Exploring students' perceptions of self-assessment in the context of problem solving in STEM

Joni Lämsä[1,2], Anne Virtanen[3], Päivi Tynjälä[3], Jussi Maunuksela[4] and Pekka Koskinen[4]

[1] Learning and Educational Technology Research Lab, University of Oulu, Finland
[2] Department of Education, University of Jyväskylä, Finland
[3] Finnish Institute for Educational Research, University of Jyväskylä, Finland
[4] Department of Physics, University of Jyväskylä, Finland

Central goals of higher education in STEM domains include learning of problem solving and self-assessment skills. To achieve these goals, we propose a novel self-assessment method called the Solve-Correct-Assess-Negotiate (SCAN) method of assessing problem solving that includes both formative and summative elements. We study students' learning experiences in courses involving different methods of assessing problem solving (Course 1 including teacher-led assessment, N_{trad} = 53; Course 2 including SCAN method, N_{SCAN} = 56) and specifically associations between these learning experiences and students' perceptions of the SCAN method. We found that the students relied on teacher-led assessment more than the self-assessment. The perceived utility of the self-assessment was positively associated with a deeper approach to learning in Course 2 than in Course 1. Students who found the self-assessment less beneficial also perceived less support from the learning environment. Our findings suggest that the successful implementation of novel self-assessment-based practices for problem-solving requires personalized support for self-assessment, teachers' awareness of the different perceptions that students have towards these practices, and discussion among teachers and students on the rationale, utility, and reliability of the different practices.

Keywords: assessment, higher education, problem solving, self-assessment, STEM

ARTICLE DETAILS

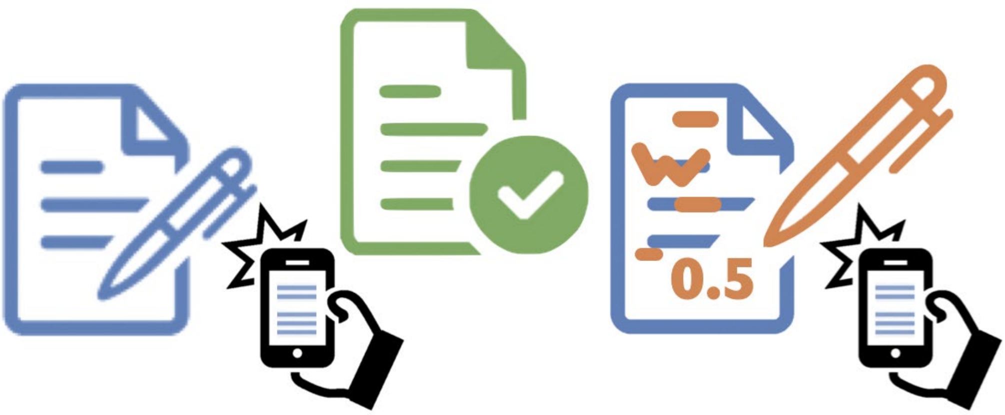
LUMAT General Issue
Vol 11 No 2 (2023), 35–59

Received 22 May 2023
Accepted 7 September 2023
Published 15 September 2023

Pages: 25
References: 52

Correspondence:
joni.lamsa@oulu.fi

https://doi.org/10.31129/LUMAT.11.2.2028



1 Introduction

The ability to solve problems is one of science, technology, engineering, and mathematics (STEM) professionals' main skills (de Jong, 2019). When the problems have well-defined solutions, which happens frequently in STEM domains, self-assessment methods are useful for teaching students problem-solving procedures, their critical analysis, and self-assessment skills (Panadero et al., 2017; Randles et al., 2018; Yan & Carless, 2021). While students' self-assessment has been recommended as a formative tool with teacher still assigning the grades (Bourke, 2018; Yan & Carless, 2021), summative self-assessment can further foster studying and learning (Häsä et al., 2021). Here, to facilitate the use of innovative and sustainable self-assessment practices in problem solving, we aim to increase understanding of students' perceptions of the self-assessment when it includes both formative and summative elements (Andrade, 2019).

