

Designing an observation instrument for Assessment for Learning practice

Drs. Niek van Benthum: Vilentum University of Applied Sciences / Wageningen University, n.van.benthum@stoasvilentum.nl, the Netherlands

Dr. Judith Gulikers: Wageningen University, Judith.Gulikers@wur.nl, the Netherlands

Dr. Frank de Jong: Vilentum University of Applied Sciences, f.de.jong@stoasvilentum.nl, the Netherlands

Prof. Dr. Martin Mulder: Wageningen University, Martin.Mulder@wur.nl, the Netherlands

Paper presented at the JURE conference, Munich, Germany, August 2013

Abstract

In Assessment for Learning (AfL), assessment is a tool used to support students' learning. In the last decades, research has shown the potential of AfL on student learning. Implementation of AfL in practice however, has proven to be complicated and a concrete discriminative operationalization of AfL practice is lacking. Aim of this study is to support the implementation of AfL and to operationalize AfL in an observation instrument for AfL practice.

For the design of the observation instrument, a literature search was conducted to find an applicable theoretical framework for AfL and to describe teachers' activities in AfL practice in concrete processes. The search resulted in the identification of the three main processes of AfL for teachers, students and peers: 1) identifying where learners are going; 2) where they are in their learning and 3) how to get there, as the theoretical framework for the design of the observation instrument. Theory on feedback (dialogues) and monitoring and scaffolding was used to describe teachers' activities in the three main processes of AfL. The resulting observation instrument is presented in this paper.

To determine if the instrument meets its intended goals, the instrument was tested in observing teachers in their practice. Results show that the instrument is appropriate to identify and discriminate between AfL practices and that the instrument is suitable for coaching teachers in their change to AfL.

Introduction

In the last few decades, Assessment for Learning (AfL) has had large research attention (Segers & Tillema, 2011; Swaffield, 2011). In AfL, assessment serves as a tool for learning (Gardner, Harlen, Hayward, Stobart, & Montgomery, 2010). Since Black and William's (1998) influential literature review on AfL and student learning, the number of studies that confirms the positive effect of AfL on student learning is growing (Birenbaum, Kimron, & Shilton, 2011). Positive effects are reported on student attitudes towards school and learning, student motivation, higher-order thinking and self-regulation (Black & Wiliam, 1998).

There is however, also growing consensus about the complicated implementation of AfL (Segers & Tillema, 2011; Wiliam, Lee, Harrison, & Black, 2004). Implementation of AfL requires teachers' professional development (Smith, 2011). The proposed strategy for teachers' professional development, is to do a personal inquiry on AfL in own classroom practice (Gardner et al., 2010; Smith, 2011). In this personal inquiry, the teacher applies AfL in practice to create a personal action theory (Smith, 2011). This inquiry in AfL however, has

proven to be complex and time-intensive and affects teachers' pedagogy of teaching fundamentally (Black & Wiliam, 1998; Gardner et al., 2010; Smith, 2011; Wiliam et al., 2004).

Another problem with the AfL concept, is that there is still lack of agreement about the artefacts that represent AfL and how AfL might work in practice (Bennett, 2011; Swaffield, 2011; Torrance, 2012). This lack of agreement is due to the initial ill-defining of AfL and the fuzziness about the differences between AfL and the closely related concept of formative assessment (Wiliam, 2011). This on-going discussion on AfL, has resulted in a considerable number of valuable theoretical studies on AfL and in more assessment-related concepts like 'Assessment *as* Learning' (Torrance, 2007), conformance, deformative and transformative assessment (Torrance, 2012). This discussion is of great importance, but has been going on a high theoretical level and is just scarcely supplemented with empirical research. One example of empirical research on AfL is Pat-el and colleagues' (2011) questionnaire study on teachers' and students' perceptions of AfL practice. Their results showed robustness of the two identified constructs 'monitoring' and 'scaffolding', but also the difficulty to discriminate between different assessment practices. In sum, there is a considerable number of valuable studies on AfL on a conceptual level, whereas a concrete operationalization of how AfL works in practice to discriminate between assessment practices, is scarce (Bennett, 2011).

