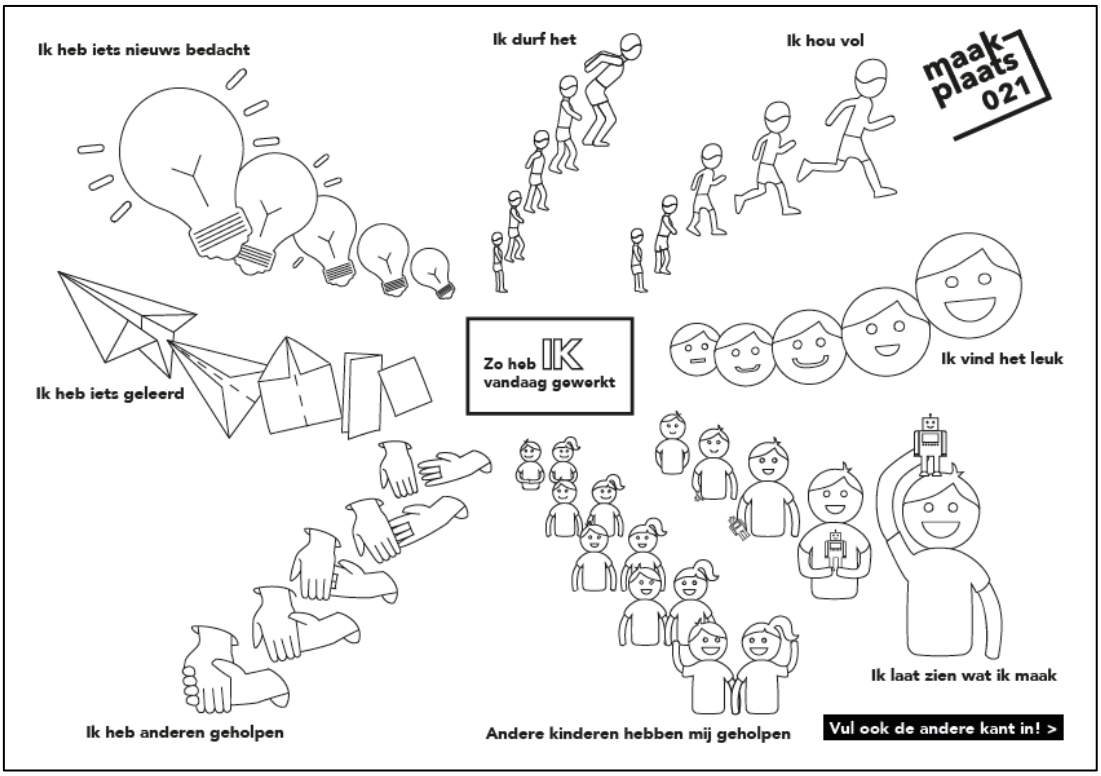

Figure 2. SET, final design. Compared to the 3rd version, the text box in the center: 'My nickname is…' is replaced by 'This is how **I** worked today'; the text top right ('I did persist') is simplified into: 'I endure'; the text bottom left ('I helped another child') is changed into: 'I helped others' and the text center-bottom ('Another child helped me') is changed into: 'Other children helped me'.

Validation data analysis and interpretation

In the various testing rounds, data for validation of the SET was generated by (i) additional behavioural observations during the making process, (ii) interviews after the completion of the SET, and (iii) comparison with standard test data provided by the school.

In the domain of qualification (creativity and maker skills), high (4-5) scores were given to 'I invented something new' even when only one or two 'inventions' were mentioned. Similarly, high scores on 'I learned something' were even given when only one or two examples could be recalled in the interviews. Also, high scores on 'I learned something' negatively correlated with psychosocial problems (Table 4).

For the domain of socialization (social scaffolding and sharing), low scores were given for 'I helped another child' when only one or no other child was helped, higher scores when two or more children were helped. Significant negative correlation with social problems (Table 4): children with low SDQ PL-scores (less social problems) gave high SET-scores. 'Another child helped me': low scores when being helped by only one other child or not being helped, higher scores when being helped by more than one child and a significant negative correlation with math scores: children which scored higher in being helped by others show lower math scores.

