Interdisciplinary Journal of Problem-Based Learning

Volume 11 | Issue 2 Article 6

Published online: 6-19-2017

# Formative Assessment to Support Students' Competences in Inquiry-Based Science Education

Regula Grob
*University of Applied Sciences and Arts Northwestern Switzerland*, regula.grob@fhnw.ch

Monika Holmeier
*University of Applied Sciences and Arts Northwestern Switzerland*, monika.holmeier@fhnw.ch

Peter Labudde
*University of Applied Sciences and Arts Northwestern Switzerland*, peter.labudde@fhnw.ch



Recommended Citation

Grob, R. , Holmeier, M. , & Labudde, P. (2017). Formative Assessment to Support Students' Competences in Inquiry-Based Science Education. *Interdisciplinary Journal of Problem-Based Learning, 11*(2).
Available at: https://doi.org/10.7771/1541-5015.1673

The Interdisciplinary Journal of
Problem-based Learning

SPECIAL ISSUE ON COMPETENCY ORIENTATION IN PROBLEM-BASED LEARNING

# Formative Assessment to Support Students' Competences in Inquiry-Based Science Education

Regula Grob, Monika Holmeier, and Peter Labudde (Centre for Science and Technology Education,

University of Applied Sciences and Arts Northwestern Switzerland)

## Abstract

Inquiry-based education has been part of innovative science teaching for the last few decades. With the competence orientation now underlying many national curricula, one of the emerging questions is how the development of student competences can be fostered in the context of inquiry-based science education. One approach to supporting students in their learning is formative assessment, which is, however, not frequently used in a structured way in daily teaching practice. The aim of this study therefore is to explore what kinds of measures might support science teachers in implementing formative assessment activities in their inquiry-based education. For this, firstly, we investigated the challenges that occur on a classroom level when using formative assessment methods in inquiry-based science education from teachers' perspectives. Secondly, we analyzed the teachers' suggestions on measures of support. Based on the respective results, this paper discusses implications for an implementation of structured formative assessment in inquiry-based science education to enhance student competences.

*Keywords:* student competences, inquiry-based science education, formative assessment

## Introduction

### Inquiry-Based Science Education

Inquiry has been "a distinguishing feature of innovative science education programs since the 1960s science curriculum reform movement" (Duschl, 2003, p. 41). Inquiry-based science education is an umbrella term subsuming a wide range of approaches to teaching and learning, such as inquiry-based teaching and learning, authentic inquiry, model-based inquiry, modeling and argumentation, project-based science, hands-on science, and constructivist science (Furtak, Seidel, Iverson, & Briggs, 2012). In definitions, inquiry-based science education is often described as a set of activities. Bell, Urhahne, Schanze, and Ploetzner (2010) conceptualize such a set of activities, based on a literature review, that they call the "nine main processes of inquiry learning." The nine processes include orienting and asking questions, hypothesis generation, planning, investigation, analysis and interpretation, model, conclusion and evaluation, communication, and prediction. The order of these nine main processes of inquiry learning is not fixed, but arranged as needed and potentially containing iterations (Bell et al., 2010).

The conceptualization of inquiry-based science as a set of activities has the advantage that it is immediately clear what students should learn: By being active in, for example, "hypothesis generation," the students should improve their hypothesizing skills. This perspective is called "inquiry as ends" (Abd El Khalick et al., 2004) and describes inquiry skills (rather than conceptual knowledge) as the main outcome of inquiry activities. In this study, the conceptualization from Bell and colleagues (2010) will be taken as the basic set of activities that are considered inquiry competences. The reason for the choice of this particular conceptualization is its close alignment with the frame provided by Swiss curricula such as Lehrplan 21 (D-EDK, 2014).



http://dx.doi.org/10.7771/1541-5015.1673

**Formative Assessment as a Way to Support the Students' Development of Competences**

An effective way of fostering students' learning is the use of formative assessment, as reported in a number of meta-analyses (Black & Wiliam, 1998; Hattie, 2009). Subsequently, several researchers have also suggested supporting student development of inquiry competences in science education through formative assessment methods (Barron & Darling-Hammond, 2008; Black, Harrison, Lee, Marshall, & Wiliam, 2004; Hume & Coll, 2010).