The aforementioned problems, the complicated implementation and the lack of a concrete discriminative operationalization of AfL practice, are related. Due to the lack of a suitable operationalization of AfL, teachers will not be able to benchmark their actual assessment practice and to discriminate between assessment practices. To identify how AfL is applied in practice, it needs to be clear what elements in assessment practice can or cannot be characterized as AfL. The lack of an operationalization to discriminate between AfL practices, hampers implementation of AfL because it's impossible to pinpoint what elements of assessment practice need improvement to change into AfL. Because concrete instantiations of AfL are lacking, the discussion on AfL is just on theoretical level and cannot be broadened to a more practical level.

Aim of this study is to support the implementation of AfL and to contribute to the theory, by operationalizing AfL in an observation instrument to identify and typify teachers' AfL practice. Doing observations is a fundamental research method in social sciences (Angrosino, 2005) and an appropriate method for identifying quality of educational practice (Praetorius, Lenske, & Helmke, 2012). Another motivation for the observation method is that Pat-el and colleagues (2011) survey method did not bring forth any discriminating results between teachers.

Identifying and typifying AfL practice, serves to support teachers in improving their AfL practice. Identifying AfL and discriminating between practices will mean a major contribution for implementing AfL in practice, because it will support teachers' personal inquiry and the coaching of experimenting schools. Moreover, operationalization of AfL will contribute to common conceptual understanding about AfL and will allow future empirical research on AfL practice and its impact.

Research questions in this study are: 1) What theory must be integrated into an observation instrument to pinpoint, differentiate and characterise teachers' activities in AfL practice?; 2) Is the observation instrument suitable to pinpoint, differentiate and characterise teachers' activities in AfL practices? In this paper, the literature study for research question one, is brought together in the theoretical framework section. Aim of the literature study was to find an applicable theoretical framework for AfL and to find descriptions of activities which could function as clues to describe AfL in concrete processes within the identified AfL framework. The resulting observation instrument is presented in the first part of the results section. To determine the practical and scientific relevance of the instrument, the instrument

was tested with a co-observer. The results of this procedure as result for research question two, are reported in the second part of the results section.

Assessment for Learning

Klenowski (2009) formulated an acknowledged definition for AfL. He stated that AfL is: “Part of everyday practice by students, teachers and peers that seeks, reflects upon and responds to information from dialogue, demonstration and observation in ways that enhance on-going learning” (Klenowski, 2009, p. 264). Reflected in this definition is the aim of AfL is to promote learning itself with the premise that this contributes to students’ performance (Klenowski, 2009). In AfL, the information gained by teacher, students and peers during classroom activities, is used to adapt teaching and to promote students’ learning (Black & Wiliam, 1998). Strategies are: sharing criteria for success, providing timely feedback and applying peer- and self-assessment (Cooper & Cowie, 2010).

AfL is often falsely equated with formative assessment. Formative assessment and AfL are both part of everyday teaching practice and offer the teacher and students insight in the areas that need more attention. AfL however, refers to the assessment purpose while formative assessment refers to the function the assessment actually serves (Bennett, 2011). Another distinction is that in AfL, the assessment gives insight to the student in how to make learning progress. Formative assessment does not necessarily hold that characteristic (Wiliam, 2011).

AfL is often misinterpreted as continuous summative assessment (Derrick, Gawn, & Ecclestone, 2008). In this interpretation, only the procedures for AfL are applied: the ‘letter’ of AfL (Marshall & Drummond, 2006). This is different than applying the ‘spirit’ of AfL, in which AfL is used to stimulate students’ autonomy (Marshall & Drummond, 2006) and which is at the heart of genuine AfL (Swaffield, 2011). In this paper, AfL is interpreted as the ‘spirit’ of AfL.

Methodology

Literature study

The literature study for the first research question was restricted to articles in peer-reviewed journals, dissertations and books on AfL or on related topics. The selected literature was selected on two main criteria: 1) Presentation of an applicable widely acknowledged theoretical framework for AfL; 2) Descriptions of phases, procedures which could function as clues to describe AfL in concrete processes.

