Using rubrics to improve the assessment lifecycle: a case study

Marius Miknis , Ross Davies & Clare S. Johnson

Published online: 16 Sep 2020.

Submit your article to this journal ↗

Article views: 31

View related articles ↗

View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

ARTICLE

Using rubrics to improve the assessment lifecycle: a case study

Marius Miknis, Ross Davies ⓘD and Clare S. Johnson ⓘD

Department of Computing and Maths, University of South Wales, Pontypridd, UK

ABSTRACT

This paper explores whether students' learning outcomes can be improved through the use of self-assessment rubrics. Students on a computer programming module in a Higher Education Institution were required to complete a self-assessment using the same rubric as the assessors. Observing discrepancies between the grades the students were receiving, and the grades the students *thought* they should be receiving, the lecturers made improvements to the pedagogical approaches taken for some elements of the course by changing the format and focus of classroom activities. This resulted in both improved grades and improved self-regulation by students. The process was facilitated through a system created by the authors of the paper called SAFE (Self-Assessment Feedback and Evaluation Learner Lifecycle), which greatly enhances the learner feedback lifecycle of an assignment. The research corroborates existing studies around the importance of revisiting feedback both for assessor and student.

Introduction

The QAA ([2018](#)) defines two types of assessment: formative assessment, which has a developmental purpose, using feedback to help make learners learn more effectively; and summative assessment, which is used to 'indicate the extent of a learner's success' in meeting given assessment criteria (QAA, [2018](#)). When feedback is only provided as a summative grade it can have a negative effect on learning, and comments are only effective if they are used to 'close the learning gap' (Welsh Assembly Government, [2010](#)). In the traditional assessment cycle, an assignment is given, graded and returned and feedback is rarely revisited. Indeed, Nicol and MacFarlane-Dick ([2006](#)) suggest that feedback is very often a process of transmission: 'Teachers "transmit" feedback messages to students about what is right and wrong in their academic work, about its strengths and weaknesses, and students use this information to make subsequent improvements'. How often students actually act upon this feedback is questionable, especially when the feedback is summative. What is less questionable is the value gained by becoming self-regulating, and how the self-regulatory process of self-moderation and self-observation can be used to mentally map strategic processes and their outcomes, as discussed by

---

CONTACT Clare S. Johnson ✉ clare.johnson@southwales.ac.uk 🖳 Department of Computing and Maths, University of South Wales, Pontypridd CF37 1DL, UK

Zimmerman and Schunk (2004). In addition, it is often the case that there is a mismatch between what students understand is required of them, and staff expectation. In her article 'Analytical rubrics in higher education: A repository of empirical data' Hack (2015) suggests that there is a need for 'greater transparency and clarity in the assessment process, for students and markers', and comments on the improved consistency in marking when rubrics are used, but notes that evidence is not so clear on whether the use of rubrics is effective in assessing specified learning outcomes.

Cockett and Jackson (2018), in their article on the use of assessment rubrics provide a comprehensive literature review, and make several observations: that rubrics can be useful in standardising feedback and ensuring consistency across markers and student submissions; that rubrics may enhance students' higher-order thinking skills, but more interestingly appear to improve learner satisfaction; that rubrics should not be seen as a panacea for all student concerns relating to feedback ('quality, usefulness and consistency'); and finally that rubrics are most effective when they have been co-created with students themselves (students as partners).

Falchikov and Boud (1989) discuss self-assessment and describe how more experienced learners, and learners who have expertise in a subject, tend to be more accurate in self-assessment process and lead to greater student-teacher agreement in terms of outcomes. Self-regulated learning refers to learning that is guided by metacognition, strategic action and motivation – including taking control of one's learning and self-evaluation. Self-regulated learners are very aware of their academic strengths and weaknesses and have a good understanding of effort and use of strategies to improve success. Panadero, Jonsson, and Strijbos (2016) describe self-regulation as an assessment for learning practice which increases student involvement in the assessment process, particularly through the use of rubrics. Falchikov and Boud (1989) also note that student-teacher agreement is not, in itself, the most valuable part of the process, but that greater self-awareness of performance is, even if the students' grades don't match those of the assessor.

The aim of this research was to establish whether the use of rubrics for self-assessment improves learner outcomes, with the following research questions:

- To what extent do rubrics assist in improving learner outcomes?
- What benefits to the student do self-assessment rubrics offer?
- What benefits to the assessor do self-assessment rubrics offer in terms of content delivery?

