

A TEACHER PERSPECTIVE ON BENEFITS AND CHALLENGES OF PEER-ASSESSMENT

Regula Grob^{1,2}, Monika Holmeier¹ and Peter Labudde¹

¹Center for Science and Technology Education, University of Applied Sciences and Arts
Northwestern Switzerland, Basel, Switzerland

²University of Teacher Education, Fribourg, Switzerland

Formative assessment has been suggested as a means to support student learning in inquiry-based science education. However, teachers need support in implementing formative assessment practices, such as peer-assessment, in their daily teaching. As a prerequisite for shaping suitable means of support, primary and upper secondary teachers' perspectives on benefits and challenges of peer-assessment in inquiry learning have been explored. Data was collected from 7 primary and 10 upper secondary school teachers from Switzerland who implemented peer-assessment in their science classes. The data included teaching plans, evaluation forms, individual interviews, and group interviews. Inductive coding of the data revealed that the teachers perceived challenges of peer-assessment at the level of teaching practice but also at the level of educational policy. These results suggest that different measures of support such as professional development programmes, but also concrete examples and tools as well as guidelines from educational policy are needed. Considering the benefits of peer-assessment, the teachers from both school levels did not only believe that peer-assessment enhances student learning but also anticipated social and motivational effects. This result implies that formative assessment theories should be more closely connected to learning theories in which student motivation has been identified as a main contributor to learning.

Keywords: formative assessment, peer-assessment, inquiry-based science education

INTRODUCTION

Problem statement

Inquiry and other competence-oriented approaches have become important parts of science education in the recent decades. One issue, however, has been how to support students in their inquiry learning and how to assess respective student competences (e.g. Harlen, 2013). A possible answer to this is the promotion of formative assessment at an international (e.g. OECD, 2005; 2013), but also at a national level (e.g. in the curriculum for the compulsory school levels for the case of Switzerland, D-EDK, 2014). But as a number of studies show, the use of formative assessment in teaching practice varies greatly between teachers (Black, 1993; Bell & Cowie, 2001; Heritage, 2010; Herman, Osmundson & Silver, 2010; Stiggins, Griswold & Wikeland, 1989). The quality of formative assessment rests to a high degree on the 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; Ruiz-Primo, Furtak, Ayala, Yin, &

Shavelson, 2010). Subsequently, the need of help for the teachers is stated: “Simply embedding assessments in curriculum does not guarantee improved learning and teaching. Teachers need tremendous support using assessment in their teaching practice” (Yin, et al., 2008, p. 356). The focus of this study will therefore be on science teacher perspectives on peer-assessment, a formative assessment method relatively well-described in the literature (e.g. Topping, 2003), in the context of inquiry learning.

Literature review

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” (Assessment Reform Group ARG, 2002, p. 2). The following four characteristic features for an operationalisation of ‘formative assessment’ were found: (1) Clarity in expectations (e.g. Black, Harrison, Lee, Marshall & Wiliam, 2004); (2) Diagnosis of student level with respect to expectations (Ruiz-Primo et al., 2010); (3) Presence of feedback (Furtak & Ruiz-Primo, 2008) (4) Opportunity to use this feedback (e.g. Andrade, 2010).

For the context of inquiry-based science education, a number of concrete methods of formative assessment have been suggested (e.g. Barron & Darling-Hammond, 2008). The focus of this study will be on peer-assessment which is defined as a process in which students assess their peers’ work and provide feedback on it (e.g. Topping, 2003). Peer-assessment follows the idea of "activating students as instructional resources for one another" (Leahy, Lyon, Thompson & Wiliam, 2005, p. 21): Students take both the role of the assessor and the assessee by assessing each other’s work. The aim of peer-assessment is to assist peers in identifying the strengths and weakness of their work and to provide suggestions for improving it (Dochy, Segers & Sluijsmans, 1999; Topping, 2003).

A number of advantages and challenges that are associated with peer-assessment have been identified in the literature. The advantages of peer- assessment are, firstly, that feedback from peers who had the same difficulties in the learning progress might suggest direct ways to overcome those difficulties, and formulate them in a language that is naturally used by the students (Black et al., 2004). Secondly, students who assess their peers’ work engage in cognitively demanding activities, such as critical thinking (Hanrahan & Isaacs, 2001; Harlen, 2007; Lin, Liu & Yuan, 2001; Lindsay & Clarke, 2001; Topping, 2003; Tsivitanidou, Zacharia & Hovardas, 2011). Thirdly, students get the opportunity to see examples of other students’ work. This can potentially lead to self-assessment: By comparing their own work to that of their peers, students can be prompted to reflect on their own learning achievements (Hanrahan & Isaacs, 2001; Lin et al., 2001; Topping, 1998; 2010). Fourthly, peer-assessment may be easier to accept since it is perceived less authoritative than feedback from adults and therefore open to negotiation (Cole, 1991; Topping, 2010). Fifthly, feedback from peers can be more immediate, timely, and individualized than feedback from the teacher (Topping, 2010) simply because there are many more students than teachers in a classroom. Lastly, providing feedback

to peers develops the social, communicative, meta-cognitive and other personal and professional skills on the way (Topping, 2010).