Procedure

The observation instrument was designed by the first author in cooperation with the second and third author. The first author produced a concept on which author two and three provided feedback. The feedback was processed by the first author before the next discussion with the same co-authors. This procedure was repeated four times. After the third discussion, the instrument was also submitted to two colleague researchers, experienced in doing observations, and to the involved teacher participants to get feedback on the concept. When feedback raised theoretical questions, literature was screened again on that aspect and if needed processed in the instrument. The actual designing process of the instrument was thus iterative in nature. Eventually, this resulted in a preliminary version of the observation instrument with 29 items. Each item was formulated in such a way that it was distinctive, to

discriminate between assessment practices. The activities were described in observable behaviour to avoid ambiguous interpretations (Good & Brophy, 2003).

Participants

The instrument was used in a longitudinal multiple-case study on teachers' professional development in AfL. Seven teachers participated in the study. The participants teach at a school for Agricultural Vocational Education and Training (AVET) in the Netherlands. Five teachers teach at prevocational level; of which three teach in the subject 'Human and nature', one in 'English' and one in 'Landscaping'. Two teachers teach at senior secondary level; one in 'floristry' and one in 'animal-husbandry'.

The participants engage in the educational reform project 'Power of assessment'. Project's goal is to stimulate teachers' professional development in AfL with improved AfL practice as desired outcome.

Observations

For the second research question, if the instrument meets the intended goals, the instrument was used in observations to provide feedback on the assessment practice in lessons of the participating teachers. The lessons were video-recorded and observed by the first author. In the lesson's introduction, the observer was introduced by the teacher and the students were prompted not to pay attention to the observer. The observer did not interfere with the natural flow of the classroom to avoid bias (Good & Brophy, 2003). Observations took place in February 2012.

Immediately after the observation, the author and teacher discussed the occurrence of AfL in the observed lesson using the instrument. Herewith, facilitating teachers' understanding of AfL and recognizing concrete strengths and possible improvements in their practice. Afterwards, the video-recording was coded by the first author by putting a mark per item if it was present or not.

To be able to determine the suitability of the instrument, three lessons were coded by a second observer by watching the video-recordings. The second observer was an experienced researcher in education and had expertise on AfL. Afterwards, the given codes (if item was present or not) were compared to determine the inter-rater reliability.

Theoretical framework

Theoretical framework for AfL

AfL involves three main processes for teachers, students and peers: 1) identifying where learners are going; 2) where they are in their learning and 3) how to get there (Wiliam, 2011). These processes are derived from Sadler's (1989) study on formative assessment (Balan, 2012).

In '*where the learner is going*', the teacher clarifies the learning goals and shares the criteria for success for which the students make an effort to understand (Wiliam, 2011). The intention of this first main process is to make sure that teachers and students have a common reference to direct their activities (Sadler, 1989) and that the areas which need improvement are 'scaffolded' (Pat-El et al., 2011). Students' understanding and teachers' communication of learning goals and assessment criteria helps students to focus their learning and improves their performance (Balan, 2012).

In '*identifying where learners are*', teachers provide opportunities for students to learn and observe students' performance (Wiliam, 2011). Observation of students' performance can help teachers to offer extra support to students and to reflect on the provided instruction

(Balan, 2012). Students are working on their learning task (in interaction) with their peers (William, 2011). The learning tasks offer students a meaningful and demanding challenge for which they can set their own self-referenced goals (Black & William, 1998). Students work in interaction or collaboration with peers because of the positive effects on their achievement (Balan, 2012).

In *'how to get there'*, the student is actively regulating own or its peers' performance (William, 2011). Regulation of learning is regarded as ultimate goal of AfL (Birenbaum et al., 2011). Students regulate their learning by processes like planning, monitoring, control and reflection (Pintrich, 2004). The teacher regulates students' performance by providing feedback on the observed gap between the students' actual performance and the intended learning intentions, aiming to move learners forward and to close the observed gap (Birenbaum, Kimron, Shilton, & Shahaf-Barzilay, 2009; William, 2011). In closing this gap, 'monitoring' to check student progress is essential (Pat-El et al., 2011).