Since assessment methods cannot be separated from other components of learning environments, in what follows, we first elaborate on the role of assessment methods in learning environments (Section 1.1). We then illustrate how assessment methods may be associated with learning outcomes, approaches to learning and support provided by the learning environment (Section 1.2). Based on this literature review, we propose a novel self-assessment-based method of assessing problem solving that includes both formative and summative elements (Section 2).

1.1 Assessment methods as part of learning environments

Biggs and Tang (2011) presented the concept of constructive alignment, a principle according to which teaching and assessment methods should align with the intended learning outcomes. For example, if students are expected to learn problem solving, their studying should involve conceptual knowledge and practical exercises on solving problems, and their assessment tasks should not only measure the development of the corresponding skills and knowledge, but they should also support learning.

Assessment methods are often divided into summative and formative categories (Ashenafi, 2017). In the context of problem solving, summative assessment methods have been based on grading of the solutions, which often emphasises the correctness of the final answer. Such grading may lead to a weak understanding of problem-solving procedures (van Merrienboer et al., 2003) instead of developing deeper understanding or supporting lifelong learning skills (Struyven et al., 2006). While summative assessment focuses on learning outcomes, formative assessment supports learning processes (Ashenafi, 2017; Boud & Falchikov, 2006). Formative assessment provides students feedback to help them recognise strengths and weaknesses and regulate learning and progress (Ashenafi, 2017; Broadbent et al., 2018). Furthermore, it helps teachers to adapt to students' progress (Ashenafi, 2017). Formative assessment can foster students' reflection on the problem-solving process and thereby improve their related skills (Virtanen & Tynjälä, 2019).

Assessment methods differ also in terms of the assessing person. While teacher-led assessment methods rely on teachers' grading, various forms of student self-assessment have also been developed (e.g. Andrade & Du, 2007). Boud and Falchikov (2006) proposed that students must become assessors to meet the needs of their future work. Self-assessment is particularly suitable in problem solving because it supports critical comparison between a student's procedure and a 'model solution'; that is, students' active role may help them critically analyse their actions and decisions (Panadero et al., 2017). Professionals in STEM fields must assess the correctness of their calculations independently. Self-assessment may not only develop self-assessment and reflective skills of students (Virtanen & Tynjälä, 2019) but also improve student engagement (Kearney, 2013), motivation (Andrade & Du, 2007) and the quality of achievement (Andrade & Valtcheva, 2009).

Self-assessment has mainly been recommended as a formative tool (Bourke, 2018). Indeed, students may find self-assessment practices appropriate for formative rather than summative purposes (i.e., assigning grades) if they find their own grading less reliable than teachers' grading (Boud & Falchikov, 1989). However, self-assessment can also be used for summative purposes and assigning grades (Taras, 2016). In that case, negotiating about assessment with peers and teachers may help students to discern the central actions in the problem-solving process, externalise their misconceptions and understand the alignment between assessment, learning goals and learning and teaching practices (Biggs & Tang, 2011; Yan et al., 2023; Yan & Carless, 2021). Summative assessment and grading can be supported by clear rubrics, which highlight the importance of the problem solving process and scientific reasoning during different actions (Docktor et al., 2016). A recent study by Nieminen et al. (2021) showed that summative self-assessment may be related to a high level of deep approach and an increased self-efficacy. Häsä et al. (2021) concluded that summative self-assessment can be used to support students' studying when self-assessment is aligned with future-driven pedagogical purposes, as suggested also by Boud and Falchikov (2006).

Assessment is an integral element of broader pedagogical practices and learning environments and inseparable from its wider context. Consequently, assessment must be investigated in its context, including student approaches to learning, their perceptions of support provided by the learning environment and intended learning outcomes.

