Report from the FP7 project:

Assess Inquiry in Science, Technology and Mathematics Education

ASSISTME

Assessment method description for 'Modeling' competence

Peer-assessment on students' artifacts

Olia Tsivitanidou, Peter Labudde

In collaboration with:

Christiana Nicolaou, Constantinos Constantinou

Delivery date	31th December 2014
Lead participant	University of Applied Sciences and Arts Northwestern Switzerland FHNW School of Teacher Education Center for Science and Technology Education
Contact person	olia.tsivitanidou@fhnw.ch
Dissemination level	PP

1. Summary

This descripiton will provide ideas and inspiration on how to formatively assess the 'modeling competence' using the peer-assessment method. There will be a description of what students and what the teacher are expected to do (their task) and how students' learning working process could be formatively assessed.

The "Peer-assessment" method certainly could be used in many fields of competence. Here the focus lies on the modeling competence in a paradigmatic example in Science.

Subject	• Modeling competence generally integrateable in all science subjects, in mathematics and technology education. • Paradigmatic example in Physics unit: moon phases, 5 lessons of 40 minutes over 2 weeks.
School level	• Modeling competence integrateable in lower and secondary education level • Paradigmatic example in lower secondary school level
Assessed competences in the paradigmatic example	In modeling "Model construction (Stratford, Krajcik, & Soloway, 1998); model use (NRC, 2012); comparison between models (Penner, Giles, Lehrer, & Schauble, 1997); model revision (Schwarz & White, 2005) and model validation have been identified as the practices in which students can be usefully engaged during modelling" (taken from ASSIST-ME report D4.7)
Data collection about student learning	• Students' artifacts: constructed model about the phenomenon of the moon phases
Feedback method	• Reciprocal peer-assessment, written comments
Combination with summative assessment	• Description, guidelines and paradigmatic example for formative assessment, assessment criteria also usable for summative assessment.

Table 1. Main characteristics of assessment method "Peer-assessment on students' model about the phenomenon of the moon phases".

2. Modeling competence

Modeling is the process of constructing and using scientific models (Hestenes, 1987) and it is considered an integral part of science (NRC, 2012). Efforts to design modelling-based learning (MBL) instruction have relied on a theoretical framework about the modelling competence, which analyses its constituent components into two broad categories, namely *modelling practices and meta-knowledge* (figure 1). Underlying this framework is the idea that student modelling competence can emerge as a result of active participation in specific modelling practices and can be reinforced by meta-knowledge about models and modelling (2009). Model construction (Stratford, Krajcik, & Soloway, 1998); model use (NRC, 2012); comparison between models (Penner, Giles, Lehrer, & Schauble, 1997); model revision (Schwarz & White, 2005) and model validation have been identified as the practices in which students can be usefully engaged during modelling. Meta-knowledge, on the other hand, is analysed into the metacognitive knowledge about the modelling process; this refers to student ability to explicitly describe and reflect on the actual process of modelling, but also on the knowledge about the nature and the purpose of models (Schwarz & White, 2005). In other words, this framework posits what scientists do during modelling and at the same time what we want students to do, so as to be modelling competent.