Formative assessment has the purpose of assisting learning and for that reason is also called "assessment for learning." It involves processes of "seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning and where they need to go and how best to get there" (ARG, 2002, p. 2). These processes are collaborative between teacher and student, and both information about the student's level of knowledge or performance and information about their strengths and areas for improvement allow the teacher to plan subsequent instruction and the student to adapt his or her learning (Cizek, 2010).

Many formative assessment methods have been described for enhancing science learning (e.g., Angelo & Cross, 1993). Some of them are best applicable for diagnosing the students' conceptual understanding (e.g., the traffic lights with which students indicate their level of comprehension; Keeley, 2008). Other methods are expected to provide scaffolds to the students in autonomous learning activities such as inquiry-based learning (Barron & Darling-Hammond, 2008; Ruiz-Primo & Furtak, 2007). From the literature, four ways in which formative assessment supports student learning in autonomous activities have been identified: firstly, the clarification of the intentions of learning assessment criteria (Mansell, James, & the Assessment Reform Group, 2009); secondly, the diagnosis of students' levels of achievement (Maier, 2014) as a basis for subsequent actions; thirdly, the provision of feedback to the individual student in order to decide on the next steps in learning (Wiliam, 2010); and lastly, the fostering of self-regulated learning abilities (Perrenoud, 1998). Suitable methods of formative assessment to support students in autonomous activities include written teacher assessment (the teacher providing written feedback to the individual student; Burke, 2006), self-assessment (students assessing their own work; Andrade & Valtcheva, 2009), and peer-assessment (students assessing their peers' work; Dochy, Segers, & Sluijsmans, 1999).

There is research informing how formative assessment could be articulated in inquiry and in related approaches like problem-based learning (Belland, French, & Ertmer, 2009; Weiss & Belland, 2016). But in teaching practice, structured, formal methods of formative assessment are not widely used (Maier, 2015). Several measures of support for an implementation of formative assessment have therefore been discussed: establishing a continuity between formative and summative assessment (OEDC, 2005; 2013); enhancing the perception of formative assessment from a "soft" and time-consuming approach to a valuable part of the assessment framework (Looney, 2011); and fostering better links between different groups of stakeholders to tackle the lack of coherence in the purposes of assessment between the policy, school, and classroom levels (Looney, 2011). Furthermore, a lack of formative assessment skills or "assessment literacy" amongst teachers is reported (Bennett, 2011; Stiggins, 1999) and subsequently, professional development is suggested. Finally, the possibilities of technology are discussed as an aid (Chudowsky & Pellegrino, 2003) to overcome logistic barriers like large classes with diverse student needs and extensive curriculum requirements (OECD, 2005).

The above measures of support have been suggested by researchers and political stakeholders. However, the success of formative assessment heavily depends on their effective implementation (Black, 1993; Black & Wiliam, 1998; Stiggins, Grisword, & Wikelund, 1989). This is because the quality of formative assessment rests to a high degree on strategies teachers use to elicit evidence of student learning, and on the use of this evidence to shape subsequent instruction and learning (Bell & Cowie, 2001; Heritage, 2010). Consequently, the individual teachers need support in implementing formative assessment in their teaching (Black & Wiliam, 1998; OECD, 2013; Stiggins et al., 1989). Since such support should be adapted to the teachers' difficulties with formative assessment, we investigated the challenges that occur when using the above-introduced formative assessment methods (written teacher assessment; peer-assessment; self-assessment) in inquiry-based science education from teachers' perspectives, and we present our findings in the first part of the results. In order to get first ideas of what measures of support teachers will accept, we also explored their respective suggestions, presented in the second part of the results.

# Methods

**Setting**

In order to explore potential measures of support for teachers in implementing formative assessment activities in their inquiry-based science education, a collaboration lasting for three semesters with 11 science teachers in Swiss upper secondary schools was established. Teachers, all of them volunteers, were recruited by contacting school heads and by announcing the project in professional development courses.