1.2 Role of the learning environment in students' learning processes and outcomes

Early studies on student approaches to learning by Marton and Säljö (1979) identified two approaches, the surface and the deep approach. Typical of the deep approach is a student's aim to establish a profound understanding of given topic, whereas typical of the surface approach is rote learning. Further studies identified a third approach, the strategic approach or organised studying (e.g. Entwistle & McCune, 2004). In this approach, students aim for high achievement and organise their study habits to attain high grades. Research has shown that although some students tend to adopt a constant approach, other students vary their approaches depending on the pedagogical practices (e.g. Postareff et al., 2015). Research conducted among higher education students has also shown that study success correlates positively with deep approach and negatively with surface approach (Lindblom-Ylänne et al., 2019). However, such correlations are absent in some studies (Fryer & Vermunt, 2018), suggesting that the surface approach may produce high grades when assessment favours mere rote learning.

Recent studies on student learning in higher education have focused on wide relationships between students' learning approaches and perceptions of their learning environment—how students see that teaching and assessment support their learning (McCune & Entwistle, 2011). Several measures have been developed to examine students' perceptions of their academic context, such as the Experiences of Teaching and Learning Questionnaire (ETLQ, Entwistle et al., 2003). In addition to the scales of approaches to learning, the ETLQ and its modified versions (e.g. Parpala et al., 2013; Utriainen et al., 2018) include several learning environment-related scales, such as teaching for understanding, constructive feedback and support from other students. Studies using these scales suggest that students' views of their learning environment have an indirect effect on their study achievement through their approaches to learning. For instance, students who perceive teaching as supportive, interesting and conducive to understanding are highly likely to adopt a deep approach and achieve good learning outcomes (Utriainen et al., 2018; Vermunt & Donche, 2017).

Learning outcomes in higher education can be divided into generic and domain-specific outcomes (Tremblay et al., 2012). While the focus in assessment has traditionally been on domain-specific knowledge, recent research has extended attention to generic skills, such as problem solving. Studies suggest that the use of diversified assessment methods, active forms of learning, and linking theory and practice are important for the development of generic skills (e.g. Tynjälä et al., 2016; Virtanen & Tynjälä, 2019).

In the following section, we present the teacher-led assessment method and propose a novel self-assessment method of problem solving grounded on the presented literature review.

2 The teacher-led and SCAN methods of assessing problem solving

The teacher-led methods of assessing problem solving in STEM contexts have typically followed a similar structure. Students start by solving the problems, often handwriting the solutions on paper. They send the solutions to the teacher by a given deadline, after which the teacher assesses and grades the solutions. The number of assessed solutions is often so large that the teachers can afford to add only scant feedback notes or none at all. Assessment is followed by a meeting, where the solutions—frequently only the correct model solutions—are discussed in small groups, typically using a narrative dominated by the teacher. In another popular method, students mark the problems they solved on a piece of paper, followed by teacher randomly selecting students to present their solutions on a blackboard in front of the class; often the points for marked problems are obtained regardless of the success in presenting the solution. Sometimes the solutions are discussed in small groups instead of the whole class (Koskinen, 2012). Nevertheless, common to these methods is that assessment is rarely discussed, and students accept grading mostly without any questions. No systematic mechanism exists to encourage students to reflect on the assessments or on the mistakes, misconceptions, and chosen actions.

To address the above shortcomings, we have proposed a novel, student self-assessment-based method to assess problem solving (Koskinen & Lämsä, 2019). We called this method the Solve–Correct–Assess–Negotiate (SCAN) method. In the method, students solve problems, then correct and assess their own solutions, and discuss and negotiate the assessment with the teacher and peers. The SCAN method aims to make the problem-solving process visible to the students by shifting the focus from the correctness of the final solutions to the critical evaluation of the actions conducted when reaching the solution (Panadero et al., 2016). In physics, for example, the central actions include physical modelling, such as applying principles and laws; mathematical calculation and related procedures; and critical analysis of the solution and its rationality (Docktor et al., 2016; Wancham & Tangdhanakanond, 2022). The SCAN method is visualized in Figure 1.