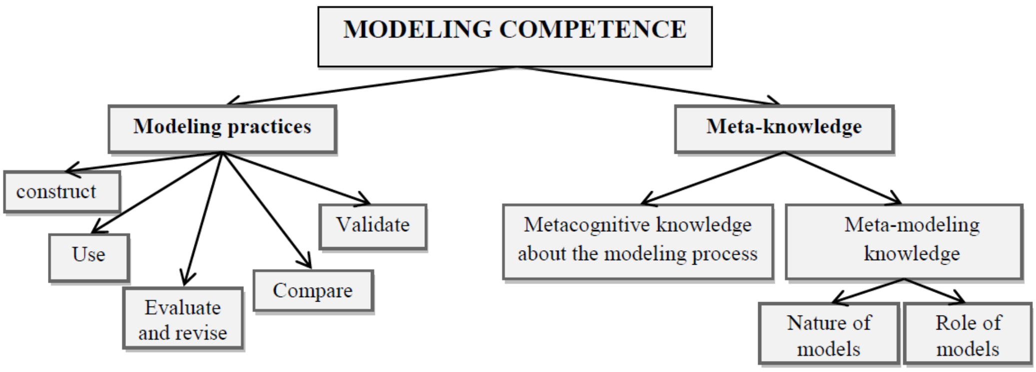


Figure 1: Modelling Competence Framework (Papaevripidou, Nicolaou, & Constantinou, 2014).

(Taken from ASSIST-ME report D4.7, p. 43)

3. Description of the assessment method with guidelines how to use it

A novelty approach in formative assessment context is the active involvement of students when assessing a peer's work, known as peer assessment. When students get engaged in peer-assessment activities, they produce peer-feedback, which could potentially assist peer assesses but also peer assessors in their learning process. Therefore, in this chapter, the use of peer-feedback as a means of formative assessment will be described. Peer assessment is conceptualized as an educational arrangement where students judge peers' performance by providing grades, and/or offering written or oral feedback (Topping, 1998). The method of peer-assessment presented here was inspired by the strategy "activating students as instructional resources for one another" (Leahy et al., 2005). Peer-assessment is seen as particularly powerful since "students may accept criticisms of their work from one another that they would not take seriously if the remarks were offered by a teacher. Peer work is also valuable because the interchange will be in language that students themselves naturally use […]" (Black et al., 2004, p. 14). Finally, the same authors find evidence that "when students do not understand an explanation, they are likely to interrupt a fellow student when they would not interrupt a teacher." (Black et al., 2004, p. 14).

Peer-assessment can be one-way or two-way (reciprocal). In one-way peer-assessment, students undertake either the role of the assessor or the assessee. On the other hand, in reciprocal (or two-way) peer-assessment, students undertake both the role of the assessor and the assessee, by assessing each other's work. The rationale lying behind reciprocal peer-assessment is that all students are given the opportunity to experience both the role of the assessor and the assessee and benefit from both practices. In order to implement reciprocal peer-assessment pairs of individual students or pairs of students' groups need to be formed. Then the pairs of students and/or groups share their work/ learning outcomes from the learning process. Initially in the peer-assessor role, the students are asked to assess their peers' work and to produce peer feedback. The peer feedback could either be of quantitative (e.g. grades) and/ or qualitative nature (e.g. oral or written comments which could include suggestions and recommendations for future action). The aim of the qualitative peer feedback is to assist peers in identifying the strengths and weakness or their work and in addition to provide suggestions for improving their learning process (Topping, 2003). Performing the peer-assessor role requires students to have and practice their assessment skills, namely: *defining criteria, judging the performance of a peer, and providing feedback* (Sluijsmans, 2002). Students could be supported through the provision of scaffolds while performing each one of these assessment skills.

Caveat! If students are novices in peer-assessing and have no prior experience on how to define assessing criteria or what has to be measured in the learning process and thus compose assessment criteria, they could alternatively be provided with those criteria from the teacher, in order to better execute their task. In that case, usually criteria are provided to peer assessors in the form of a rubric (see, for instance, Hafner & Hafner, 2003; Kocakülah, 2010). If the students have already acquired some expertise on peer-assessment, then they could define by themselves which elements of the task determine how success of the performance is measured (Topping, 2003) and therefore compose their own assessment criteria in regard to this. As far as the skill of *judging the performance of a peer* is concerned, students are responsible to critically analyze and judge a peer's performance, by applying the assessment criteria that have been given by the teacher (Topping, Smith, Swanson, & Elliot, 2000), or the assessment criteria that have been defined by themselves. With regard to the skill of *providing feedback*, peer assessors need to communicate their judgments to peer assessees and provide constructive feedback about their learning process. After having completed their task as peer-assessors, students change roles and become the assessees. Students receive the peer-feedback initially created by their peers. In the peer-assessee role, students are called to critically review the peer feedback received and decide on the actions to be taken. The skills required for enacting this role are different in nature from those of the peer-assessor. Peer-feedback might include flaws, since peer assessors are most probably novices in giving feedback. Therefore peer assessees need to filter the peer feedback and then decide whether there is a need to adopt peers' suggestions and recommendations and therefore whether there is a need to revise their work and/or making considerations in their future work.