Methodology

This research takes a case study approach over a 3-year period. A module on a level 4 undergraduate (first year) computing course with an average of 42 students was used to review the assessment lifecycle, both from the perspective of the assessor and the student to establish mechanisms for improving learner outcomes. The module was delivered by two full time lecturers who taught alternating lecturers and tutorials. Over the 3-year period, a grading rubric was implemented and supplemented with a self-assessment rubric mirroring the grading rubric. The module in question contained seven assessment points, with a slightly different rubric used at each point. Three cohorts of students

studying this module over the duration of the study were required to complete self-assessments prior to reviewing the results given to them by the assessor, and were then able to compare their own grading of their work with the grading received from the assessor. In the latter stages of the study, the assessor also reviewed the rubrics to establish areas of weakness; areas where the students perceived they were either strong or weak but that did not correlate with the assessment grades; and finally to review how the students develop their self-assessment and metacognitive skills throughout the process. The use of a rubric was chosen due to its versatility in different types of assessments and it lends itself well to being digitised so that the same system could be used by assessors and students and implemented on the SAFE platform. In addition to completing the rubric, which provided quantitative data, students were also asked to rate several aspects of their studies (quality of lectures, tutorials and lecturers) and finally were invited to add comments, which resulted in some qualitative data. For the purposes of this paper, only the quantitative data provided by the rubrics will be discussed.

Assessment lifecycle

In its simplest form, the typical assessment lifecycle from the lecturer's perspective consists of two parts; creation and assessment (see Figure 1). During creation the assessment is developed and validated, whilst during assessment the student's work is graded and feedback produced. This approach is not inherently reflective as once the mark/feedback is given back to student there is no obligation on the student to reflect on how well the assessment has gone.

This research proposes that in addition to the criteria for the rubric, development and creation of standardised and consistent feedback comments should also take place when the assessment is created. Customised personal feedback should only be required at the marking stage. With feedback development in mind at the creation stage the assessment itself can be improved as the assessor considers what students need to do to achieve a specific level of feedback (Figure 2).

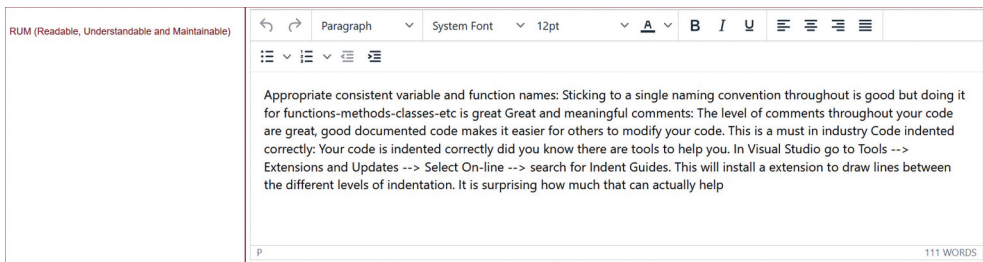**Figure 1.** The simplified lecturer assessment lifecycle.

Figure 2. Example of standardised feedback comment that can be selected during the grading process, created in SAFE.

This use of rubrics provides a structure for the assessor to ensure more consistent marking across the cohort and also an easier way for a student to reflect on their work. This is corroborated by Jonsson and Svingby (2007) who comment that rubrics assist in the reliable scoring of performance assessments.

Self-assessment is useful not only for students to self-reflect on their work but also for the assessors to better understand their students. This also has a substantial benefit for the students in directing them to fully analyse and comprehend the marking scheme. Over time this can lead to a significant increase in students' understanding as well as the lecturer's understanding of the students. As students become part of the learner lifecycle their self-assessment will not only improve their understanding of a concept but have a much clearer picture of what the assessor is looking for – a feature of rubrics identified by Jonsson and Svingby (2007, p. 141) who note that the transparency of a rubric facilitates greater student learning because 'rubrics make expectations and criteria explicit'.

Following marking, the assessor can review the results, identify areas that were misunderstood or lacked clarity, and add content to supplement those gaps for revision purposes or to improve the teaching for future cohorts. It is useful to show students any improvements made in this way as it demonstrates that they are part of the development process. This changes the lecturer assessment lifecycle (see Figure 3).

