

Formative Assessment; Pivot of Masterly Development

Miriam Goes, Open University, The Netherlands
 Marieke Dresen, Open University, The Netherlands
 Marcel van der Klink, Open University, The Netherlands

Summary

In this paper we present our research findings concerning the development of competency-based assessments that support the learning process of novice teachers and career switchers. More in detail this paper addresses the first step in the designing process: collecting and describing professional situations in such a way that it serves as valuable input for the design and delivery of assessments.

One of the main activities of the Ruud de Moor Centre (RdMC) of the Open University of the Netherlands (OUNL) consists of design and delivery of on-line assessments based on situations that are typical for the teacher profession. These on-line assessments will assist teachers to gain better insight in their own professional growth and will support teachers in selecting and performing learning activities. In other words, assessment is perceived as formative assessment.

As mentioned, these assessments are based on a list of professional situations set up by SBL, the association for the professional quality of teachers as a representative of the Dutch Teachers' Unions. For example, the situation: "parental evening" and "cooperation with colleagues".

However, this list of situations does not provide enough input for designing assessments. More detailed information is needed. That is why the RdMC developed a method for collecting, describing and presenting in-depth information.

The method consists mainly of two steps. The first step; participants are requested to give information about the situation using a format to ensure that four areas are covered; 1) output of the situation, 2) behaviour in the situation 3) factors that affect the complexity of the situation 4) descriptions of the situations. This is done individually by the participants.

The next step is; participants attend a session and discuss this information in groups and the group facilitator assures that collective agreement is reached upon output, behaviour and complexity factors of the situation.

At the moment six groups of participants, consisting of

teachers employed in primary, secondary and vocational education, have worked with this method.

This research project focuses on the evaluation of this method. The project addresses the following questions 1) does the method provide useful descriptions for the design of formative on-line assessments, 2) how do participants and facilitators experience and value this method, 3) what are the recommendations to improve the method.

Written questionnaires for participants and facilitators are used to collect data. In addition an interview with each group is scheduled.

After presenting the research findings the paper discusses some key issues in the design and delivery of the method.

Introduction

Teacher education in the Netherlands is competency based. The key words are: diversity, constructivist education, competence based learning and life long learning and qualification requirements. (Klarus 2004). The Occupations in Education Act (Wet op de beroepen in het onderwijs, Wet BiO) ensures the quality of the educational staff. This Act describes the competencies required for certification. In other words it is the national standard of competencies for teachers (Ministerie OCW 2004).

In the Netherlands we work with the concept of accreditation of prior learning. This so called 'zij-instromerstraject' (learning pathways for fast track teachers) enables persons, if subject to satisfactory performance, to get qualified within two years. So they need a testimonial of aptitude and a job in the teaching profession. This testimonial can be acquired after an assessment during which it is determined what competencies still need to be developed. The fast track teacher is employed by a school and is guided by a teacher educator and a coach in the school.

When the fast track teacher has sufficiently developed his competencies, he takes an assessment. During this assessment it is established if the candidate meets the requirements for certification

From interviews with fast track teachers we learn that the communication between school, fast track teacher and teacher training institute leaves much to be desired for. Moreover, the assessment is not transparent enough (Directory Ruud de Moorcentrum 2004).

The Ruud de Moor centre, a department of the Open University in the Netherlands, is aimed at professionalisation of the fast track teachers and helps to establish learning communities to improve the qualities of novice teachers and teachers (Stijnen, 2003). The assessment project of the Ruud de Moor Centre provides on-line formative assessments thus increasing the transparency to all parties involved. The formative assessments have been based on situations typical for the teaching profession. Therefore an analysis and descriptions of these situations were needed. However, teachers themselves have not always been involved in the discussions on professional situations and competencies so far (Sluijsmans 2002). This research project chose for an active participation of teachers in order to analyse and describe the professional situations. The Ruud de Moor centre developed a method for this.

In this report we present an evaluation study of this method for analysing and describing characteristic professional situations. In the second chapter we discuss the theoretical framework. Then we describe the methodology, set-up and the implementation of the study. The results are the theme of the fourth section. This report ends with conclusions and discussion.

Theoretical framework

Definition of concepts

We define formative assessments as follows: assessments that show and monitor the development of competencies and help to structure the learning path. In this project formative assessments on line are designed in order to make assessment independent of time and place and accessible for each individual. For that purpose the assessments provide feedback that helps the fast track teacher to monitor his own learning process and choose his own learning path.