Why you should do this

In reciprocal peer-assessment students could potentially benefit from experiencing both roles. Firstly in the peer-assessor role, the students practice and develop the aforementioned assessment skills (Hanrahan & Isaacs, 2001; Lin, Liu, & Yuan, 2001; Topping, 1998). Second, when writing feedback, students have more opportunities to engage in important cognitive activities, such as critical thinking (e.g., deciding what constitutes a good or poor piece of work), reflection etc. Third, students' informational resources expand by viewing and reviewing peers' work since they are given the opportunity to see examples of other students' work. This could potentially lead to experiencing implicitly self-assessment too, by comparing their own work and that of their peers', hence reflecting on their own learning achievements.

A number of benefits for learning could also be associated with the peer-assessee role (Tsivitanidou, Zacharia, & Hovardas, 2011; Hanrahan & Isaacs, 2001; Harlen, 2007; Lin et al., 2001; Lindsay & Clarke, 2001; Topping, 2003). Firstly students get the opportunity to receive additional feedback, compared to a more traditional setting where feedback usually comes from the teacher. Secondly, peer feedback might be more comprehensible to students since they share a common language/ coding. Thirdly, feedback derives from peers who have experienced the same learning process and possibly who have faced the same difficulties while performing the tasks of the learning sequence. As a result, the peer feedback could detect in a more direct way possible ways to overcome those difficulties and in a comprehensible language. Finally students while enacting the peer assessee role engage in important cognitive activities, such as critical thinking (e.g. while filtering peer feedback and deciding what constitutes a good or not peer feedback).

In this peer-assessment method, the students should exchange the models they have constructed based on data given by the teacher or a simulator or after having noted their own observations. The peer-assessment method endorsed in this example is *reciprocal* peer-assessment. This means that students both assess their peers' model and are being assessed by other peers. The students should try to evaluate their peers' model in regard to: (1) its functionality, whether the model explains the phenomenon of moon phases (2) its predictability, whether predictions could be done based on the model, (3) any other aspects that the students might consider as important to be med in a scientific model. After the completion of the model's evaluation, each peer-groups receives its feedback (could be more than one, please see below for more details) and is allowed to filter it and make revisions if there is such a necessity.

Finally, even though the assessment method chosen for this paradigmatic example is *peer-assessment*, it incorporates features of the feedback method *marking (grading and written comments),* since the peer-feedback described in this example is supposed to be delivered via written comment from a peer to a peer. As described in more details later on, the written peer feedback could emerge from an unsupported peer-assessment method or a supported one. In the first case, the students are free to give written feedback in any form/ way that they wish to. In the second case, the students are given a rubric (see subchapter *Assessment criteria for students' artifacts*) in which the peer assessors are required to assign a score as well as to give a qualitative comment supporting the assigned scoring.

"Rubrics articulate the expectations for an assignment or a learning goal by listing the relevant criteria the teacher looks for, or what counts that students should show to demonstrate various levels of performance." The peers in that case "indicate the student's scoring by placing a cross in the correct level of performance". "Written comments should identify what has been done well and what still needs improvement, and give guidance on how to make that improvement (Black et al., 2003). The same authors explain that simple 'good', 'well done', etc. is not sufficient since these general evaluations do not say what has been achieved nor what should be the next steps to be taken" (taken from ASSIST-ME report D4.7, pp 21).