Method

For the initial trial a C++ programming module was chosen which has its assessments split into tutorial exercises. For the first 2 years of the study, the rubric was used by the assessors, and the students were required to complete the self-assessment at each assessment point, prior to receiving their grade. The assessors did not review the self-assessment reports for the first 2 years, and did not adjust their teaching methods. Over the module term, students undertake seven assessments which are allocated a variety of maximum marks. The distribution is as follows: Aptitude Exercise: 2%, Tutorial Exercise 1.1: 4%, Tutorial Exercise 1.2: 4%, Tutorial Exercise 2.1:10%, Tutorial Exercise 2.2:10%, Assessment 2:30% and Assessment 3:40%. These assessments focused on the base concepts of C++ which included: arrays, vectors, functions, classes and inheritance. The benefit of this split assessment approach is the ability to better track students' progress through the year as more samples can be taken instead of relying one
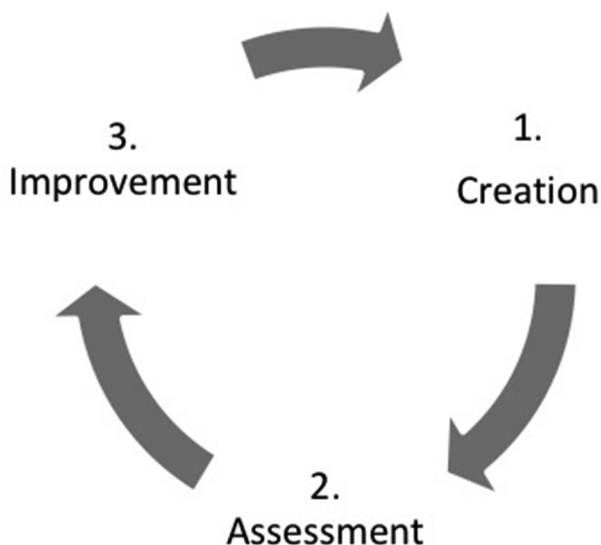
Figure 3. The improved lecturer assessment lifecycle.

or two large assessments placed late during the year making it hard to pick up any gaps in students' understanding. The first five assessments were small and were completed by students in a two-hour tutorial slot as they only focused on a single programming concept, while the last two larger assessments were done over the time period of a month as they were assessing multiple concepts. The rubric for the Tutorial Exercise 1.1 can be seen in Figure 4. Full ethical approval was obtained for publication of the findings.

Results

The observation was run over a period of 3 years and was able to show a disparity between students' perception of a given concept (e.g. the use of constant variables) to

	Fail (0/29)	Narrow Fail (30/39)	3rd Class / Pass (40/49)	Lower 2nd Class / Pass (50/59)	Upper 2nd Class / Merit (60/69)	1st Class / Distinction (70/100)
Pseudocode (15%)	☐ Very poor or no pseudocode	☐ Poor use of pseudocode design	☐ Pseudocode is satisfactory	☐ Pseudocode is good	☐ Very good use pseudocode design	☐ Excellent use pseudocode design
RUM (Readable, Understandable and Maintainable) (10%)	☐ Variable names lack useful meaning ☐ No comments ☐ Code not indented correctly ☐ No constant variables used	☐ Variable names lack useful meaning ☐ Invalid badly worded comments	☐ Some variable names lack useful meaning ☐ Lack of comments ☐ Some parts of the code need better indentation ☐ Poor use of constant variables	☐ Variable names are mostly appropriate ☐ Some good comments	☐ Most variable names are appropriate ☐ Great comments	☐ Appropriate consistent variable and function names ☐ Great comments ☐ Code indented correctly ☐ Good use of constant variables
Assignment and arithmetic operations (25%)	☐ Very poor assignment operations ☐ Very poor arithmetic operations	☐ Poor assignment operations ☐ Poor arithmetic operations	☐ Assignment and arithmetic operations is satisfactory	☐ Assignment and arithmetic operations is good	☐ Very good assignment and arithmetic operations	☐ Excellent assignment and arithmetic operations
IF statements (20%)	☐ No IF statements used	☐ Poor use of IF statements	☐ IF statements are satisfactory	☐ IF statements are good	☐ Very good IF statements	☐ Excellent IF statements
Input validation (10%)	☐ No input validation implemented	☐ Poor input validation	☐ Input validation is satisfactory	☐ Input validation is good	☐ Very good input validation	☐ Excellent input validation
Testing (20%)	☐ Very poor testing	☐ Poor testing	☐ Testing is satisfactory	☐ Testing is good	☐ Very good testing	☐ Excellent testing
Global:						

Figure 4. The rubric of tutorial exercise 1.1 used to assess students and for them to self-assess themselves.