William's (2011) three main processes in AfL build on earlier research and are based on different meta-studies (Balan, 2012). The three processes, summarize the core of AfL in three *easy-to-grasp-questions* for each of the *three involved actors*. Moreover, it's a widely acknowledged framework with descriptions of strategies involving the teaching-learning environment of AfL. The descriptions help to get a more detailed picture of teachers', students' and peers' activities in AfL and help to discriminate between assessment practices. The firm theoretical grounding, the well understandable three questions, the distinguished three actors, the wide acknowledgement of the framework and the descriptions of the strategies to help discriminate between practices, are the motivation to use the framework as foundation for the design of the observation instrument.

Feedback

To further describe teachers' activities in the three AfL processes, theory on feedback is used because feedback is a central concept in AfL (Balan, 2012). Feedback can operate on four different levels: 'task level' (whether work is correct or incorrect), 'process level' (the process used to create a product), 'self-regulation level' (self-monitoring and regulating of actions) and 'self level' (directed to the 'self' of the learner). Feedback on self-level is the least effective for student learning. Most effective is to guide students from task-level, to process-level and subsequently to self-regulation level (Hattie & Timperley, 2007). The effectiveness of feedback also relies on students' capacity and the opportunity to process and act on feedback (Sadler, 1989). Feedback is more effective if the teacher demands a response to the information given (Balan, 2012). A response which can naturally be given in feedback dialogues in which feedback and information about learning and performance are exchanged by teacher and student (Van der Schaaf, Baartman, Prins, Oosterbaan, & Schaap, 2011). Students perceive the feedback in feedback-dialogues as more useful and the dialogue stimulates their reflective thinking (Van der Schaaf et al., 2011).

Being one of the agents in AfL, peers are also involved in giving feedback. Peer-feedback enhances students' and peers' learning because both are reconstructing their knowledge when processing and transferring the information (Liu & Carless, 2006). Practicing peer-feedback actively engages students in understanding standards and criteria and will positively affect students' motivation and attitudes (Boud, 2000; Liu & Carless, 2006). Peer-feedback is an essential element in the 'spirit' of AfL: it helps students to become actively engaged in regulating their own or peers' learning.

In summary

William's (2011) framework was processed in a matrix with three lines for teacher, student and peer and three columns for the AfL processes. In each of the nine cells, teachers'

activities were described. To specify teachers' activities in the matrix, the concepts of 'monitoring' and 'scaffolding' were used. These concepts are important in AfL (Pat-El et al., 2011) and give cues to describe concrete teachers' behavior in AfL. The four types of feedback (Hattie & Timperley, 2007) are used to describe teachers' activities because of the direct relation between type of feedback and the aim of AfL, improving student learning. Another cue for describing teachers' activities in AfL is the importance of responding to feedback in feedback dialogues (Van der Schaaf et al., 2011).

Results

In this section, the observation instrument is presented as a result of the literature study. Subsequently, results on the second research question are described.

The observation instrument for AfL practice

As discussed in the methodology section, the observation items are processed in a 3*3 matrix. In the line 'teacher', teachers' activities regarding instruction, observation and feedback provision are described. In the line 'peers', teachers' activities are described in stimulating students to function as learning resources for each other. In the line 'students', teachers' activities to stimulate students in becoming owners of their own learning are described. In figure 1 the final version of the observation instrument is presented. The instrument is supplemented with examples of observed practice on each item to illustrate behaviour that corresponds with the item.