4. Paradigmatic example: Physics, lower Secondary level

In this chapter, the use of a method for formatively peer-assessing students' modeling competence will be illustrated by an example. The assessment method used in this paradigmatic example is reciprocal peer-assessment. In reciprocal peer-assessment students enact both the roles of the assessor and the assessee (for more details see above the description of reciprocal peer-assessment). The purpose of students' artifact, that is a constructed model, should meet the criteria of representing, interpreting and predicting the phenomenon. In this paradigmatic example students should develop their modeling practices and in particular: model construction, use, comparison and evaluation. The two first components of the modeling practices (construction and use) should be accomplished within the home group work; the comparison and evaluation components should be practiced during the peer-assessment procedure, since their model about moon phases should be reviewed be peers; the revision component should take place within the initial home groups.

Students will initially work in their home groups (or pairs). Following the teaching material and collecting data they will be requested to construct their own model in order to explain the phenomenon. The example of using the method of peer-assessment is presented for a unit where the students are asked to construct their own model regarding the phenomenon of moon phases which will gradually lead them to the development of their model which will explain the aforementioned phenomenon and will be able to be used for future predictions. In this paradigmatic example the teaching material is derived from Physics by Inquiry (McDermott, & P.E.G., U. Wash., 1996). For the purposes of this example, the sub-chapter "Astronomy by Sight- Section 5" is used. The teaching approach embraced is guided inquiry; this means that the instructions provided in this curriculum guides students to gradually construct their own model. At the beginning of the teaching sequence, the students are asked to collect by themselves recordings of the moon phases in a piece of paper, thereby creating a "Moon Observation Summary Chart". For the purposes of this example, the data collection could be part of the teaching sequence; that is to say, students could be engaged in the procedure of collecting that data for a month. Due to the long period of time (one month) that is required for the data collection purposes, another option could be that the teacher provides students with data (data collected by students in previous years, or data that the teacher has already collected for a whole period of a month). Even in the latter case, a good idea could be to engage students in the data collection process (e.g. for two or three days) just to give them the opportunity to get familiar with the process and understand the provided data.

As the next step, the students are asked to carefully examine the data, try to identify patterns in the behavior of the moon and if possible group the data based on those patterns. This work should be done in the class within students' home groups and no lecture is taking place during the courses. On the contrary students follow the teaching material which is comprised by guidance questions and tasks and they (the students) are guided to reason through available data and carefully prepared worksheets. The teaching material incorporates “check points” with the teacher (or trained facilitators) whose duty is to assist the students finding their own path to understanding the under study phenomenon by guiding them with carefully chosen questions. In this specific paradigmatic example, the peer-assessment method is incorporated in the first “check point” where students are given the opportunity to exchange feedback between the various home groups (for more details see sub-chapter “Astronomy by Sight, Section 5. Phases of the moon, in *Physics by Inquiry*: McDermott, & P.E.G., U. Wash., 1996, p. 353).

In brief, the teaching material until the first check point, guides students through guided questions to: (1) carefully examine the provided data during the course of a single day/ the entire period of observations, and observe whether the appearance of the moon varies and how, (2) try to identify patterns in the way that the shape of the moon changes, (3) put in the right order several sketches of the moon’s appearance which are given, and label each phase based on a convection that is described in a given text (e.g. waxing gibbous moon, waning crescent moon etc.), (4) answer to a series of questions given, in regard to the sun position in the sky in various moon phases, supporting their answers with examples from the observations, (5) identify whether the angle between the sun and the moon is always the same when the moon has a particular phase, (6) try to illustrate the phenomenon using a light bulb and after having darken the room, along with a reasoning explaining the phenomenon so far, (6) discuss the following statement: “*The phases of the moon are caused by the shadow of the earth*”.

In other words students are provided with guided questions and texts, which will allow them to gradually construct their own scientific model in their home groups. Their final models (at the end of the whole unit) should be adequate to represent, interpret and make predictions about the moon phase’s phenomenon (for more details see *Physics by Inquiry* by McDermott, & P.E.G., U. Wash., 1996, p. 349).