In order to design these formative assessments it was necessary to describe the characteristic professional situations of teachers. Our definition of professional situations is; situations which regularly occur, require professional qualities and are characteristic for the profession. (Fontys 2005). Not every characteristic professional situation is a critical professional situation (Flanagan 1954). The latter is complex of nature and in case of unprofessional performance resulting in major negative consequences. In a situation like this the professional is faced with a choice, problem, dilemma, tension and is therefore complex by nature. Critical professional situations are crucial and characteristic. Characteristic professional situations are not always crucial or critical.

The method we use aims at collecting information that is necessary to design formative assessments, based on the characteristic professional situations. The method consists of a few steps. First we send the participants a task for homework. This task concerns the description of characteristic situations by means of:

- The desired results that have to be realised in this situation
- The factors that make the situation complex

- The behaviour necessary to realise the desired results
- Descriptions of three situations; one easy, one average and one complex situation
- The competencies required for the situation

This last element is included because fast track teachers are very often placed in highly complex situations. This means that it is not a safe learning environment. It is generally better to start in relatively simple situations which can become more complex as the competencies of a student develop. Factors complicating a situation are necessary for the design of formative assessments. Examples of these factors are: The number of students in the class (group size) or the number of problem students (group composition).

The homework task acts as input for the next step which is a session. During this session these tasks are discussed and a collective agreement is reached on results, complexity factors and behaviour and the competencies.

Theories

We carried out a literature research. We looked at various methods..

1. The 'critical incidents' method (Flanagan 1954). This is a way to categorise people's experiences. Experiences are difficult to explain in words (Dekker 2000) because they also entail implicit knowledge. The critical point here is that the method is not very structured.
2. Observation methods. A method solely based on observation can cause too much bias of the observer and this method is often experienced as subjective. The observer has to be an expert in the educational sector concerned.
3. Wisdom of practice- study: This is a method to determine competencies. Analysis of competencies can only take place within the framework of a development centre (Tillema 2004). The author distinguishes seven steps: demarcation, preparation, competencies analysis, data analysis, making a tree diagram, linking the results and selection of the exercises. Critical here is that too much emphasis is on competencies and not on situations in which these competences have to be demonstrated.
4. The use of structured interviews is not valid and difficult to combine.

The conclusion is that none of these theories is fully equipped for the method of the expert sessions although some elements are excellent for use. That is why we combined elements of aforesaid methods. Moreover, we looked at methods from industrial psychology such as functional analysis (Klarus 1998, 2003) and the development of professional profiles. These two methods include the collecting of information about professional situations. The functional analysis is especially concentrated on the inventory of the activities and tasks. The information that is proceeded by the functional analysis focuses on details and because of this the complexity of the total situation is not discussed.

Ideally a professional profile includes:

- a description of the most important and frequent activities in a profession
- an overview of the tasks and responsibilities within a profession
- the competencies which the professional needs to have
- Indications of future developments and the consequences of this for the tasks and responsibilities in a professional field (Klink, van der and Boon 2001). Critical point is that this method gives too much overall information

We also looked at the 4C/ID model, a research-based comprehensive instructional design model for complex learning (Janssen-Noordman and Van Merriënboer 2002). The 4C/ID model specifically states that component skills and associated knowledge have to be coordinated and integrated. The model helps to design a learning environment. The model distinguishes four interrelated blueprint components (4C):

- Learning Tasks
 - Supportive Information
- Just-in-Time Information
- Part-Task Practice.

A basic assumption of the 4C/ID model is that environments for complex learning can be described in terms of these four components. The most important components are concrete, authentic and meaningful learning tasks, based on situations from professional practice. Every learning task includes the whole professional task and is carried out in a realistic professional environment. The learning tasks provide the backbone of the programme. The other components are developed in relation with the learning task. Learning tasks are organised in so called task classes. Each new task class is more complex than the previous one, this partly depends on the degree of guidance given to the learner. Sufficient variation between the different learning tasks within a task class is required. A learning task in the highest task class performed independently can be considered as a test.

Characteristic for this model is that the user gets decreasing guidance and the tasks become more and more complex. Every task is analysed carefully. The fast track teacher finds himself in situations that vary from simple to complex. This explains why we have chosen to work with descriptions of the situations at three levels: easy, average and complex.

