

# When students assess each other: How effective is peer assessment for learning in the classroom?

One of the core tasks of a teacher is to provide feedback to his or her students on their learning performance - in the form of grades, or oral and written feedback. In the peer assessment method, this task is taken on by the students themselves, by giving each other feedback on their performance. Double, McGrane, and Hopfenbeck (2020) examine whether students can benefit from peer assessment in their meta-analysis, »The impact of peer assessment on academic performance: A meta-analysis of control group studies«. The authors also explore the question of how peer assessment can be effectively implemented in the classroom.

**INTRODUCTION.** Peer assessment appears in the literature under different terms, for example, *peer feedback*, *peer grading* or *peer evaluation*. In the meta-analysis by Double and colleagues, peer assessment is understood as a method in which members of a peer group (e.g. students of the same school level) give each other feedback on their performance. This includes oral, written, or digital exchange of task-relevant feedback.

In the classroom, peer assessment can be implemented in a variety of ways. For example, teachers can provide students with assessment criteria in the form of rubrics or grading systems, they can enable the digital exchange of feedback, and they can establish peer assessment as a regular feedback practice in their classrooms.

Previous research on peer assessment indicates a strong relationship between students' peer assessment ratings and teacher ratings of the same students (e.g. Li et al, 2016; Sanchez et al, 2017). This meta-analysis examines the effectiveness of peer assessment on students' academic performance based on experimental research.

### META-ANALYSIS AT A GLANCE

<b>Focus of the study</b>	Effectiveness of peer assessment on achievement compared to other forms of assessment (e.g., by teacher)
<b>Target group</b>	Primary and secondary school students and university students
<b>Average effect size</b>	Positive overall effect of peer assessment on achievement ( $g = 0.31$ )
<b>Further findings</b>	Effect of peer assessment is stable across different implementation approaches and age groups

**WHAT IS THIS STUDY ABOUT?** In their meta-analysis, Double and colleagues investigate how the use of peer assessment affects the performance of students (primary, secondary level, and tertiary level) in comparison to instruction without assessment, with self-assessment, or with teacher assessment.

In total, the meta-analysis includes 55 individual (quasi-)experimental studies with a total of 143 effect sizes. These studies comprised articles published in international journals as well as *gray literature* (e.g., dissertations, conference papers) from between 1966 and 2018. In the studies included, peer assessment was integrated into instruction in a wide variety of ways (see *Table 1*), and each was compared to a passive control group (i.e., no assessment) or an active control group (i.e., other form of assessment, e.g., by the teacher).

Student performance was recorded either in terms of written performance (e.g., test or essays) or practical performance (e.g., application tasks). The authors used moderator analyses to investigate the conditions under which peer assessment was particularly effective in the classroom. In doing so, they not only examined the influence of different classroom implementation options, but also in which subject peer assessment was used.

*Table 1: Ways of implementing peer assessment in the classroom.*

<b>Role:</b>	Students can take on different roles within the peer assessment: as the assessor, as the one to be assessed, or both.
<b>Type:</b>	The most common forms of peer assessment are grading, oral exchange, and written assessment. All forms can be combined in different ways and can also be implemented online.
<b>Support:</b>	Possible forms of support are, for example, the specification of a rubric, an evaluation scheme, or a grading system. Peer assessments can also be designed freely and without support.
<b>Anonymity:</b>	Peer assessments can take place openly or anonymously. Thus, it can be made transparent or not as to which student is assessing whom, or who is being assessed by whom. In addition, the type of assignment can be varied: the roles can be assigned openly or secretly, or role assignment can be freely chosen by the students.
<b>Frequency:</b>	The use of peer assessments can vary from one-time use to regular classroom practice. It can be assumed that with increasing experience and practice on the part of the students, the quality of the feedback can also improve.