In this example, after having completed all the aforementioned tasks, students are asked to construct their first model about the moon’s monthly motion and shape, which could be a written description or a drawing or even an illustration using a light bulb; in any case the relative objects, variables, quantities, processes and relations should be presented along with the “story” behind the model (its interpretive power).Therefore their task is to develop an external representation of this physical phenomenon which should be accompanied with an underlying mechanism which will rule the model. In this particular example the external representation should include the moon, the earth and the sun and the underlying mechanism of the moon phases’ phenomenon should refer to the changing angles (relative positions) of the earth, the moon and the sun, as the moon orbits the earth. The completion of a draft model by each group would determine the accomplishment of phase 1 in this peer-assessment method (see figure 2).

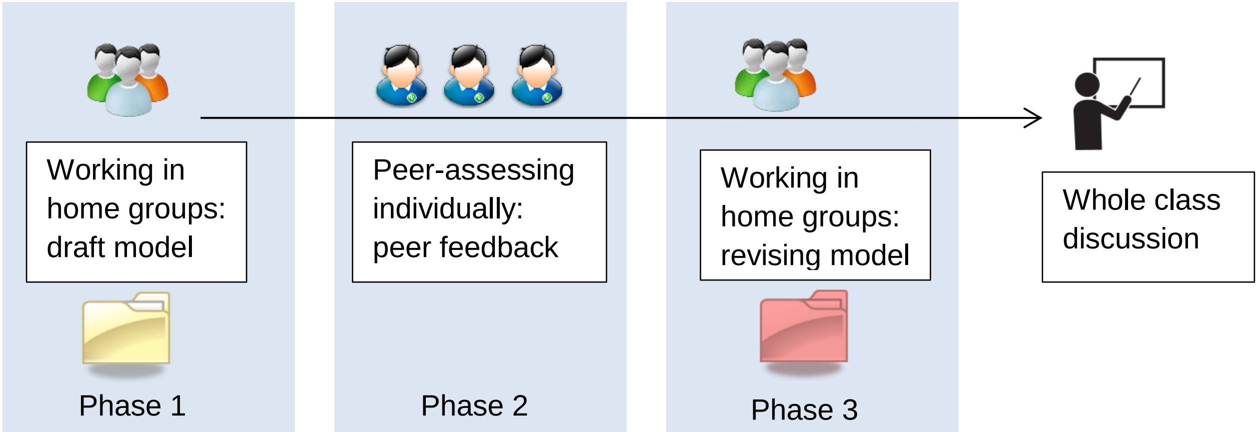


Figure 2. Reciprocal peer-assessment method

Once a draft model is completed, students of each group are asked to exchange their model with those of another group. In particular, one group of students gives its model to another group of students and vice versa. An alternative option would be to have the assesses demonstrating the model (e.g. in the case they use their own bodies to model the phenomenon, where the head represents the earth, a light bulb the sun and a peer holding a ball the moon) to the peer assessors, supplementing the demonstration with a written and/or oral rationale (see for example Figure 3).