‘I show what I make’: mostly high scores, even when product was said to be shown to only one other person (for example the teacher or a parent). Significant positive correlation with ‘I like it’, ‘I show what I make’ and ‘I dare to’ (see also Table 5).

In the domain of subjectification (intrinsic motivation, persistence and self-efficacy), ‘I like it’ scored the highest of all categories, and had significant positive correlations with all other categories, except with ‘another child helped me’ (Table 5). ‘I did persist’ also gave high scores. Interview data suggests that the children associated this with the number of times they paused during their task because of distraction, like in the classroom. ‘I dare to’: in the interviews, children explained their (mostly maximum) scores with remarks such as ‘there was nothing really scary’. Furthermore, all categories associated with subjectification had the lowest SD.

Correlations between SET-scores from the programs followed during the school hours and standardized learner-test data provided by the school are shown in Table 4.

Table 4. Correlations between SET-scores and school test results. A: Invented something new; B: Learned something; C: Helped another child; D: Another child helped me; E: Show what I make; F: Like it; G: Did persist; H: Dare to. SDQ: Strengths and Difficulties Questionnaire; TOT: total; PL: social problems; EM: emotional problems; PR: pro-social behaviour).

	Age	**Language**	**Maths**	**SDQ TOT**	**SDQ PL**	**SDQ EM**	**SDQ PR**
A	-.12	-.08	.05	.12	-.07	-.13	-.14
B	.07	-.13	.03	-.31*	-.22	-.26	.15
C	-.03	-.01	-.21	-.30	-,33*	-.07	.05
D	-.15	-.28	-,37*	.06	-.19	-.15	-.02
E	.00	.14	.02	-.02	-.08	-.15	.20
F	.26	-.04	.03	.02	.05	.11	.04
G	.19	.12	.02	.05	-.15	.06	.10
H	.20	-.02	.06	-.10	-.30	-.09	.21

*. Correlation is significant at the 0.05 level (2-tailed).

Self-evaluation data analysis and interpretation

Figure 3 shows the distribution of the average scores and SD per item. The lowest average scores were on the items ‘I invented something new’ (A), ‘I helped another child’ (C) and ‘Another child helped me’ (D). Highest averages were on the items ‘I like it’ (F), ‘I did persist’ (G) and ‘I dare to’ (H). The differences in SD are also notable: greater differences between scores on ‘I invented something new’, ‘I helped another child’ and ‘Another child helped me’ (high SD), and smaller differences in the scores on ‘I like it’, ‘I did persist’ and ‘I dare to’ (low SD).

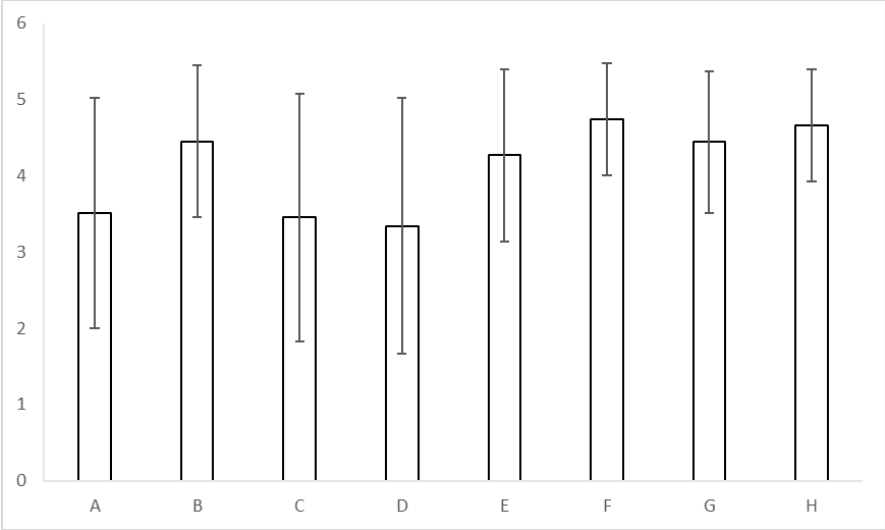

Figure 3. Distribution of scores and SD per item (N=90). A: Invented something new; B: Learned something; C: Helped another child; D: Another child helped me; E: Show what I make; F: Like it; G: Did persist; H: Dare to.