The participating teachers were selected so that variable extents of teaching experience, both genders, and all three science subjects (biology, chemistry, and physics) were covered.

**Data Collection**

In every semester, the teachers were asked to choose one of the following formative assessment methods: written teacher assessment, peer-assessment, or self-assessment; and to trial this method in an inquiry-based unit. The teachers were asked to fill out a self-reporting tool for each of their trials. The self-reporting tool consisted of two parts: In the first part, the teachers reported on their trials (five open questions plus lesson plans and teaching materials). In the second part of the self-reporting tool, the teachers reflected on their units and particularly on the formative assessment activities (eight open questions). The questions focused on benefits and challenges of the formative assessment activity. No later than 14 days after the end of the trials, semi-structured individual interviews of 45 to 60 minutes were conducted with a subsample of $n = 6$ teachers on their trials and on their assessment practices in general. The six teachers were selected so that all three assessment methods were covered. Both written and oral data were transcribed for further analysis.

**Data Analysis**

Twenty trials were completed and sufficiently documented and therefore included in the analysis. Challenges in trialing the formative assessment methods mentioned by the teachers in the self-reporting tools were coded, resulting in 14 codes. The codes were afterwards summarized to five categories of challenges using content analysis (Mayring, 2010). Measures of support for the implementation of formative assessment mentioned by the teachers in the individual interviews were coded resulting in six codes. They were transferred into six categories on measures of support using content analysis (Mayring, 2010). Fifteen percent of the data was coded by two people, resulting in an interrater reliability of $\kappa = 0.89$. Landis & Koch (1977) consider values $\kappa \geq 0.81$ as almost perfect agreement.

## An Illustrative Example of a Trial

In one of his trials, a teacher embedded peer-assessment in a unit on electricity. Its focus was that students should learn about electric capacity and condensators using a model with different boxes and water. The students were expected to explore this model, to conduct measurements on the amount of water in the individual boxes per time, and to document these measurements in their lab journals. They were then expected to describe the analogies between the model and a condensator in their lab journals. After two lessons of inquiry, the students exchanged their journals so that everybody got the report of a peer to assess. For the peer-assessment, the students were provided with a rubric that contained questions focusing on the documentation of the measurements and on the connections between the measurements and the model that were formulated by the students. These questions can be linked to the investigation competence (documentation of measurements) and to the modeling competence (connection between measurements and model), both defined in Bell and colleagues (2010). The students were asked to rate the quality of their peers' lab journals with respect to the criteria and to provide written advice for improvement, which took 20–30 minutes. Afterward, all students got back their own lab journals and their peers' assessment. The students had time to consider the feedback and to adapt their lab journals if needed.

## Results Part 1: Challenges in Using Formative Assessment

The challenges in using formative assessment methods mentioned by the teachers were grouped into five categories: "embedment of formative assessment activities"; "content and structure of the feedback"; "students' engagement with the feedback"; "relation between formative and summative assessment"; and "effort needed."

The challenges related to the *embedment of formative assessment activities* fell into organizational and pedagogical issues. On an organizational level, there were long-term aspects, namely the necessity to carefully plan the formative assessment in the course of a semester so that it did not become too much of a boring routine for the students. But there were also short-term organizational issues: One teacher mentioned the complication that "in peer- and self-assessment, not all students need the same amount of time to provide feedback or to reflect upon their own work." On a pedagogical level, the following challenges were mentioned: firstly, the preset criteria of assessment, which cannot be adapted during the course of the unit. One of the teachers argued that this "greatly hinders flexibility when unexpected aspects arise." Secondly, the choice of the right time for providing feedback in the course of an inquiry was mentioned. A teacher said,

> I conducted the peer-assessment too late, at a point where the students were almost done with their work. The peers could only comment, yes, this is quite ok. . . . The choice of the right timing is the most difficult "in formative assessment": If I do it too early, then they are . . . not far enough in their process and the feedback

> can only be something like "complete your work." If I am too late, almost everybody already [has] their [solution] in the lab journal.