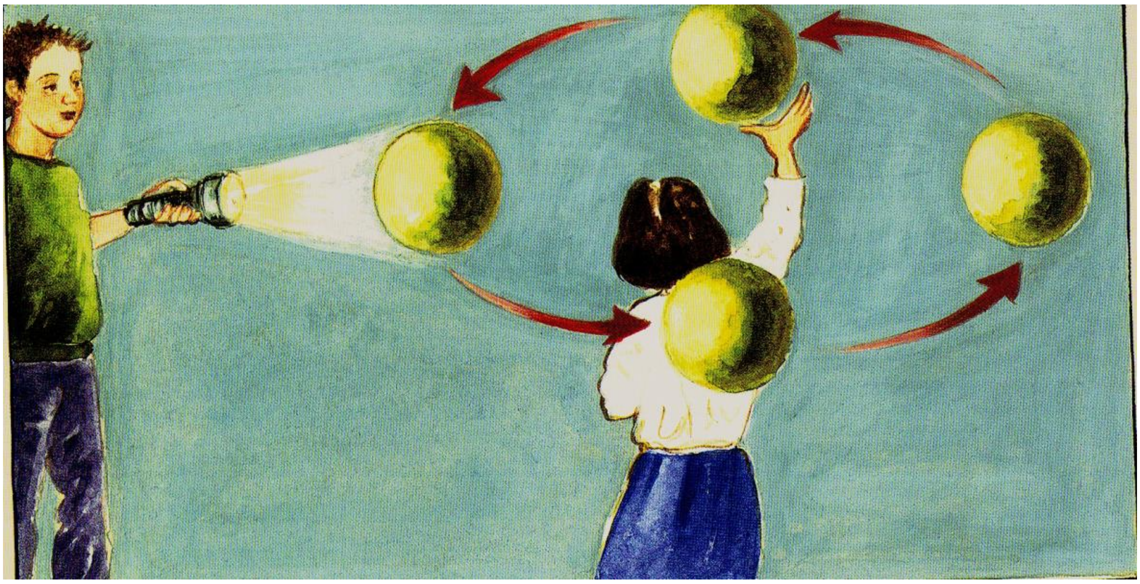


Figure 3: Demonstrating the model of the moon phases phenomenon (Halkia, 2006)

The pairs of group assessors- assessees could be predefined by the teacher or randomly. In this type or peer-assessment each group's model is assessing the model of another peer-group and is being assessed by the same peer-group. The assessment method could be supported or unsupported. In the first case, scaffolds could be given to students in order to implement the peer-assessment procedure (e.g. assessment criteria in the form of a rubric; see sub-chapter below); in the second case no scaffolds could be given to peer-assessors. In the latter case the students are free to assess their peer's model following their own rationale and giving emphasis on what they by themselves consider as important to be assessed. The decision on what type of peer-assessment to use (supported/ unsupported) could rest on the teacher. If the teacher feels that his/ hers students are novices in peer-assessing, then supported peer-assessment could be employed using the assessment criteria given in the next sub-chapter. The peer-assessment should be implemented individually. Therefore, if each group is comprised of two students, two feedback texts or filled in rubrics should emerge from each peer-assessor group. Consequently, each peer-assessee group should receive two different feedback texts or filled in rubrics from the corresponding assessor- group. This is phase 2 of this peer-assessment method (see figure 1).

After the peer-assessment completion, students within their home groups should receive the peer-feedback and collaboratively decide on the actions to be taken after a short discussion within the group. That is to say the students will have time to review peer-feedback and make revisions on their models if they consider it as significant to do in order to improve their models. Even if the peer-feedback's content might not be useful for a peer-assessee group, revisions in their models could be also emerge if the peer-assesses have noticed a better structured of the peer-model that they had assessed as peer-assessors. In either way students could benefit from this procedure. This is phase 3 of this peer-assessment method (see figure 1). At the end of the lesson, the teacher should lead a whole-class discussion where students could present their initial and revised models. During this discussion the students could explain how they experienced the peer-assessment method, what they liked about it and what not and if this assessment method assisted them in revising their models.

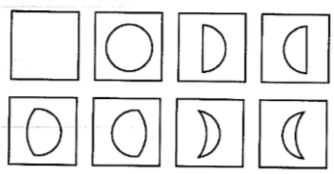
5. Assessment criteria for students’ artifacts

The following paragraph gives an overview of assessment criteria that could be used while peer-assessing a peer’s model. The assessment criteria provided in the rubric concern: the representational power of the model (inclusion of objects, variable quantities, processes, relations) (see criteria 1.1- 1.4), its interpretive power (the story/mechanism behind it) (see criteria 2-3) and its predictive power (see criteria 4-8). In regard to the criteria 1.1- 1.4, the necessary components that must be mentioned are: earth, sun, moon (objects), moon position in regard to earth and sun (variables), motion of the moon around the earth and around itself (processes), the sunlight hits the moon and the light that hits the moon reflects to the earth (interactions) etc.