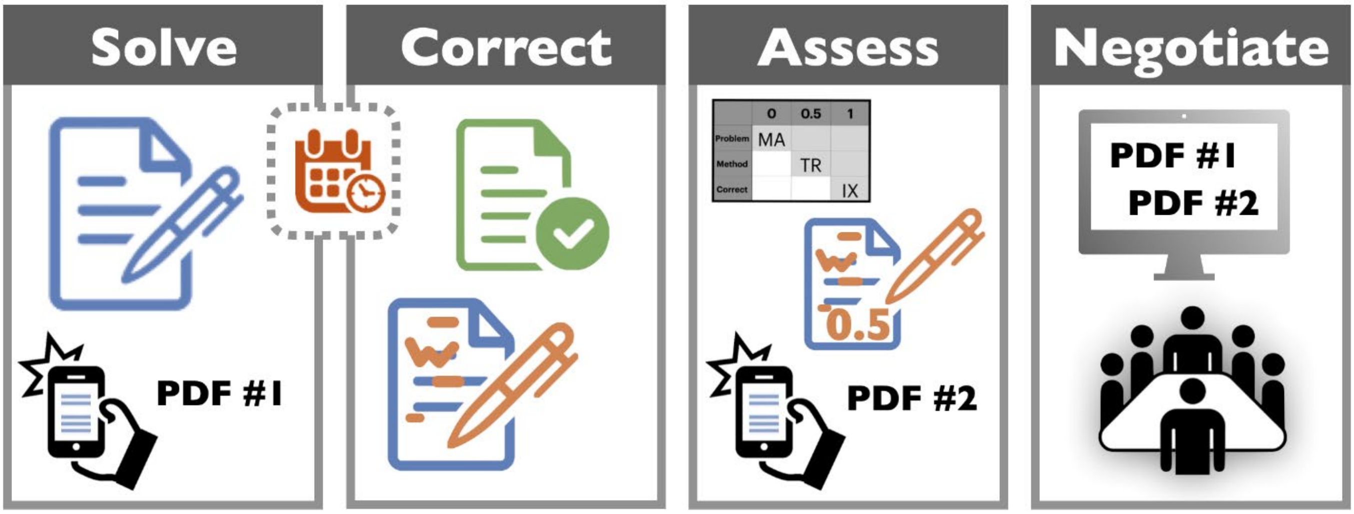


Figure 1. Four stages of the Solve-Correct-Assess-Negotiate (SCAN) method to assess problem solving.

1. Solve: Students solve problems, write solutions on paper, scan the papers using smartphones, and send the PDF files to a digital learning environment. Scanning a few pages can be done in less than a minute by smartphones' ubiquitous scanning apps. Students may also use tablets, Latex, or some other software to produce PDF files directly.
2. Correct: The work is submitted by a given deadline, after which the digital learning environment automatically publishes the model solutions. Students use the model solutions to identify and correct mistakes in their solutions.
3. Assess: With the help of a customisable rubric (Docktor et al., 2016) that considers the central actions of the problem solving process (Figure 2), students assess their original and corrected solutions to determine their grades. Depending on the specific STEM field in which the SCAN method is applied, it may be necessary to rephrase the descriptions of the assessment criteria to account for

the unique characteristics of each field. All notes, marks, corrections and grades are appended visibly on top of the original solutions. Students scan the papers again, now with all the notations and grades, and send the PDF file to the digital learning environment.

4. Negotiate: The teacher now has two sets of PDF files—the first set reveals what the students accomplished independently, and the second set shows students' attempts at grading. The teacher and a small group of students (3–6) then have discussions on the problem-solving process, outcomes, and assessment. Because the correct solutions are public, the time students share with the teacher can be focused on discussing assessment and matters that remain unclear or challenging.