Other challenges related to the *content and the structure of the feedback provided* and were grouped into method-specific subcategories. Considering the written teacher assessment, the decision of what feedback to give, in terms of both focus and extent, was difficult: The teacher has to consciously decide what feedback to share since the quantity of comments students can handle is limited. Very extensive feedback would contradict the student-oriented nature of inquiry-based science education. One of the teachers phrased it this way:

> I always find it difficult to decide to what extent I should guide the students in such student-oriented activities. Of course I see a lot of potential for optimization, but how should I decide whether to interfere or not? Such interference would also lead to a decrease in openness. I also find it hard to provide support for the next steps in learning without giving away the optimal solution.

Considering peer- and self-assessment, the teachers communicated their uncertainty on the quality of the feedback (in peer-assessment) or the reflections (in self-assessment). They were unsure whether the students had the necessary content knowledge, language and communication skills, and abilities to be honest to their peers or to themselves, respectively. Furthermore, the teachers were uncertain whether all students took their task as assessors seriously. A teacher expected "peer-assessment may not always be reliable and understandable." Consequently, another teacher uttered her discomfort by saying "I cannot control all feedback. But when I do see students oversee mistakes, I never know whether I should interfere."

The teachers mentioned challenges relating to the *students' engagement with the feedback received*. These were grouped into three subcategories: Firstly, it was mentioned that the use of the feedback does not only depend on the quality of the feedback but also on the eagerness and discipline of the recipients. One teacher said, "Some students are at times quickly satisfied with their work." Secondly, the students may understand the feedback differently from the teacher or from peers. This may lead to misunderstandings. An example of such a misunderstanding is that a teacher's comment saying "record the results more precisely" may lead to the students recording more data rather than documenting them more exactly. Thirdly, the teachers mentioned that the transfer of the feedback to a new situation is difficult: "It is a hard task for students to transfer my feedback to the next lab journal entry."

The *relation between formative and summative assessment* was mentioned as a challenge several times. Two subcategories were differentiated: firstly, the relevance of formative assessment. More specifically, the teachers addressed the problem that from a student perspective the imminent importance of summative assessment is much higher than of feedback in the context of formative assessment. One teacher said, "In the end, only what is graded is important." Secondly, it was brought up that formative assessment, even though it is supportive, still has the character of a check. A teacher worried that "the formative assessment hinders the joy and the interest in conducting experiments." Similar to this last point, the teachers also mentioned that criteria for assessment (which necessarily have to be provided in the beginning of a unit) hinder the openness and the degree of freedom of an inquiry.

A challenge specified many times was that formative assessment takes some *effort*. This effort is twofold: Firstly, it takes time. One of the teachers mentioned "the preparation time for written teacher assessment," whereas other teachers referred to lesson time in the case of self- and peer-assessment. Secondly, self- and peer-assessment also need a certain amount of practice to be effective. One of the teachers said, "The students need some exercise in assessing and providing feedback before peer-assessment can really improve the artifacts."

## Results Part II: Supportive Measures for the Uptake of Formative Assessment

This second part of the results presents the supportive measures that could facilitate the uptake of formative assessment in everyday science teaching, as mentioned in the individual interviews. The answers are grouped into six categories.

The first measure mentioned was the provision of *examples of good practice*: rubrics; teaching sequences with integrated formative assessment activities that can be used and adapted or can serve as a source of inspiration. One of the teachers said, "There should be a collection with good examples of formative assessment activities somewhere, so that not everybody [has] to reinvent them. These examples could also be adapted or used as inspiration for [their] own teaching."

The second measure was *time*: for planning and preparing the formative assessment activities before the lessons; for conducting formative assessment during regular lessons; and for analyzing the results and deciding on next steps after the lesson. As one of the teachers described it: "Formative assessment contains several steps, and each of these steps need time."

The teachers also mentioned training and coaching to enhance their *assessment literacy*. They had specific questions such as how to develop motivating prompts for self-assessment, how to document oral feedback, and how to provide usable feedback. They also asked for an overview of the different methods of formative assessment from which they could choose depending on a particular context. One of the teachers phrased it this way:

> I would like to build up a [stockpile] of different methods, an overview of what is out there. . . . I also need some ideas on how to make the formative assessment visible for the parents; I need clear statements so that they know what their children have to work on.