Content-related criteria:

Table 1. Example of a rubric with pre-defined assessment criteria

	1: dissatisfied, 4: satisfied				Explain your reasoning
Assessment criteria	1	2	3	4	Comments
1. The model includes all necessary components: 1.1 Objects (earth, sun, moon)					
1.2 Variables (moon position in regard to earth and sun)					
1.3 Processes (motion of the moon around the earth and around itself)					
1.4. Interactions (the sunlight hits the moon and the light that hits the moon reflects to the earth)					

2. The model is ruled by an underlying mechanism which interprets the phenomenon.					State here which is the mechanism: ______________________ ____________
3. The underlying mechanism that rules the model is scientifically correct.					
4. Using the model we can define the sun's position, when there is a full moon.					
5. Using the model we can define the sun's position, when the moon is in the first quarter phase.					
6. Using the model we can define the sun's position, when the moon is in the third quarter phase.					
7. Using the model provided by our peers, we are in the position to order the following moon phases as they would appear during the course of a month. 					
8. The model can be used for prediction purposes. This means that the model could be used to predict how the moon shall appear in a specific day of the month.					
9. Criteria added by students.					

References

Black, P., Harrison, Ch., Lee, C., Marshall, B., and Wiliam, D. (2004): *Working inside the black box: assessment for learning in the classroom.* Phi Delta Kappan, Sept. 2004.

Hafner, J. C., & Hafner, P. M., (2003). Quantitative analysis of the rubric as an assessment tool: an empirical study of student peer-group rating. *International Journal of Science Education, 25,* 1509-1528.

Halkia K. (2006). *Το Ηλιακό Σύστημα μέσα στο Σύμπαν: Η διαδρομή από την επιστημονική γνώση στη Σχολική γνώση* (The Solar System within the Universe: The path from scientific knowledge to school knowledge). Heracleion, Crete: Crete University Press, (588 pages).

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.

Kocakülah, M. S. (2010). Development and application of a rubric for evaluating students' performance on Newton's laws of motion. *Journal of Science Education and Technology, 19,* 146-164.

Leahy, S., Lyon, Ch., Thompson, M., and Wiliam, D. (2005): *Classroom assessment: minute by minute, day by day.* Assessment to promote learning, Vol. 63 No. 3, pp.19-24.

Lindsay, C., & Clarke, S. (2001). Enhancing primary science through self– and paired–assessment. *Primary Science Review, 68*, 15–18.

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.

McDermott, L. C., & The Physics Education Group. (1996). *Physics by Inquiry*. New York: J. Wiley.

Papaevripidou, M., Nicolaou, C. T., & Constantinou, C. P. (2014). On Defining and Assessing Learners' Modelling Competence in science Teaching and Learning. Paper presented at the Annual Meeting of American Educational Research Association (AERA), Philadelphia, Pennsylvania, USA.

Penner, D. E., Giles, N. D., Lehrer, R., & Schauble, L. (1997). Building Functional Models: Designing an Elbow. *Journal of Research in Science Teaching, 34*(2), 125-143.

Schwarz, C., & White, B. (2005). Metamodeling Knowledge: Developing Students' Understanding of Scientific Modeling. *Cognition and Instruction, 23*(2), 165–205.

Sluijsmans, D. M. A. (2002). *Student involvement in assessment, the training of peer-assessment skills.* Interuniversity Centre for Educational Research.

Stratford, S. J., Krajcik, J., & Soloway, E. (1998). Secondary Students' Dynamic Modeling Processes: Analyzing, Reasoning about, Synthesizing, and Testing Models of Stream Ecosystems. *Journal of Science Education and Technology 7*(3), 215 - 234.

Topping, K. (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., 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.