Methodology, set-up and implementation study

This research is performed to establish the workability of the method for designing learning paths, more in particular formative assessments. It aims to collect information on the method used and to make recommendations for improvement. We formulated the following three questions:

1. Does the method provide useful descriptions for the design of formative online assessments?
2. How do participants and facilitators (discussion leaders) experience and value this method?
3. What are the recommendations to improve the method?

The method is implemented with groups consisting of ten teachers all from similar educational sectors. The educational sectors are: primary education (PO), junior vocational schools (VMBO), secondary education (VO) and senior vocational schools and adult education (BVE). The group consisted of experienced and beginning teachers.

At present six groups have been active. Two PO groups one VMBO group two VO groups and one BE group. Generally the groups work out five situations thus performing five homework tasks followed by five sessions.

The questionnaire consists of 67 open and closed questions on seven clusters: knowledge of the method and terms, results of the method, qualities of the expert and the discussion leader, the value of the method, the method itself, personal motives and general data. The closed questions are measured on a 4

point scale (I agree strongly, I agree, I disagree and I disagree strongly).

Examples of questions are:

The general questions concern data on age, gender, educational sector, number of years working in education etc.

A. Knowledge of the method and terms	1	2	3	4
	I agree strongly	I agree	I disagree	I disagree strongly
I am informed about the Occupations in Education Act (BIO)	0	0	0	0
The term characteristic professional situation is known	0	0	0	0
The term competency is known	0	0	0	0
In my school we work with competence based learning	0	0	0	0
I am involved at competence based learning	0	0	0	0
The background information (circle and competencies (appendix?)) is interesting	0	0	0	0
The background information (circle and competencies) is clear	0	0	0	0

Figure 1: questions from the questionnaire

A group interview was held, consisting of open questions, in order to get more background information regarding the questionnaires. Open questions were asked as to the advantages and disadvantages of the expert sessions, suggestions for improvement, any missing elements, the time spent and if the participants felt the number of descriptions was substantial. Moreover, we inquired if they are going to use the method or parts of it in the near future. The remarks and answers to the questions of the group interview are used as background information and incorporated in the results and conclusion.

Examples of the questions of the group interview are:

In addition interviews were held with the designers of the method and more information for example reports of the sessions was gathered.

2. How do you value the homework and the sessions?

How much time did you spend ?

Did you find working with the templates significant and time-saving?

What would you change ? And what would remain the same?

Did you find it difficult? Did you understand everything?

Results

The research group

The research group (N=49, response 83 %) consisted of 17 men and 32 women, of which 25 younger than 40 years and 24 older than 40. 12 respondents work in primary education (PO), 13 in junior vocational school (VMBO), 11 in senior vocational school and adult education (be) and 11 teachers in secondary school (VO), 2 are unknown. The years of working experience was less than 10 years for 25 persons whereas 23 persons had 10 years or more, 1 person did not give the information. The research group is divided into six groups. For the classification of the categories see appendix 1.

The variables

In the following table the mean, the standard deviation and the reliability of the variables is given

Scales were constructed for 7 variables, which vary from 4 up to 14 items per scale. The reliability is acceptable to good. The instrument is good and the response is high. No item is eliminated. A lower score on the scale means a higher degree of agreement as to the statements belonging to the scale.

Table 1: Number of items, reliability, mean and standard deviation of the scales (n=49)

Name of the scale	Number of items	Reliability	Mean	Standard deviation
Knowledge of the method	7	.69	2.07	.41
Results for RdMC	5	.72	1.84	.34
Output for individual	6	.70	1.87	.39
Qualities of expert	14	.84	1.74	.34
Qualities of discussion leader	9	.80	1.51	.34
Method	9	.79	1.92	.37
Personal motives	4	.82	1.48	.44

The scale Knowledge of method and terms consists of 7 items. Examples of items are:

- The term characteristic professional situation is familiar
- The term competency is familiar

The mean score of 2.07 indicates that respondents have relatively sufficient knowledge of the method and terms.

The scale Output of the method (for the Ruud de Moor centre) consists of 5 items. This scale partly covers the first research question namely the usefulness of the method. Examples of items are:

- The set descriptions of professional situations is familiar
- The behaviour that is necessary to realise the desired results is well defined

The mean score of 1.84 indicates that respondents think that the output for the Ruud de Moor centre is relatively good.