	Where the learner is going	Where the learner is right now	How to get there
Teacher	<p>A1 Teacher indicates the learning goals.</p> <p><i>“Goal is that you know of which animals you are competent in handling and of which animals you need to learn to handle it better.”</i></p>	<p>B1 Teacher compares students’ execution of the learning activity with the assessment criteria.</p> <p><i>“Are you making progress?”</i></p>	<p>C1 Teacher helps students to understand how the execution of the learning activity can be improved.</p> <p><i>“Do you understand my feedback? How could you improve that part?”</i></p>
	<p>A2 Teacher indicates the assessment criteria.</p> <p><i>“The grade will be determined with the criteria on page 2.”</i></p>	<p>B2 Teacher indicates what parts of the learning activity the students performed well or less.</p> <p><i>“You can grab the animal a bit firmer. The rest was fine.”</i></p>	<p>C2 Teacher helps the students to understand how they can use theory to improve the execution of the learning activity.</p> <p><i>“When the latin plant name starts with a capital letter, you know it’s the first name. The first word of a latin plant name always starts with a capital letter.”</i></p>
	<p>A3 Teacher justifies the importance of the learning activity.</p> <p><i>“Its’ important that you pay attention, because this is the last exercise before the assessment.”</i></p>	<p>B3 Teacher indicates the theory the students have applied to get to the performance on the learning task.</p> <p><i>“I notice that you applied the theory on verb spelling in your writing.”</i></p>	<p>C3 Teacher indicates concrete improvements on how to proceed (the next time).</p> <p><i>“Talk to the animals, that’s part of animal husbandry.”</i></p>
	<p>A4 Teacher shares criteria for success.</p> <p><i>“You will be assessed on ‘cooperation and discussion’. Are you discussing together, does one of you perform everything...?”</i></p>	<p>B4 Teacher compares the performed learning behavior with the demanded learning behavior for completion of the learning activity.</p> <p><i>“I saw you reading the assessment criteria several times.”</i></p>	<p>C4 Teacher adjusts instruction on the basis of the results.</p> <p><i>“I noticed that it’s hard to use the right technique in your flower arrangement. Please pay attention, I will demonstrate the technique again.”</i></p>
	<p>A5 Teacher indicates the learning behavior, needed for the completion of the</p>	<p>B5 Teacher indicates the observed students’ learning behavior.</p>	

	<p>learning task.</p> <p><i>“Before you start, please do read the assignment carefully!”</i></p>	<p><i>“You should have filled in the working schedule together. That enables me to address you on your work.”</i></p>	
Peer	<p>A6 Teacher instructs students to jointly set the learning goals.</p> <p><i>“Discuss potential additional personal learning goals during this activity.”</i></p>	<p>B6 Teacher instructs students to use the assessment criteria to assess peers’ performance.</p> <p><i>“Your flower arrangement is ready? Okay, then you can assess it with your peer, using the assessment sheet.”</i></p>	<p>C5 Teacher instructs students to jointly name concrete improvements for executing the learning activity.</p> <p><i>“Discuss future improvements for executing the assignment.”</i></p>
	<p>A7 Teacher instructs students to jointly pay attention to the assessment criteria.</p> <p><i>“Help each other to keep an eye on the assessment criteria to check whether you are still on track.”</i></p>	<p>B7 Teacher instructs students to name strengths in peers’ execution of the learning activity.</p> <p><i>“When assessing your peer’s work, please don’t forget to name his strengths.”</i></p>	<p>C6 Teacher instructs students to jointly name their learning experiences as a result of the learning activity.</p> <p><i>“Discuss what you’ve learned from this assignment.”</i></p>
		<p>B8 Teacher instructs students to name improvements in peers’ execution of the learning activity.</p> <p><i>“When you think that your peer is not handling the animal right. Tell him.”</i></p>	<p>C7 Teacher helps the student to understand how to assess their peers’ performance.</p> <p><i>“I am curious how you graded her work and why.... For technique, I would grade it a bit lower because the oasis is still visible.”</i></p>
Student	<p>A8 Teacher instructs students to set their own learning goals.</p> <p><i>“Set your own additional learning goals in this learning activity.”</i></p>	<p>B9 Teacher instructs students to assess their own performance using the assessment criteria.</p> <p><i>“Assess your own work using the assessment sheet”</i></p>	<p>C7 Teacher instructs students to name own concrete improvements for executing the learning activity.</p> <p><i>“Write down possible improvements in your own assignment.”</i></p>
	<p>A9 Teacher instructs students to pay attention to the assessment criteria.</p>	<p>B10 Teacher instructs students to name own strengths in acting in the learning activity.</p>	<p>C8 Teacher instructs students to name own learning experiences as a result of the learning</p>