**WHAT DID THIS STUDY FIND?** Based on the 55 individual studies, the authors found a significant positive overall effect of peer assessment on achievement:  $g = 0.31$  (confidence interval  $g = 0.18$  to  $g = 0.44$ ). When secondary school students were considered separately, there was also a significant positive effect of peer assessment on academic achievement:  $g = 0.44$  (13 studies).

Looking at the control groups considered, it appears that students with peer assessment performed significantly better in learning than students in control groups without assessment ( $g = 0.31$ ). This effect was also evident when compared to control groups with teacher assessment ( $g = 0.28$ ), but not when compared to control groups with self-assessments.

In the moderator analyses, the authors found that the positive effect of peer assessment is robust across different implementation options. That is, for nearly all<sup>1</sup> forms of peer assessment examined (type, frequency, etc.), a positive effect on student achievement could be found. An overview of all moderators can be found in the [overview of all individual findings](#).

**HOW DOES THE CLEARING HOUSE UNTERRICHT EVALUATE THIS STUDY?** The *Clearing House Unterricht Research Group* evaluates the meta-analysis using the following five questions, guided by the Abelson criteria (1995):

**How substantial are the effects?** Overall, the meta-analysis finds a significant overall positive effect of  $g = 0.31$ . The size of this effect indicates that more than 60% of students who used peer assessment achieved better learning performance than the average student in control conditions without assessment, or with other forms of assessment. Results from moderator analyses indicate that different implementation options for peer assessment do not account for statistically significant differences in this effect. Thus, overall, the findings suggest that peer assessment has a positive effect on student achievement regardless of the particular implementation.

**How differentiated are the results?** The authors consider different levels of education (primary, secondary, and tertiary) and school subjects in their analyses. However, there are no significant differences between the educational levels and school subjects with regard to the effectiveness of peer assessment. Both K-12 students and university students seem to benefit from peer assessment in different subject matter and with different implementation options. Although the performance criterion studied is differentiated into written and practical tasks in the study selection, it is not examined separately in the analyses.

**How generalizable are the findings?** The results of the meta-analysis show positive effects on performance for almost all peer assessment implementation approaches. No statistically significant differences were found for different subjects or educational levels. In principle, therefore, it can be assumed that the positive effect of peer assessment can be regarded as generalizable.

It should be noted, however, that geographic location was not taken into account as a moderator, which is why no statement can be made about the generalizability across different countries. Further experimental studies are also needed to make a differentiated statement about the effectiveness of various options for peer assessment implementation, especially in secondary education and among the various mathematics and science subjects.

**What makes this meta-analysis scientifically relevant?** Double and colleagues' meta-analysis is scientifically significant because it systematically and reliably estimates the effec-

---

<sup>1</sup>Only peer feedback with a grade showed a difference: While peer feedback with a grade had a significant positive effect on university students' performance ( $g = 0.55$ ), this effect could not be shown for primary or secondary school students ( $g = 0.002$ , n.s.).

tiveness of peer assessment for learning success in addition to existing qualitative reviews and literature (Dochy et al., 1999; Topping, 1998).

In addition, the meta-analysis reveals two important points for future research: first, it highlights the lack of experimental studies in which effects of assessing and being assessed on performance are considered in a differentiated way. Attention toward these respective roles in peer assessment is important in order to be able to make more specific statements about the mechanisms that promote learning. Moreover, the meta-analysis shows that in the studies considered, students were predominantly assigned as whole classes in experimental and control conditions, i.e., quasi-experiments. For the future, more rigorous study designs with randomized assignment at the individual level (experimental study design) would be needed.

**How methodologically reliable are the findings?** The transparency and justification of the methodological approach largely meets the standards criteria of common requirement guides (e.g., APA Meta-Analysis Reporting Standards). In particular, the study search and selection are documented in an exemplary manner. In the area of coding and analysis of primary studies, more detailed information would be required for better comprehensibility - e.g., details on the samples, designs, and survey instruments of the studies considered. More detailed information on the methodological assessment can be found in the [rating sheet](#).