The fourth measure of support was *opportunities to reflect on their assessment practices*. One of the teachers described their opportunities as follows:

> To become aware of the function of formative assessment; to become aware of the fact that I have always been doing this but also to learn that there are many more methods and approaches; some of them much more structured and formal than what I have always been doing.

Fifthly, teachers mentioned possibilities to *exchange experiences* about different forms of assessment and to collaboratively develop them with other teachers. One teacher suggested, "A school development project would be very good. This would provide us with the opportunities to exchange ideas and to develop our assessment culture in our school." Another teacher proposed that examples of good practice should be discussed amongst teachers, possibly online, so that these examples could evolve.

The sixth means of support was a *clarification of the role of formative assessment* and also its relation with summative assessment. It was suggested that this should be done in official documents such as curricula or guidelines from educational ministries. One of the teachers said,

> Educational policy at upper secondary school is going into a very different direction at the moment: It is all about certificates and grades. So it is rather difficult for us to fit in the formative assessment just by ourselves, without guidance.

## Interpretation

The first part of the results shows that the teachers who collaborated in the project generally did not have doubts about the eligibility of formative assessment to support students in developing their inquiry competences. Instead, concrete challenges on a classroom level were presented. These suggest that the teachers might be supported by providing advice in embedding formative assessment in their long-term and in their short-term planning, in choosing appropriate foci to feedback upon, in handling self- and peer-feedback, in fostering the students' use of the feedback, in clarifying the role of formative assessment and its relation to summative assessment, and in keeping the effort needed for good formative assessment affordable.

The second part of the results suggests that there are different levels of support: The teachers do not only recognize the potential to tackle the above-mentioned challenges in the classroom themselves by enhancing their assessment literacy, by reflecting upon their assessment practices, and by exchanging their experiences and co-constructing knowledge with their peers. Instead, they also see other stakeholders in charge of an enhanced education. These other stakeholders include the providers of teaching materials such as textbook authors, but also the people working at the level of educational policy.

Most of the challenges could be approached with different measures of support: The provision of examples of good practice, the enhanced assessment literacy, the reflection upon assessment practices, and the co-construction of knowledge amongst peer-teachers could all feed into overcoming several challenges mentioned in the first part of the results; namely, the embedding of formative assessment into a unit, the content and the structure of the feedback, the students' use of the feedback, and, to some extent, also the effort needed. Time as a means of support matches the effort needed on the challenge side. The clarifications of the assessment policy will tackle the unclear relation between formative and summative assessment. Overall, the teachers' perceptions of challenges with the different assessment methods in the classroom appear consistent with the measures of support on different levels they suggest.

Challenges related to designing the assessment activities as expected from the literature were not mentioned by the teachers: Nobody mentioned difficulties in formulating assessment criteria, in finding an appropriate artifact to diagnose student learning, or in diagnosing student learning as a teacher. A possible interpretation is that teachers are used to these activities from summative assessment.

Comparing the measures of support suggested in the international literature (from the researcher and the educational policy perspectives) to the results of this study (from teachers' perspectives), similarities and differences occur: The clarification and strengthening of the role of formative assessment (Looney, 2011; OEDC, 2005; 2013) is highlighted, similar to the ideas of the Swiss teachers. The links between different groups of stakeholders (Looney, 2011) were not mentioned by the teachers in this study, perhaps because they mostly consider the problem at the classroom level. The lack of formative assessment skills is identified in the literature (Bennett, 2011; Stiggins, 1999), but approached by providing professional development rather than the many different measures voiced by the teachers in this study. The teachers' suggestions (e.g., reflections on assessment practices, exchange of experiences, and co-construction of knowledge on assessment) reflected the idea that the teachers wish to develop their own teaching themselves rather than having new strategies imposed externally. The possibilities of technology are suggested in the literature (Chudowsky & Pellegrino, 2003; Looney, 2011) but not in this study, perhaps because Swiss classrooms are generally conservative in that respect. Finally, the logistical barriers such as class sizes and extensive curricula identified in the literature (OECD, 2005) are presented in a similar fashion by the teachers in the study.