Student's original solution					
Assessing:	**0 points**	**0.25 points**	**0.5 points**	**0.75 points**	**1 point**
Assignment, modeling	Problem assignment merely repeated, assignment misunderstood or not understood at all.	Assignment understood correctly, modeling began, solution advanced significantly beyond initial assumptions.			
Idea, solution, principle		Central idea, method or solution principle still unidentifiable. Modeling incomplete or central idea or assumption wrong.	Central idea or solution principle identified and written explicitly. (Idea can be different from the one in model solutions.)		
Method, procedure, calculations			Principle is not put into action or the procedure is fundamentally incomplete.	Central idea or solution principle identified and written explicitly. The procedure is also carried out completely.	
Solution, answer, inspection				Solution includes mistake in math or logic that goes beyond inaccuracy or carelessness. The result is qualitatively wrong or physically unrealistic.	The solution is complete, the answer (nearly) correct. May include mistakes, but the answer is sensible qualitatively correct. *Or:* correct answer is obtained by different (acceptable) route.

Student's corrected solution					
Assessing:	**+0.25 points**	**+0.25 points**	**+0.25 points**	**+0.25 points**	**0 points**
Identifying mistakes, correcting and amending own solution	Identifying the reasons for why the assignment was not understood and what exactly was learned from this. Merely copying the model solutions is not enough: points are not given without *subjective* reflection and *a proper original attempt for solution*.	Identifying the reasons for why the idea or procedure remained incomplete; modeling and assumptions are corrected and amended. Points require *subjective* reflection; they are not given by mere copying of model solutions.	Identifying why implementing the idea or procedure remained incomplete; the procedure is made complete with the help of model solutions. Points require *subjective* reflection; they are not given by mere copying of model solutions.	Identifying the mistake and the reasons that lead to it. Reflecting the implications of the mistake. The solution is amended with the help of model solutions. Points require *subjective* reflection; they are not given by mere copying of model solutions.	No additional points. Stand by for more challenging problem assignments.

Figure 2. Generic assessment rubric consisting of two parts. The first part relates to the student's original solution. It is structured using the central actions of the problem-solving process to enhance the learning of these actions. The second part relates to the student's efforts to identify and correct their own mistakes.

Thus, in addition to being an acronym, SCAN refers to students' task of scanning their solutions, which makes the method easy to implement. Although the time taken by technology is minimal, its role is central in enabling the method. The clear procedures and regular deadlines of the SCAN method may foster organised studying, the lack of which is frequently associated with the surface approach to learning (Lindblom-Ylänne et al., 2019). The instant feedback on one's own solutions (as opposed to delayed feedback) and time to process the feedback before and during the negotiating stage may also foster learning (Candel et al., 2020; Yan et al., 2021; Yan & Carless, 2021). The SCAN method could particularly benefit STEM students in higher education, who have achieved higher scores on surface approaches to learning than students in other disciplines (Parpala et al., 2010).

On the one hand, students themselves make use the SCAN method is understood as a method that includes elements of both summative and formative assessment. On the other hand, students themselves make use of summative assessment (Struyven et al., 2006) when they give themselves a grade for every problem they solve that also contributes to their final grade. On the other hand, their self-assessments are reviewed during the course through evaluation criteria and discussions. In this case, formative assessment (Ashenafi, 2017) is used, which refers to assessment and feedback during the learning process. In the self-assessment literature, many studies have found congruence between student and teacher's summative assessments (Carroll, 2020; Jax et al., 2019). Still, students may perceive teacher's grading more appropriate than self-grading, and self-assessment rarely contributes to the student's final grade directly (Häsä et al., 2021). Thus, an improved understanding of students' perceptions of the assessment in their learning experiences may foster the successful and sustainable implementation of self-assessment-based practices for problem solving in the future (Andrade, 2019), particularly when the self-assessment includes both formative and summative elements.