Table 1. Distribution of classifications between students' self-assessment grades and the assessor grades over the 3-year study.

	1st year		2nd year		3rd year	
Classification	Students Grading	Assessor Grading	Students Grading	Assessor Grading	Students Grading	Assessor Grading
First	52.17%	6.38%	88.89%	12.50%	87.50%	31.82%
Pass	13.04%	12.77%	5.56%	17.50%	6.25%	31.82%
Fail	34.78%	80.85%	5.56%	70.00%	6.25%	36.36%

what they actually understood about it. The data summary can be seen in Table 1. Each given concept was marked as either First class, Pass or Fail, as the individual concepts were not significantly complex enough to warrant more detailed grading (e.g. Upper Second or Lower Second) as would be the norm for the institution concerned and as such feature in the standardised rubric template above.

Grading for Tutorial Exercise 1.1 (4%) showed a significant disparity in a programming concept (the use of constant variables) between the students and the assessor. In this exercise, it showed that most students (52%) thought they did well and 13% of the students thought they adequately while 34% thought they had failed to implement the concept. This however was different to what the assessor has marked as only 6% of the students did well, 12% adequately and 80% of the student had failed to implement the concept. This shows that many students didn't truly understand the concept but thought they did, the difference can be seen in Figure 5.

The teaching content was not changed for the following year to see if this would repeat itself with new cohort of student and the same exercise was given out to the new students to perform. The results were similar to that of the previous year where the majority (88%) of students thought they did well. However, the assessor had failed the majority of them (see Figure 6).

For the third year the teaching content was adjusted to have more detail of the programming concept that students had misunderstood in the previous 2 years, after
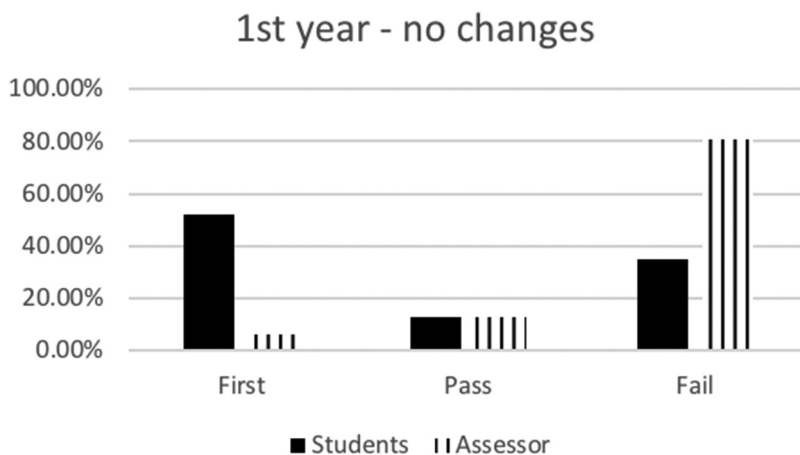

Figure 5. Distribution of grades between 3 criteria (First, Pass and Fail) for the programming concept during the first year of running the exercise.
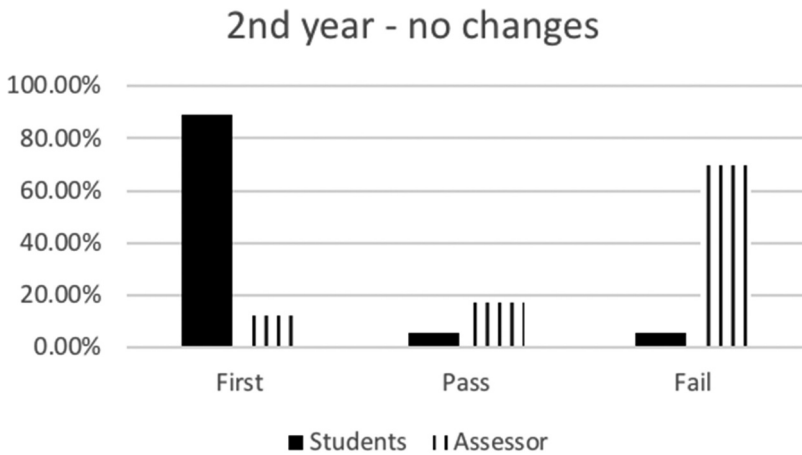
2nd year - no changes