Beside the aforementioned advantages, a number of challenges of peer-assessment have also been identified in the literature: When doing peer-assessment, students need to judge the performance of a peer. This needs a certain degree of knowledge in the field that is assessed (Topping, Smith, Swanson & Elliot, 2000). Furthermore, students need to communicate the judgments to their peers and need to provide constructive feedback about their learning process for which communication skills are necessary (Black, Harrison, Lee, Marshall & Wiliam 2003). Thirdly, the recipients need to critically review the feedback and decide on the actions to be taken: Since peer-feedback might include flaws, the recipients need to filter it and then decide whether there is a need to adopt the peers' suggestions and to revise their work (Sluijsmans, 2002). Fourthly, peer-assessment costs lesson time for organization, training and monitoring, particularly in the beginning, if it should be provided at a good level of quality (Topping, 2010). Lastly, social processes influence and contaminate the validity and reliability of assessment provided by peers (Topping, 2010).

Statement of intentions

Following the problem statement, the exploration of the teacher perspective on formative assessment methods such as peer-assessment is considered relevant for a successful implementation of respective approaches. Teachers' perceptions of the benefits and challenges of peer-assessment will therefore be investigated and the implications for supportive measures for the implementation of peer-assessment in inquiry-based science education will be discussed. Furthermore, a widening of the conceptual framework for formative assessment is suggested based on the results.

METHODS

For this study, a 3-semester cooperation with 20 science teachers in Switzerland (9 primary, 11 upper secondary) was established. In every semester, the teachers incorporated a formative assessment method from a pre-defined list (including peer-assessment) in one of their normal inquiry units. The methods were used to assess one or several student competences from another pre-defined list (including, for example, investigation competence, argumentation competence, and modelling competence). The cooperation also included regular meetings with all the teachers, and a teacher manual on the assessment methods which also included illustrative examples.

Data collection

The teachers provided their teaching plans and -materials (student worksheet etc.) from their trials and filled out an evaluation form in which they reflected upon the benefits and challenges of the assessment method. No more than ten days after the trials, individual interviews were

held with a sub-group of the teachers (consisting of $n=8$ teachers from both school levels) in order to speak about the trials and about general issues related to assessment in more detail.

Data analysis

Based on the teaching plans and the teaching materials, it was decided whether the trials included a formative assessment activity. This was evaluated with the four characterizing features of formative assessment as introduced in the literature review. Afterwards, it was decided whether the formative assessment activity was peer-assessment. The respective criterion was whether the students diagnosed and provided feedback on their peers' work. This resulted in 7 primary and 10 upper secondary school cases.

For the analysis of the benefits and the challenges of peer-assessment, the evaluation forms ($n=17$ evaluation forms) and the transcripts from the individual interviews ($n=8$ interviews) were inductively coded. This led to a coding frame with 8 categories for the challenges and 5 categories for the benefits which will be presented in the results part. 18% of the data was double-coded ($\kappa = 0.89$).

RESULTS

Looking at the challenges, the teachers mentioned difficulties related to the planning of peer-assessment activities (challenge 1). Furthermore, the teachers expressed their doubts about the quality of the diagnosis done by peers (challenge 2), about the quality of the feedback provided by peers (challenge 3), and their uncertainty about their own role (challenge 4). The teachers also anticipated that some of the students might not consider the feedback received from peers to revise their work (challenge 5) or that assessing peers could be boring for students (challenge 6). Another aspect was the role of peer-assessment within the assessment framework, for example the relation between peer-assessment and grading from the teacher (challenge 7). Peer-assessment was also considered rather time-intensive and dependent on a good training of the students (challenge 8).

Considering the benefits, the teachers mentioned that the feedback is provided in a language that is naturally used by the students and it is accepted because the assessor is a peer (benefit 1). Furthermore, the responsibility for the learning in a peer-assessment setting lies with the students, resulting in a lower workload for the teachers and a higher capacity for individual support (benefit 2). The teachers anticipated learning effects in inquiry-specific but also in transversal competences (benefit 3) as well as effects on the classroom climate and the students' motivation (benefit 4). Lastly, the low preparation time for the teacher (benefit 5) was mentioned.