The scale Output of the method (for the respondent himself) consists of 6 items. Examples of items are:

- I learn to reflect by formulating and clarifying what I do during my lessons and why
- I use this method to look at colleagues at work and to learn from these observations

The mean score of 1.87 indicates that respondents think that the output for themselves is relatively good.

The scale Qualities of the expert consists of 14 items. Examples of items are:

The expert has to:

- Be motivated to think about the method and how to improve it
- Have some years of working experience

The mean score of 1.74 indicates that respondents fairly often agree with the statements.

The scale Qualities of the discussion leader consists of 9 items. Examples of items are:

The discussion leader has to:

- Formulate well, explain clearly
- Dare to intervene during discussions

The mean score of 1.51 indicates that respondents fairly often agree with the statements.

The scale Method consists of 9 items. Examples of items are:

- This working procedure is a good method to categorise the characteristic situations

- I think the time necessary for the tasks is realistic
- The mean score of 1.92 indicates that respondents valued the method pretty highly.

The scale Personal motives consists of 4 items. Examples of items are:

- I think it was important to participate
- I enjoyed the atmosphere

The mean score of 1.48 indicates that the personal motives of respondents are relatively positive and that respondents did find it pleasant to participate.

Apart from descriptive analyses, comparisons between the six groups have been made.

Using variance analysis (F-test) we have examined whether factors such as educational sector, years of working experience, age or expert group to which he has taken part, influenced the scores on the scales.

The results of this analysis show that the educational sector did not influence the scores on the scales. The years of working experience influenced the score on the scale 'Knowledge of the method' ($F = 3.897$, $df 38$, $p.01$). Respondents with fewer than two or more than fifteen years of experience scored lower on this scale. This means that these two groups have relatively more knowledge of the method and terms. For the former group this is quite logical because these teachers just received their degree and the terms are still fresh in their mind.

The age of the respondent influenced the scores on the scale 'Qualities Expert' ($F = 3.656$, $df 46$, $p.02$). The score in the group 'between 20 and 30 years of age' is the highest and in the group 'above 50 years' the lowest. This means that the older group relatively agrees less with the statements and the younger group relatively agrees more. The 'older group' probably found their qualities good enough. They have more self confidence.

Furthermore it appeared that the expert group in which the respondent participated did not influence the scores on the scales. There are some differences but not significant.

The group interview

The second research question concerning the experience of the participants was one of the topics of interest during the group interview.

Nearly all participants mentioned positive points such as more awareness of own behaviour, the teamwork and the opportunity to compare knowledge and experience with other colleagues. They recognised the situations and because of this they gained better insight as to the complexity of their own profession.

As a negative point they mentioned: every session followed the same procedure which made it a bit dull. Other negative points were the fact that the sessions combined with the homework were time consuming and the limited information available.

Suggestions for improvement were given by the teachers from the BVE-sector. They found it very important that the discussion leader is familiar with the educational sector. Another expert group preferred to receive more information as to the handling of the homework. The discussion leader of this group did not distribute this information (which was done by other discussion leaders). They also suggested to vary the sessions a bit more.

The majority of the participants experienced the homework and the sessions as useful, it made them think about their work. The time spent on homework was between approximately half an hour up to four hours per session. We suspect that the average time needed is 2,5 hours. They thought the tasks set were overall not so difficult but they had to get used to

the presentation of the questions and tasks. After one or two sessions they understood the procedure.

Concerning the first research question as to the output of useful situations one did not all agree. The groups BVE, VMBO and PO were positive and content about the amount of descriptions. The VO group had more doubts about the amount of descriptions and stated there are many more situations in order to cover the whole field of teaching.

About using the method in their daily work several respondents indicated that they actually do that. A striking aspect is that it involved coaches and assessors.

The interviews with (co)discussion leaders and the reports

From the interviews with the (co)discussion leaders and the reports of the sessions it became clear that the (co)discussion leaders invested a lot of time on the preparation of the session, the processing of the homework and the outcome of the sessions. The average time required for the co-discussion leader was 20 hours per session. For the discussion leader this was 16 hours per session (this included the session itself that lasted 3 hours on average and the required travelling time) and 7 hours for an assistant.

The remaining information from these interviews and reports did not add anything to the information from the questionnaires and the group interviews.

Conclusions and recommendations

The first research question which is whether the method produced the desired output is positively answered by the research group. In the group interview experts indicated that enough situations were well described. The discussion leaders were satisfied about the outcomes.