	<i>“Keep an eye on the assessment criteria to check whether you are still on track.”</i>	<i>“Don’t forget to name your own strengths!”</i>	activity. <i>“What did you learn from this learning activity?”</i>
		B11 Teacher instructs students to name own weaknesses in acting in the learning activity. <i>“Look critically at your own work to name possible improvements in your work.”</i>	C9 Teacher helps the student to understand how to assess their own performance. <i>“Read the assessment criteria and be critical in assessing your own work.”</i>

Instrument in AfL practice

For the second research question, the instrument was used to determine if the instrument was suitable to pinpoint, differentiate and characterize AfL practice. The researchers’ experiences with the instrument are very positive. During the observation, the integration of William’s framework in the design of the instrument helped to easily determine teachers’ activity in regard to the three AfL processes and to characterize the assessment practice. The items helped to focus on teachers’ activity on a micro-level and helped to identify quality of practice. The items enabled to pinpoint what elements of practice were applied in the ‘spirit’ of AfL (Derrick et al., 2008) and to differentiate between different assessment practices.

In discussing the observed assessment practice, the instrument was used in a similar way. The framework with three lines and three rows helped the teachers to get a grip on what’s AfL and to determine to what extent their practice aligned with AfL. The items enabled a more focused discussion on observed practice and on what specific elements the practice could be improved in terms of AfL. In line with the AfL processes, the instrument supported the teachers to get a grip on the concept AfL, to ‘grade’ their practice and to improve future practice.

Unfortunately, above described experiences cannot be validated and generalized because the second observer has not been able to use the instrument in coding the observations. As a result inter-rater reliability cannot be determined.

Conclusion/discussion

Although the study has not been finished yet, concluding remarks can already be made in relation to the aims of this study. First aim of this study was to contribute to the theory by operationalizing the still-under-construction-concept of AfL (Bennett, 2011) in an observation instrument. This study confirms the theoretical framework of William (2011) which summarizes the core of AfL. The framework proved to be a firm foundation for building an observation instrument. The operationalization of AfL on micro-level in the cells, contributes to the agreement on the concept AfL and enables a discussion on the concept on a more specific level and helps to get a grip on the ‘spirit’ of AfL in practice (Marshall & Drummond, 2006).

Second aim was to determine if this instrument was appropriate to support implementation of AfL. Conclusions in relation to this second aim are rather weak because of the lack of a co-observer's judgment and impression on the instrument. Due to the same circumstance, the inter-rater reliability was not determined yet. However, the researcher's and the observed teachers' experiences show that the instrument certainly helps to identify AfL. The items make it possible to exactly pinpoint the elements that improve quality of AfL practice. The instrument however, focused on teachers' activity in practice, which neglected the general principle to not limit observations in classrooms to teachers' behaviour only (Good & Brophy, 2003). This is even more important in relation to AfL, because teachers' assessment practice and students' learning activities are strongly related (Brown & Hirschfeld, 2008). An important suggestion for improvement of this study and for further research is to design an observation instrument for students' activity in AfL.

Teachers' assessment practice is slow to change into AfL and fundamentally affects their classroom practice (Wiliam et al., 2004). This instrument helps teachers to get to a theoretically grounded personal action theory (Smith, 2011) to change their own practice. This observation instrument has proven to help teachers to identify AfL in practice and to differentiate between practices to coach teachers in their implementation of the 'spirit' of AfL.