One of the emerging results from the benefits of peer-assessment as mentioned by the teachers is that the teachers from both school levels did not only perceive learning effects (see benefit 3) from peer-assessment but also social and motivational effects (see benefit 4; illustrative quotes: "Peer-assessment enhances the relation between the students"; "Peer-assessment is a

way to take students serious and to give value to what they say. This motivates them in their work”). This aspect will be discussed in more detail in the next section of the paper.

DISCUSSION AND CONCLUSIONS

Comparison of the results to the literature

Comparing the benefits and challenges of peer-assessment as mentioned by the teachers in the study to the results found in the literature, a number of aspects are similar. The specific language characteristics of feedback formulated by peers and the responsibility for learning have been previously reported in Black et al. (2004). No references on the resulting capacities of the teachers were found in the research literature, however. The effects of peer-assessment on the students’ transversal competences (Topping, 2010) and on self-regulated learning (Hanrahan & Isaacs, 2001; Lin et al., 2001; Topping, 1998; Topping, 2010) have also been previously mentioned but not the effects on the classroom climate and on the students’ motivation as anticipated by the teachers in this study. The preparation time was not covered in the literature either.

Considering the challenges, the planning issues as brought up by the teachers in this study are not mentioned in the literature. The quality of the diagnosis (Topping et al., 2000; Topping, 2010) and the quality of the feedback (Black et al., 2003) have been previously discussed. The uncertainty about the own role that resulted, according to the teachers in this study, from the questionable quality of the diagnosis and the feedback, was not found in the literature. The lesson time and the training needed were recognized by Topping (2010), too. None of the teachers in the study spoke about the difficulties in what feedback to use for revision as reported in Sluijsmans (2002).

Overall, the benefits of peer-assessment perceived by the teachers in this study are similar to what is mentioned in the research literature. These effects appear to be independent of the school level and the country-specific context. The social and motivational benefits from peer-assessment have not been found in the literature, though. This will be discussed in more detail in the paragraph ‘widening of the theoretical concept needed’ below.

The challenges of peer-assessment in the literature were not specifically focussed on the perspective of the teachers nor on organisational issues, resulting in a smaller congruence between the results of this study and the research literature. However, it becomes apparent that the challenges of peer-assessment cannot be neglected.

Support needed

The challenges of peer-assessment appear to need support at different levels to be overcome: Professional development as well as concrete teaching resources could help teachers to enhance their own assessment literacy (see challenges 1, 4) but also to let the students improve their abilities in diagnosing, providing and using peer-feedback (see challenges 2, 3, 5, 6, 8). The role of peer-assessment in the assessment framework (see challenge 7) was the only challenge

mentioned that is not situated at the level of teaching and learning practice. Rather, it refers to a more strategic level, with teachers needing help in understanding the relation between formative assessment methods and summative as well as evaluative methods. Guidelines from educational policy representatives could help to clarify the relation between formative and summative assessment.

Widening of theoretical concept needed

Regarding the benefits of peer-assessment, the teachers did not only perceive learning effects but also social and motivational effects. This is not aligned with formative assessment theory which focusses on the former by conveying the idea that formative assessment supports student learning (Black & Wiliam, 1998; Natriello, 1987). Interdependencies between formative assessment and student motivation (Black & Wiliam, 1998) and a relation between formative assessment and student confidence (Smit, 2009) have been suggested, but literature on these effects is generally scarce. The result suggests that the formative assessment theory should be widened towards learning theories in which student motivation has been identified as a main contributor to student learning.

Retrospects and prospects

The aim of this study was to explore teachers' perceptions on benefits and challenges of peer-assessment in order to shape suitable means of support for teachers. The study was conducted with a small number of participants and in an open setting where the teachers designed the inquiry units themselves. It is therefore hard to decide on the specificity of the results (e.g. to what extent the challenges refer to peer-assessment specifically rather than to formative assessment methods in general). Nevertheless, the participating group of teachers included different school levels, subjects, years of teaching experience and gender. Furthermore, the rich data on the teachers' trials and their reflections upon them provide a dense picture of the teachers' perspectives on peer-assessment in the context of inquiry.

The study results in two main outcomes: Firstly, it offers first ideas on how to support the uptake of more peer-assessment in daily teaching practice. Secondly, it provides implications on how to further develop formative assessment theories.

ACKNOWLEDGEMENT

The work presented in this paper is part of the ASSIST-ME project which is funded by the European Commission (Seventh Framework Programme for Research; grant agreement no: 321428).