The second research question regarding the experiences and conceptions of the participants was also positively answered. The output per person is considered high (mean 1,87) and they wanted to continue participating in the sessions which indicates enthusiasm. The participants mentioned as highest output the fact that they were compelled to think about their own performance (self reflection) and the exchange of knowledge, experience and complexity of their profession with colleagues.

Improvement of the method was the focus of attention of the third research question. The participants indicated that they would like to see more variety in the sessions. They also mentioned they needed more information as to the time and tasks required. The discussion leader must be familiar with the educational sector of the teachers in the group.

The method is appropriate for competence based learning and is an aid in the design of step by step high-quality formative assessments. A point of attention is that the method costs a lot of time and work in comparison with other methods (see for example den Boer et al. 2004). Therefore one topic for discussion is the correspondence between merit and reward. Another topic for discussion addresses the transfer of this method to other professions than the teaching profession.

Appendix 1: Respondents split up into different categories

The research group has been divided in the following categories:

Table 2: gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	male	17	34.7	34.7	34.7
	female	32	65.3	65.3	100.0
	Total	49	100.0	100.0	

Table 3: age

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	20 - 30 years	13	26.5	26.5	26.5
	30 - 40 years	12	24.5	24.5	51.0
	40 - 50 years	17	34.7	34.7	85.7
	more than 50	7	14.3	14.3	100.0
	Total	49	100.0	100.0	

Table 4: educational sector

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	PO	12	24.5	25.0	25.0
	VMBO	13	26.5	27.1	52.1
	MBO/BVE	11	22.4	22.9	75.0
	VO	11	22.5	23.0	97.9
	Unknown	1	2.0	2.1	100.0
	Total	48	9.8	100.0	
Missing	9	1	2.0		
	Total	19	100.		

Table 5: working experience

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	less than 2 years	3	6.1	6.3	6.3
	2 - 5 years	7	14.3	14.6	20.8
	5 - 10 years	15	30.6	31.3	52.1
	10 - 15 years	8	16.3	16.7	68.8
	more than 15 years	15	30.6	31.3	100.0
	Total	48	98.0	100.0	
	9	1	2.0		
	Total	49	100.0		

Table 6: six groups

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	RISA	8	16.3	16.3	16.3
	RISH	9	18.4	18.4	34.7
	MDR	9	18.4	18.4	53.1
	GKE	8	16.3	16.3	69.4
	GKB	6	12.2	12.2	81.6
	FJW	9	18.4	18.4	100.0
	Total	49	100.0	100.0	

References

Fontys (2005), *Biloba op weg naar assessments. Notitie voor Fontys-medewerkers*, Eindhoven, Fontys.

Boer, P. den, Klarus, R., Tijssen, G. en Wee, E. ter (2004), *Levensechte verhalen als basis voor het competentieprofiel, Ervaring met ervaringen verzamelen*, Wageningen, Stoas.

Dekker, H. (2000), 'Critical incidents methode' in: *Werken aan kennis*, nr. 6, 28-32, Utrecht, Kessels & Smit.

Flanagan, J.C. (1954), 'The critical incident technique' in: *Psychological Bulletin*, 41, 237-358.

Janssen-Noordman, A.M.B. and Merriënboer, J.J.G. van (2002), *Via leertaken naar complexe vaardigheden*, Groningen/Houten, Wolters-Noordhoff.

Klarus, R., Schüler, Y. en Wee, E. ter (2002), *Bewezen geschikt*, Wageningen, Stoas.

Ruud de Moorcentrum (2004), *Jaarboek Ruud de Moorcentrum*, Heerlen, Open Universiteit Nederland.

Merriënboer, J.J.G. van (2004), *Competentiegericht opleiden* (presentatie voor het Onderwijskundig Netwerk UT, 8 november 2004).

Sluijsmans, D.M.A. (2002), *Student involvement in assessment* (unpublished doctoral dissertation, Heerlen, Open University of The Netherlands).

Stijnen, P.J.J. (2003), *Leraar worden: 'under construction'?, Over lerarentekorten en afstandsonderwijs voor de opleiding tot leraar* (oration), Heerlen, Open University of The Netherlands.

Tillema, H.H. (2004), 'Gericht werken met competenties in de opleiding' in: *VELON, Tijdschrift voor lerarenopleiders*, 25(2), 28-34.

Wee, E. ter and Daal, M. (2004), *VO Geschikt*, Wageningen, Stoas.

www.ou.nl/rdmc

www.lerarenweb.nl