**CONCLUSION FOR CLASSROOM PRACTICE.** The findings of the present meta-analysis provide evidence that peer assessment can be a supplement to instruction that enhances learning. The positive effect of peer assessment on achievement is stable across different ages, subjects, and other various conditions. This suggests that peer assessment can be effectively implemented in the classroom in a variety of ways.

For students, peer assessment provides opportunities to critically engage with learning content and reflect on their own performance. Based on the individual studies considered, however, it can also be noted that the research on peer assessment to date is still quite unspecific. For example, there is a lack of studies that include the different roles in peer assessment or the concrete contexts in order to be able to make evidence-based recommendations for more specific implementation. Further research efforts are therefore needed to better understand the effectiveness of different ways that peer assessment can be implemented and mechanisms that promote learning.

## EXAMPLE STUDY

Wang and colleagues' (2017) study examines the effectiveness of peer assessment in computer science classes in 9th grade. During the ten-week study, the students were to learn the programming language »Scratch«. The sample consisted of four school classes with a total of 166 students. All four classes received an introduction to »Scratch« and the task of programming their own project.

In two classes, the students were then supposed to give each other feedback on their individual projects (experimental group with 80 students, peer assessment). They were given clear evaluation criteria (e.g. on topic, content and design) and evaluation levels (from 1 to 4). In the other two classes, the students received feedback from the teacher (control group with 86 students).

To check the effectiveness of peer assessment compared to teacher assessment, all students took part in a test on programming knowledge and skills as well as questionnaires on critical thinking and their attitudes – once before starting the course (pre-test to check the entrance requirements) and once at the end of the course (post-test). In addition, the students were given programming tasks at both times to test their practical programming skills. As a result, the experimental group achieved significantly better results than the control group for all areas of the posttest.

## REFERENCES.

- Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Ahn, S., Ames, A. J., & Myers, N. D. (2012). A review of meta-analyses in education: Methodological strengths and weaknesses. *Review of Educational Research, 82*(4), 436–476.
- 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.
- Double, K. S., McGrane, J. A., & Hopfenbeck, T. N. (2020). The impact of peer assessment on academic performance: A meta-analysis of control group studies. *Educational Psychology Review, 32*(2), 481–509.
- Li, H., Xiong, Y., Zang, X., L. Kornhaber, M., Lyu, Y., Chung, K. S., & K. Suen, H. (2016). Peer assessment in the digital age: A meta-analysis comparing peer and teacher ratings. *Assessment & Evaluation in Higher Education, 41*(2), 245–264.
- Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H., & Cooper, H. (2017). Self-grading and peer-grading for formative and summative assessments in 3rd through 12th grade classrooms: A meta-analysis. *Journal of Educational Psychology, 109*(8), 1049.
- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research, 68*(3), 249–276.
- Wang, X.-M., Hwang, G.-J., Liang, Z.-Y., & Wang, H.-Y. (2017). Enhancing students' computer programming performances, critical thinking awareness and attitudes towards programming: An online peer-assessment attempt. *Journal of Educational Technology & Society, 20*(4), 58–68.

## LINKS.

To the meta-analysis from [Double et al, 2020](#)

To the study example from [Wang et al, 2017](#)

## CITE AS.

Farrell, M., Schönbeck, T. & CHU Research Group (2023). When students assess each other: How effective is peer assessment for learning in the classroom?  
[www.clearinghouse-unterricht.de](http://www.clearinghouse-unterricht.de), Short review 28.

## ADAPTED FROM.

Diery, A., Knogler, M. & CHU Research Group (2022). Wenn SchülerInnen sich gegenseitig beurteilen: Wie lernwirksam ist Peer Assessment im Unterricht?  
[www.clearinghouse-unterricht.de](http://www.clearinghouse-unterricht.de), Short review 28.



This short review is licensed under a [Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International](https://creativecommons.org/licenses/by-nc-nd/4.0/). Use and distribution permitted with attribution, no modifications permitted