Overall, the measures of support suggested from the teachers in this study are rather consistent with the suggestions from researchers and educational policy in the literature. On a more detailed level, the teachers gave more weight to and had more ideas on how to tackle the pedagogical challenges in the classroom compared to the international literature where professional development programs are suggested. In this study, the teachers gave the impression that they wish to improve their assessment habits with bottom-up strategies, based on their own experiences and collaborations with other teachers. Classical top-down strategies from the international literature were not mentioned by the teachers in the study. The high autonomy of the teachers at Swiss upper secondary schools, with the respective responsibilities for assessment, might be an explanation for these results.

## Conclusion

The aim of this study was to explore what kind of measures might support science teachers in implementing formative assessment activities to foster student competences in inquiry-based education. For this, we analyzed the challenges that occur when using formative assessment methods in inquiry-based science education from teachers' perspectives.

With the small sample size, the results are not representative of a broader community of teachers. However, since the formative assessment methods were trialed in authentic lessons, the results allow deriving first hints on what supportive measures might help the implementation of formative assessment activities at the level of upper secondary school in Switzerland.

The results of the study show that the teachers perceived a number of pedagogical challenges when using formative assessment methods in their inquiry-based education to support student learning. The measures of support suggested by the teachers appear suitable to tackle these challenges. However, a series of interventional studies with control group(s) are needed to test these hypotheses. Comparing the teachers' perspectives explored in this study to the perspectives of researchers and educational policy makers, the teachers appear to favor bottom-up strategies where they develop their assessment strategies themselves rather than receiving help externally. This tendency should be further investigated in future research and considered when tailoring the implementation strategy of formative assessment.

## References

Abd El Khalick, F., Boujaoude, S., Duschl, R. A., Lederman, N. G., Mamlok-Naaman, R., Hofstein, A., Niaz, M., Treagust, D., & Tuan, H. (2004). Inquiry in science education: International perspectives. *Science Education, 88*(3), 397–419.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory Into Practice, 48*(1), 12–19.

Angelo, T., & Cross, K. P. (1993). Classroom assessment techniques: A handbook for college teachers. San Francisco: Jossey-Bass.

ARG (Assessment Reform Group) (2002). *Assessment for learning: 10 principles.* London: ARG.

Barron, B., & Darling-Hammond, L. (2008). Teaching for meaningful learning: A review of research on inquiry-based and cooperative learning. In L. Darling-Hammond, B. Barron, P. D. Pearson, A. H. Schoenfeld, E. K. Stage, T. D. Zimmermann, G. N. Cervetti, & J. Tilson (Eds.), *Powerful learning: What we know about teaching for understanding* (pp. 11–70). San Francisco: Jossey-Bass.

Bell, B., & Cowie, B. (2001). The characteristics of formative assessment in science education. *Science Education, 85*(5), 536–553.

Bell, T., Urhahne , D., Schanze, S., & Ploetzner, R. (2010). Collaborative inquiry learning: Models, tools, and challenges. *International Journal of Science Education, 32*(3), 349–377.

Belland, B. R., French, B. F., & Ertmer, P. A. (2009). Validity and problem-based learning research: A review of instruments used to assess intended learning outcomes. *Interdisciplinary Journal of Problem-Based Learning, 3*(1), 59–89.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice, 18*(1), 5–25.

Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). *Working inside the black box: Assessment for learning in the classroom.* London: GL Assessment.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice, 5*(1), 7–73.

Burke, K. (2006). *From standards to rubrics in 6 steps*. Heatherton, Victoria: Hawker Brownlow Education.

Chudowsky, N., & Pellegrino, J. W. (2003). Large-scale assessments that support learning: What will it take? *Theory into Practice, 42*(1), 75–83.