3 Research aims

Here, we focus on two assessment methods of problem solving in a STEM context—the teacher-led method based on teacher assessment and the SCAN method based on student self-assessment. First, we study the learning experiences in courses involving different methods of assessing problem solving and answer the following research questions (RQs):

RQ1a: How do students perceive their learning outcomes in courses based on (i) teacher-led and (ii) SCAN methods?

RQ1b: How do students perceive the support for assessing problem solving provided by the learning environment in courses based on (i) teacher-led and (ii) SCAN methods?

Second, awareness of perceptions towards self-assessment-based methods may help increase the acceptability and transparency of these methods. Thus, we ask:

RQ2a: How do students' perceptions of the SCAN method differ?

RQ2b: How do these different perceptions relate to the students' comparison of the two courses, learning outcomes, and perceived support provided by the learning environment?

4 Materials and methods

4.1 Participants and data

The participants of this study took introductory physics courses on electricity (N_{tr} = 53, Course 1: teacher-led method of assessing problem solving) and electromagnetism (N_{SCAN} = 56, Course 2: SCAN method of assessing problem solving) at a Finnish university. Half of the participants were physics majors (48% in Course 1, 52% in Course 2), while most of the minor students came from chemistry, biology, mathematics, or computer science. One-fifth of the participants were in the subject teacher track (22% in Course 1, 19% in Course 2). The courses lasted for 6–9 weeks and followed the primetime learning model, with a weekly set of problem assignments (Koskinen et al., 2018). Courses did not have exams, and the assessment of problem-solving solutions constituted 45% of the final grade. The remaining portion of the grade is determined through weekly self-study activities and small group meetings prior to problem-solving, in addition to criteria-based self-assessment, group assessment, and teacher assessment at the end of the course (see details in Koskinen et al., 2018).

Although Courses 1 and 2 had different contents, they were highly similar. Both courses discussed basic electromagnetism, and the discussion was split into two courses for the sheer purpose of practical arrangements (the courses could easily be merged into a single coherent course). The courses had the same level of (mathematical) difficulty and were arranged back-to-back during one spring term; they were taught by the same teachers and completed virtually by the same group of first-year undergraduate students, and most importantly, their problem-solving and assessment procedures were identical. Therefore, for the purposes of this study, the courses are analogous and can be compared.

To address RQ1, students answered questionnaires online in Webropol (https://webropol.com) at the end of the course. The questionnaires examined perceptions of their own learning outcomes using a battery of nine items regarding relevant generic and domain-specific skills (e.g. Virtanen & Tynjälä, 2019). Questionnaires also probed the perceived support provided by the learning environment. The starting point for the questionnaires was the short version (40 items) of the ETLQ, modified for the Finnish context (Parpala et al., 2013; Utriainen et al., 2018). We only used scales related to students' perceptions of their learning environment, and we adapted some items to fit the context of our study. For example, the original item 'The feedback given on my set work helped clarify things I had not fully understood' was adapted to 'Reviewing the problem-solving solutions helped clarify things I had not fully understood'. The students were asked to respond to each item on a 5-point scale (1 = strongly disagree, 5 = strongly agree). These questionnaires were answered after each course, 34 students answering both questionnaires (Figure 3).

To address RQ2, we used the same set of items as in RQ1 but only focused on Course 2 ($N_{\mathrm{SCAN}} = 56$, Figure 3). In this course, the questionnaire also included 12 items that concerned students' perceptions of the SCAN method (Table 1, see also Section 4.2). Moreover, the questionnaire included four items that compared the courses using teacher-led and SCAN methods, including a modified question from the ETLQ to measure deep approach to learning. In addition to the questionnaire items, we used the aggregated gradings of the weekly problem-solving solutions as the measure of the students' problem-solving ability. Teachers verified these gradings and lowered the grade of the solution if students had not followed the provided assessment rubric or corrected their solutions (Figure 2). For the subsequent analysis, we standardised the problem-solving ability to scales 1–5. Figure 3 shows a summary of our data.