Table 5. Correlation matrix of SET-scores. A: Invented something new; B: Learned something; C: Helped another child; D: Another child helped me; E: Show what I make; F: Like it; G: Did persist; H: Dare to.

	A	B	C	D	E	F	G
B	.31**						
C	.29**	.17					
D	.30**	.11	.33**				
E	.27*	.21*	.06	.16			
F	.24*	.36**	.24*	.13	.36**		
G	.26*	.13	.26*	.16	.24*	.27*	
H	.16	-.02	.21*	.11	.39**	.26*	.37**

*. Correlation is significant at the 0.05 level (2-tailed).
**. Correlation is significant at the 0.01 level (2-tailed).

Qualitative assessment of the text written on the backside of the SET from the Fabschool-program showed that most children are proud of the item they made during the workshops (Figure 4).

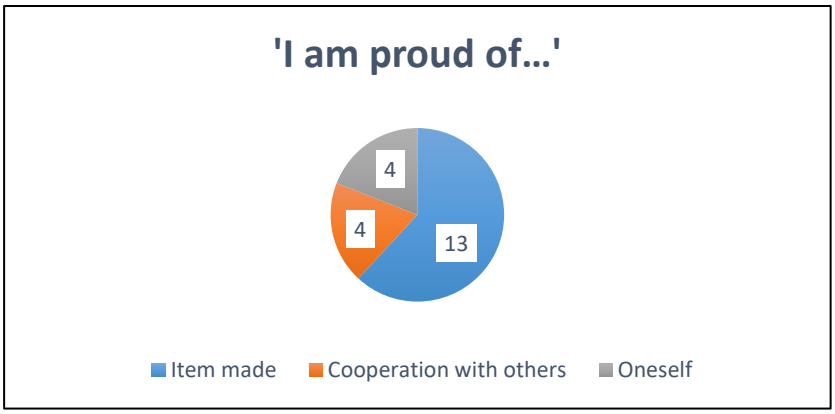

Figure 4. Overview of the items that children were proud of (Fabschool, n=21).

Conclusions

Is the self-evaluation instrument of learning activities in the maker space valid?

Summarising, it can be concluded that the self-evaluation tool is valid, but that in most categories there might be a 'ceiling effect'. An average high score (4 or 5) simply means 'yes', especially in the domains of qualification and subjectification. However, this might imply that lower scores (i.e. about 3) mean 'a bit' or 'little'. Moreover, because there is a significant positive correlation with the scores for 'I like it' (Table 5), the rich and informal learning environment of the Makerspace most likely arouses the overall motivation. Hence, this might be more an indicator for a learning *attitude* rather than for a learning *experience*.

How do the children evaluate their own learning in the various domains?

In the domain of qualification, most children find that they 'learn something', but are more divided about 'inventing something'. In de domain of socialization, helping and being helped have the lowest averages of all scores, and with the highest variation. In other words, some children evaluate themselves either helping more or being helped more than others. Scores are highest within the domain of subjectification; apparently, almost all children are having a good time in the Make space and evaluate themselves as persistent and daring.

Discussion

Although the tool might be valid for the self-evaluation of the various learning goals, there might be differences in interpretation in the text and pictures of each category; differences among the children as well as differences between children and adults (i.e. researchers). In the domain of qualification, for example, children of age 8-10 might interpret 'inventing something new' also as 'solving a problem', of which the result is not a new 'invention', but just the overcoming of a practical problem they encounter during the making process. Also, although children can indicate 'how much' they learn, it remains unclear *what* they think they learn. In order to collect qualitative data on this point, this question should be included on the backside of a new version of the SET. Also, there might be a strong overall positive attitude bias, which might have an elevating effect on the scores altogether.