Figure 6. Distribution of grades between 3 criteria (First, Pass and Fail) for the programming concept during the second year of running the exercise.

which the exercise was run again. The changes made included the additional information with practical examples and included an in-class discussion for students to participate in a chat about the said concept. Results demonstrate how the additional materials and discussion have improved students' understanding of the concept. Students' perceptions stayed very similar to the previous year, but the assessor's marks have shifted significantly. Instead of more than 75% of students on average failing to implement the concept this time around only 36% failed to do so. The number of students doing either 'well' or 'adequately' has doubled, which is a significant change (see Figure 7). When looking at other assessments that students have done it was seen that they had a far better grasp of the concept than in previous years. This demonstrates the effectiveness of the assessor being aware of the discrepancies between what the student perceives they are good at: in
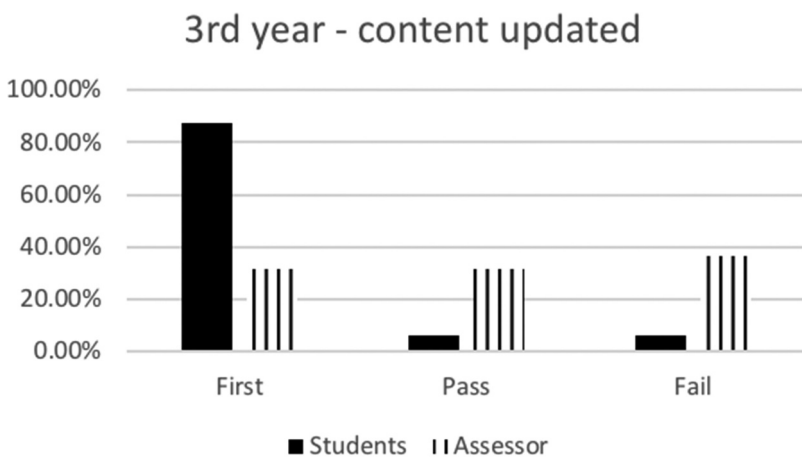
3rd year - content updated

Figure 7. Distribution of grades between 3 criteria (First, Pass and Fail) for the programming concept during the third year of running the exercise after changes have been made to the teaching content based on previous years self-assessment.
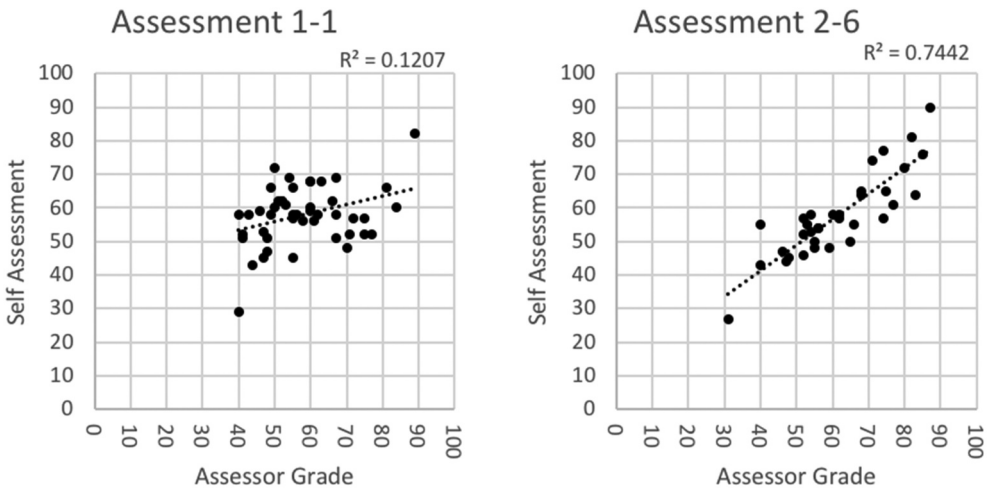
Figure 8. Self-assessment analysis comparing the assessors' grades vs the students-self assessment.

short, the student thinks they are good at something, although their grades indicate otherwise, highlighting the assessor's need to alter the pedagogical approach.

This further enforces that students can become part of the learner lifecycle as through self-assessment they can not only improve their understanding but also that of the assessor. In this way the educator and learner should not be seen differently as the educator can also be learner.