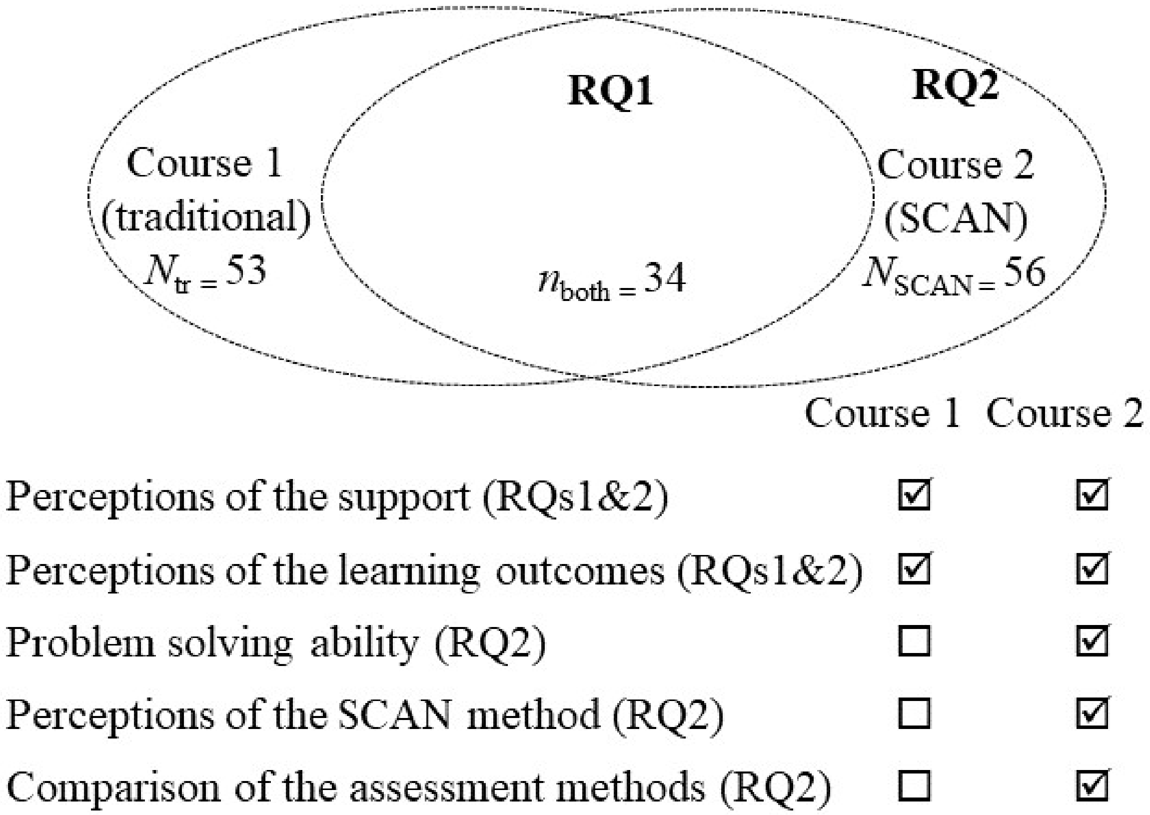


Figure 3. The participants and data relating to the different research questions (RQs).

Table 1. Items that were used to measure students' perceptions of the Solve-Correct-Assess-Negotiate (SCAN) method. The loadings, based on the exploratory factor analysis of the two-factor structure, are presented for absolute values greater than 0.4. We labelled factor 1 as 'SCAN as a beneficial assessment method' and factor 2 as 'SCAN as a reliable assessment method'.