In the domain of socialization, average scores for 'I helped another child' are slightly higher than for 'Another child helped me'. Logically speaking, these scores should be exactly the same. This might indicate that some children evaluate themselves more 'helpful' than others, although in reality the extent in which they are helping others might be the same as other children. The difference could also be caused by differences in awareness of 'helping others' between individual children or by the tendency to give socially desirable answers to questions from adults (i.e. coaches, teachers or researchers).

An overall comparison of the data suggests that, although the children show a positive attitude and self-esteem (subjectivation), they evaluate themselves less positive on creativity and social skills. This could be explained by the fact that in most of the programs there is limited occasion for child-initiated activities. When comparing future SET-data with the extent in which the maker programs allow the children to follow their own initiative, we expect a positive correlation with SET-scores in creativity and socialization.

With the SET, we designed an instrument that is valuable for assessing the impact of maker education programs on the development of 21st Century Skills, especially on learning and innovation. Since formative (self-)assessment has a positive effect on motivation and metacognition and is essential for self-regulation of learning, we will also explore how the SET can serve as a maker education tool for enabling children to become more conscious of their own development in learning-by-making, thereby increasing the effectivity of the maker education programs.

Acknowledgements

This research was funded by the municipality of Amsterdam as part of the project Maakplaats 021. Partners in this project are the Amsterdam Central Library (OBA), Waag Society, Pakhuis de Zwijger and the Amsterdam University of Applied Sciences. We kindly thank the personnel of OBA 021 Maker spaces Reigersbos, Slotermeer en Waterlandplein, and the Openbare Daltonschool Nellestein and Stichting Sirius for their cooperation.

Literature

Biesta, G. (2014). *The beautiful risk of education*. Routledge.

Bevan, B., Gutwill, J. P., Petrich, M., & Wilkinson, K. (2015). Learning through stem-rich tinkering: Findings from a jointly negotiated research project taken up in practice. *Science Education, 99(*1), 98-120.

Castelijns, J. & Baas, D. (2016). *Evalueren om te Leren. De integratie van formatief toetsen in de dagelijkse praktijk van basisscholen. Dossiernummer 405-15-509*, Lectoraat Eigentijds Beoordelen in het Onderwijs, Hogeschool de Kempel, Helmond.

Chu, S. L., Quek, F., Bhangaonkar, S., Ging, A. B., & Sridharamurthy, K. (2015). Making the Maker: A Means-to-an-Ends approach to nurturing the Maker mindset in elementary-aged children. *International Journal of Child-Computer Interaction, 5,* 11-19.

Cohen, J., Jones, W. M., Smith, S., & Calandra, B. (2017). Makification: Towards a framework for leveraging the maker movement in formal education. *Journal of Educational Multimedia and Hypermedia, 26*(3), 217-229.

Eijck, T. van, Fraij, M., Kragten, M. en Pijls, M. (2018). *Self-evaluation of learning activities by children in a library-makerspace*. Fablearn 2018, Eindhoven, Netherlands.

Fraij, M. & Zegers, J. (2018). De Ontdekapp. www.ontdekapp.nl. Accessed 05 May, 2019.

Kesteren, B.J. van (1993). *Applications of De Groot's 'Learner Report': A tool to identify educational objectives and learning experiences.* Department of Agricultural Education, Wageningen Agricultural University, Wageningen, The Netherlands.

Pijls, M., Kragten, M., van Eijck, T. (2018). *Maakplaats 021. Praktijkonderzoek naar het verwerven van vaardigheden voor de toekomst door kinderen in Amsterdamse buurten in maakplaatsen bij de openbare bibliotheek*. Kenniscentrum Onderwijs en opvoeding, Lectoraat Didactiek van de Bètavakken, Hogeschool van Amsterdam, Netherlands.

Sheridan, K., Halverson, E., Litts, B. K., Brahms, L., Jacobs-Priebe, L., & Owens, T. (2014). Learning in the making: Comparative case study of three makerspaces. *Harvard Educational Review, 84*(4), 505–532.

Sluijsmans, D. & Kneyber, R., Eds. (2016). *Toetsrevolutie. Naar een feedbackcultuur in het voortgezet onderwijs*. Uitgeverij Phronese, Netherlands.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*, 135-147.