REFERENCES

- Andrade, H. (2010). Students as the definitive source of formative assessment. In H. Andrade & G.J. Cizek (Ed.), *Handbook of formative assessment*. New York: Routledge.
- 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.
- Black, P. (1993). Formative and summative assessments by teachers. *Studies in Science Education*, 21, 49-97.
- Black, P., Harrison, Ch., Lee, C., Marshall, B., & Wiliam, D. (2003). *Assessment for learning: Putting it into practice*. London: Open University Press.
- Black, P., Harrison, Ch., Lee, C., Marshall, B., & Wiliam, D. (2004). *Working inside the black box: Assessment for learning in the classroom*. Phi Delta Kappan.
- Black, P. & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: Principles, Policy and Practice*, 5(1), 7-73.
- Cole, D.A. (1991). Change in self-perceived competence as a function of peer and teacher evaluation. *Developmental Psychology*, 27, 682-688.
- Deutschschweizer Erziehungsdirektoren-Konferenz D-EDK (2014). *Lehrplan 21 [Curriculum 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.
- Furtak, E. M. & Ruiz-Primo, M. A. (2008). Making students' thinking explicit in writing and discussion. An analysis of formative assessment prompts. *Science Education*, 92(5), 799-824.
- Hanrahan, S. J. & Isaacs, G. (2001). Assessing self- and peer-assessment: The students' views. *Higher Education Research and Development*, 20, 53–70.
- Harlen, W. (2007). Holding up a mirror to classroom practice. *Primary Science Review*, 100, 29–31.
- Harlen, W. (2013). *Assessment & inquiry-based science education: Issues in policy and practice*. Trieste: Global Network of Science Academies (IAP) Science Education Programme (SEP).
- Heritage, M. (2010). *Formative assessment: Making it happen in the classroom*. Thousand Oaks, California: Corwin Press.
- Herman, J. L., Osmundson, E. & Silver, D. (2010). *Capturing quality in formative assessment practice: Measurement challenges, CRESST Report 770*. Los Angeles: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).
- Leahy, S., Lyon, Ch., Thompson, M., & Wiliam, D. (2005). Classroom assessment: Minute by minute, day by day. *Assessment to promote learning*, 63(3), 19-24.
- Lin, S. S. J., Liu, E. Z. F., & Yuan, S. M. (2001). Web-based peer assessment: Feedback for students with various thinking styles. *Journal of Computer Assisted Learning*, 17, 420-432.
- Lindsay, C. & Clarke, S. (2001). Enhancing primary science through self- and paired-assessment. *Primary Science Review*, 68, 15–18.
- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22, 155-175.
- Organization for Economic Co-operation and Development OECD (2005). *Formative assessment. Improving learning in secondary classrooms*. Paris, France: OECD Publishing.
- Organization 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. OECD Publishing, Paris.
- Ruiz-Primo, M. A., Furtak, E. M., Ayala, C., Yin, Y., & Shavelson, R. J. (2010). Formative assessment, motivation, and science learning. In H. Andrade & G. J. Cizek (Ed.), *Handbook of formative assessment* (pp. 139 – 158). New York: Routhledge.
- Sluijsmans, D. M. A. (2002). *Student involvement in assessment, the training of peer-assessment skills*. Maastricht: Interuniversity Centre for Educational Research.
- Smit, R. (2009). *Die formative Beurteilung und ihr Nutzen für die Entwicklung von Lernkompetenz. Eine empirische Studie in der Sekundarstufe I [Formative assessment and its use for the development of learning competence. An empirical study at lower secondary school level]*. Schneider Verlag Hohengehren GmbH: Baltmannsweiler.

- Topping, K. J. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*, 68, 249–276.
- Topping, K. J. (2003). Self and peer assessment in school and university: Reliability, validity and utility. In M. Segers, F. Dochy & E. Cascaller (Eds.), *Optimising new modes of assessment: in search of qualities and standards* (pp. 55–87). The Netherlands: Kluwer Academic Publishers.
- Topping, K. J. (2010). Peers as a source of formative assessment. In H. Andrade & G.J. Cizek (Eds.), *Handbook of formative assessment* (pp. 61- 74). New York: Routhledge.
- Topping, K., Smith, F. F., Swanson, I., & Elliot, A. (2000). Formative peer assessment of academic writing between postgraduate students. *Assessment and Evaluation in Higher Education*, 25, 149-169.
- Tsivitanidou, O., Zacharia, Z. C. & Hovardas, A. (2011). High school students' unmediated potential to assess peers: Unstructured and reciprocal peer assessment of web-portfolios in a science course. *Learning and Instruction*, 21, 506-519.
- Yin Y., Shavelson, R.J., Ayala, C.C., Araceli Ruiz-Primo, M., Brandon, P. R., Furtak, E. M., Tomita, M. K. & Young, D. B. (2008). On the impact of formative assessment on student motivation, achievement, and conceptual change. *Applied Measurement in Education*, 21(4), 335-359.