Literature

- Angrosino, M. V. (2005). Recontextualizing observation: Ethnography, Pedagogy and the prospects for a progressive political agenda. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research, third edition* (pp. 729-745). Thousand Oaks: Sage.
- Balan, A. (2012). *Assessment for Learning; A case study in mathematics education*. Doctoral thesis, Malmö University, Sweden, Malmö.
- Bennett, R. E. (2011). Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice*, 18(1), 5-25. doi: 10.1080/0969594x.2010.513678
- Birenbaum, M., Kimron, H., & Shilton, H. (2011). Nested contexts that shape assessment for learning: School-based professional learning community and classroom culture. *Studies In Educational Evaluation*, 37(1), 35-48. doi: 10.1016/j.stueduc.2011.04.001
- Birenbaum, M., Kimron, H., Shilton, H., & Shahaf-Barzilay, R. (2009). Cycles of inquiry: Formative assessment in service of learning in classrooms and in school-based professional communities☆. *Studies In Educational Evaluation*, 35(4), 130-149. doi: 10.1016/j.stueduc.2010.01.001
- Black, P., & Wiliam, D. (1998). Assessment and Classroom Learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74. doi: 10.1080/0969595980050102
- Boud, D. (2000). Sustainable Assessment: Rethinking assessment for the learning society. *Studies in continuing education*, 22(2), 151-167.
- Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice*, 15(1), 3-17. doi: 10.1080/09695940701876003
- Cooper, B., & Cowie, B. (2010). Collaborative research for assessment for learning. *Teaching and Teacher Education*, 26(4), 979-986. doi: 10.1016/j.tate.2009.10.040
- Derrick, J., Gawn, J., & Ecclestone, K. (2008). Evaluating the 'spirit' and 'letter' of formative assessment in the learning cultures of part-time adult literacy and numeracy classes. *Research in Post-Compulsory Education*, 13(2), 173-184. doi: 10.1080/13596740802141295

- Gardner, J., Harlen, W., Hayward, L., Stobart, G., & Montgomery, M. (2010). *Developing teacher assessment*. Berkshire, England: Open University Press.
- Good, T. L., & Brophy, J. E. (2003). *Looking in classrooms* (Ninth ed.). New York: Pearson Education, Inc.
- Hattie, J., & Timperley, H. (2007). The Power of Feedback. *Review of Educational Research*, 77(1), 81-112. doi: 10.3102/003465430298487
- Klenowski, V. (2009). Assessment for Learning revisited: an Asia-Pacific perspective. *Assessment in Education: Principles, Policy & Practice*, 16(3), 263-268. doi: 10.1080/09695940903319646
- Liu, N.-F., & Carless, D. (2006). Peer-feedback: the learning element of peer assessment. *Teaching in Higher Education*, 11(3), 279-290.
- Marshall, B., & Drummond, M. J. (2006). How teachers engage with Assessment for Learning: lessons from the classroom. *Research Papers in Education*, 21(2), 133-149.
- Pat-El, R. J., Tillema, H., Segers, M., & Vedder, P. (2011). Validation of Assessment for Learning Questionnaires for teachers and students. *British Journal of Educational Psychology*, no-no. doi: 10.1111/j.2044-8279.2011.02057.x
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16(4), 385-407.
- Praetorius, A.-K., Lenske, G., & Helmke, A. (2012). Observer ratings of instructional quality: Do they fulfill what they promise? *Learning and Instruction* 22, 387-400.
- Sadler, R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, 18, 119-144.
- Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies In Educational Evaluation*, 37(1), 49-54. doi: 10.1016/j.stueduc.2011.03.008
- Smith, K. (2011). Professional development of teachers—A prerequisite for AfL to be successfully implemented in the classroom. *Studies In Educational Evaluation*, 37(1), 55-61. doi: 10.1016/j.stueduc.2011.03.005
- Swaffield, S. (2011). Getting to the heart of Authentic Assessment for Learning. *Assessment in Education: Principles, Policy & Practice*, 18(4), 433-449.
- Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education: Principles, Policy & Practice*, 14(3), 281-294.
- Torrance, H. (2012). Formative assessment at the crossroads: conformance, deformance and transformative assessment. *Oxford Review of Education*, 38(3), 323-342.
- Van der Schaaf, M., Baartman, L., Prins, F., Oosterbaan, A., & Schaap, H. (2011). Feedback Dialogues That Stimulate Students' Reflective Thinking. *Scandinavian Journal of Educational Research*, 1-19. doi: 10.1080/00313831.2011.628693
- Wiliam, D. (2011). What is assessment for learning? *Studies In Educational Evaluation*, 37(1), 3-14. doi: 10.1016/j.stueduc.2011.03.001
- Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: impact on student achievement. *Assessment in Education: Principles, Policy & Practice*, 11(1), 49-65. doi: 10.1080/0969594042000208994