Importance of self-assessment

During the observation period, it was also discovered that students' self-assessment became much more accurate with each of the exercises completed. A Spearmen's RHO analysis of the classification boundaries shows a significant improvement in the student's ability to self-assess accurately. This shows students understanding of the assessment has increased which helps both the student and the assessor. The student's accuracy by the 6th assessment was within 5 marks of the assessor. There has been a significant improvement from the first assessment to the 6th assessment (see Figure 8).

Initial data has also shown this trend has carried on through all the 3 years of the degree but more research will be required to further analyse this.

Conclusion

The findings from this project demonstrate that both the assessor and the learner can benefit from reflective practice and self-assessment. Over 2 years, the assessors reviewed the self-assessments of the students, and was able to observe discrepancies between how the work was graded, and the students' perceptions of how well they did. In the final year of this stage of the project, the assessor changed the content for the elements that the students were struggling with, and student outcomes improved as a result. Furthermore, over the course of the 3-year study, it is clear that students become much more proficient at understanding their own learning, and recognising their strengths and weaknesses, as

their self-assessment gets much closer to the actual grading received from the assessor. This is a skill which is fully transferrable and highly valuable within the workplace.

Returning to the first research question posed during this study 'To what extent do rubrics assist in improving learner outcomes?' it is clear that rubrics do have a role to play for several reasons: firstly, that the students can use these to become more self-regulating, and better able to understand their own strengths and weaknesses through the process of self-assessment, as indicated by the increase in Spearman correlation between the assessor's marks and the students' marks by the end of the study period, (from 0.1207 in Assessment 1–1 to 0.7442 in Assessment 2–6), and secondly, because the assessors can use this self-assessment to revise their teaching methods and scaffold the learning process more successfully.

The second research question asked 'What benefits to the students do self-assessment rubrics offer?'. Aside from the obvious advantages offered from improve self-regulation discussed above, the transparency of the marking criteria offered by a rubric assists the student in being better able to understand what the assessors are looking for in each task. This could be developed further in a future study by asking students to develop the rubric themselves, thus giving them greater ownership of the marking criteria – an idea discussed by Falchikov and Boud (1989), which would be an area worthy of further development in the SAFE system.

Finally, in terms of 'What benefits to the assessor do self-assessment rubrics offer in terms of content delivery?' the study shows that rubrics support a more consistent approach to marking, and the self-assessment element provides an insight into the learners' perceptions of their strengths and weaknesses that assists the assessor in adjusting delivery and content to facilitate improved learning outcomes.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Ross Davies 🔟 http://orcid.org/0000-0002-5460-9837
Clare S. Johnson 🔟 http://orcid.org/0000-0001-5869-957X

References

Cockett, A., & Jackson, C. (2018). The use of assessment rubrics to enhance feedback in higher education: An integrative literature review. *Nurse Education Today, 69* (October 2017), (pp. 8–13). doi:10.1016/j.nedt.2018.06.022

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, *59*(4), 395–430. doi:10.3102/00346543059004395

Hack, C. (2015). Analytical rubrics in higher education: A repository of empirical data. *British Journal of Educational Technology*, *46*(5), 924–927. doi:10.1111/bjet.12304

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review*, *2*(2), 130–144. doi:10.1016/j.edurev.2007.05.002

Nicol, D., & MacFarlane-Dick, D. (2006). Formative assessment and selfregulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. doi:10.1080/03075070600572090

Panadero, E., Jonsson, A., & Strijbos, J.W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. In D. Laveault & L. Allal (Eds.), Assessment for Learning: Meeting the challenge of implementation. The Enabling Power of Assessment, *4*. Springer, Cham. doi:10.1007/978-3-319-39211-0_18.

QAA. (2018). *UK quality code for higher education general introduction*. (October), (pp. 1–50). Retrieved from http://www.qaa.ac.uk/en/Publications/Documents/quality-code-brief-guide.pdf

Welsh Assembly Government. (2010). *How to develop thinking and assessment for learning in the classroom guidance*. Retrieved from https://www.gwegogledd.cymru/wp-content/uploads/2018/04/How_to_develop_thinking_and.pdf

Zimmerman, B.J., & Schunk, D.H. (2004). Self-regulating intellectual processes and outcomes: A social cognitive perspective. In D.Y. Dai & R.J. Sternberg (Eds.), *Motivation, emotion, and cognition integrative perspectives on intellectual functioning and development* (pp. 323–349). New York: Routledge. Retrieved from https://zodml.org/sites/default/files/%5BDavid_Yun_Dai,_Robert_J._Sternberg%5D_Motivation,_E_0.pdf#page=338Lon