Cizek, G. (2010). An introduction to formative assessment. In H. Andrade & G.J. Cizek (Eds.), *Handbook of formative assessment* (pp. 3–17). New York: Routledge.

D-EDK Deutschschweizer Erziehungsdirektoren-Konferenz (2014). Lehrplan 21. Luzern: D-EDK.

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-peer and co-assessment in higher education: A review. *Studies in Higher Education, 24*(3), 331–350.

Duschl, R. A. (2003). Assessment of inquiry. In J. M. Atkin & J. E. Coffey (Eds.), *Everyday assessment in the science classroom* (pp. 41–59). Arlington, VA: NSTA press.

Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research, 82*(3), 300–329.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London & New York: Routledge.

Heritage, M. (2010). *Formative assessment: Making it happen in the classroom*. Thousand Oaks, CA: Corwin Press.

Hume, A., & Coll, R. K. (2010). Authentic student inquiry: The mismatch between the intended curriculum and the student-experienced curriculum. *Research in Science & Technological Education, 28*(1), 43–62.

Keeley, P. (2008). Science formative assessment. 75 practical strategies for linking assessment, instruction, and learning. Thousand Oaks, CA: Corwin Press.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159–174.

Looney, J. W. (2011). Integrating formative and summative assessment: Progress toward a seamless system? *OECD Education Working Papers No. 58*. Paris: OECD.

Mansell, W., James, M., & the Assessment Reform Group. (2009). *Assessment in schools: Fit for purpose? A commentary by the Teaching and Learning Research Programme*. London: TLRP.

Maier, U. (2014). Formative Leistungsdiagnostik in der Sekundarstufe—Grundlegende Fragen, domänenspezifische Verfahren und empirische Befunde. In M. Haselhorn, W. Schneider, & U. Trautwein (Eds.), *Lernverlaufsdiagnostik* (pp. 19–40). Göttingen: Hogrefe Verlag.

Maier, U. (2015). *Leistungsdiagnostik in schule und unterricht: Schülerleistungen messen, bewerten und fördern*. Bad Heilbrunn: Verlag Julius Klinkhardt.

Mayring, P. (2010). *Qualitative inhaltsanalyse: Grundlagen und techniken*. Weinheim: Beltz.

Organisation for Economic Co-operation and Development (OECD) (2005). *Formative assessment: Improving learning in secondary classrooms*. Paris: OECD Publishing.

Organisation for Economic Co-operation and Development (OECD) (2013). *Synergies for better learning: An international perspective on evaluation and assessment*. *OECD Reviews of Evaluation and Assessment in Education*. Paris: OECD Publishing.

Perrenoud, P. (1998). From formative evaluation to a controlled regulation of learning processes: Towards a wider conceptual field. *Assessment in education: Principles, policy and practice, 5*(1), 85–102.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding of the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57–84.

Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice, 18*(1), 23–27.

Stiggins, R. J., Griswold, M. M., & Wikelund, K. R. (1989). Measuring thinking skills through classroom assessment, *Journal of Educational Measurement, 26*(3), 233–246.

Weiss, D., & Belland, B. R. (2016). Transforming schools using project-based learning, performance assessment, and Common Core standards. *Interdisciplinary Journal of Problem-Based Learning, 10*(2).

Wiliam, D. (2010). Research literature and implications for a new theory of formative assessment. In H. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 18–40). New York: Routledge.

Regula Grob is a PhD candidate in the Centre for Science and Technology Education at the University of Applied Sciences and Arts Northwestern Switzerland. Her doctorate focuses on formative assessment in inquiry-based science education.

Dr. Monika Holmeier is a senior researcher in the Centre for Science and Technology Education at the University of Applied Sciences and Arts Northwestern Switzerland. Her research focuses on school governance, quality of instruction, and assessment. She is a lecturer for qualitative and quantitative research methods and leads several research and evaluation projects in science education.

Prof. Dr. Peter Labudde is the head of the Centre for Science and Technology Education at the University of Applied Sciences and Arts Northwestern Switzerland. His research focuses on learning processes in science education, competence models and educational standards, and international comparative studies on science education, etc.