SCAN item	Factor 1	Factor 2
SCAN was easy to do.		0.62
SCAN deepened my learning.	0.75	
I found SCAN oppressive.		−0.48
I found SCAN-based assessment unreliable.		−0.71
SCAN-based assessment was as reliable as teacher assessment.		0.77
SCAN became easier with practice.	0.79	
SCAN inspired me to further analyse my learning and development.	0.95	
The assessment criteria were clear.		0.64
SCAN helped me correct my mistakes.	0.68	
SCAN developed my ability to critically analyse my own work.	0.91	
The time that I used on SCAN was not wasted.	0.69	
I am used to assessing my own learning.	Excluded from further analyses	

4.2 Analysis

It was clear to me what was expected of the students in the problem solving and in the process of reviewing the problem-solving solutions going through them. To examine students' perceptions of the support they received during their courses (Figure 3), we created three aggregate scales of the ETLQ items based on validated instruments (Parpala et al., 2013; Utriainen et al., 2018) and statistical (i.e., Cronbach's alpha) methods (see Appendix). Since we adapted some of the original items to fit the context of our study (Section 4.1), we modified the labels of the aggregated variables. The labels were as follows: 1) support from other students ($\alpha = 0.87$ in Course 1, $\alpha = 0.79$ in Course 2, four items); 2) constructive feedback related to problem solving ($\alpha = 0.74$, $\alpha = 0.85$, two items; original label: 'constructive feedback'); and 3) teaching for understanding ($\alpha = 0.79$, $\alpha = 0.68$, three items). In addition, we used two single variables, 4) enjoyment of studying ("I liked studying in this course.") and 5) understanding the demands of problem solving (original label: 'alignment'; "It was clear to me what was expected of the students in the problem solving and in the process of reviewing the problem-solving solutions.").

To answer RQ2, we first focused on 12 items that concerned students' perceptions of the SCAN method (Figure 3 and Table 1). When we analysed the items' correlation structure, one item ('I am used to assessing my own learning') correlated weakly with the other 11 items ($|r| < 0.24$) and was removed from further analyses. From the correlation structure, we also observed the multidimensionality (correlation within subsets of items was high); thus, we aimed to discern the boundaries of the subsets by performing an exploratory factor analysis (EFA) based on the 11 items. The overall Kaiser–Meyer–Olkin index (0.88) indicated that the correlation matrix was adequate for the EFA: The index was 0.73 for one item and above 0.84 for the other ten items. We then calculated the eigenvalues of the correlation matrix. Two eigenvalues were above 1, and they saturated rapidly when the number of eigenvalues was two or more. Thus, we decided to perform the EFA with a two-factor structure. We did not require orthogonal factors, and we used promax rotation. The root mean square error of approximation (RMSEA) fit statistic of the two-factor structure indicated good fit (RMSEA = 0.03, 90% confidence interval [0.00, 0.11]) (McNeish & Wolf, 2020).

Based on the loadings of the items with absolute values higher than 0.4, we identified 1) factor 1 as describing students' perceptions of the SCAN as a beneficial assessment method (six items with loadings 0.68–0.95, for example, 'SCAN developed my ability to critically analyse my own work') and 2) factor 2 as describing students' perceptions of the SCAN as a reliable assessment method (five items with loadings 0.48–0.77, for example, 'SCAN-based assessment was as reliable as teacher assessment') (see Table 1). These two factors explained 57.1% of the total variance of the variables.

Due to the disparate loadings within each factor, the factor scores of factors 1 and 2 were calculated for each student instead of unit-weighting sum scores (McNeish & Wolf, 2020). The mean and standard deviation of the factor scores were 0 and 1, respectively. Since we aimed to use these factor scores in subsequent analyses to study their relations to the students' comparison of the two courses that had different methods of assessing problem solving, learning outcomes and perceived support provided by the learning environment, we used the regression as a scoring method (McNeish & Wolf, 2020). We then used the scores of factors 1 and 2 as independent variables and students' comparison of the courses, learning outcomes and perceptions of the support as dependent variables; and we employed linear regression analyses. We excluded the interaction between the factor scores of the two factors from our models, since those had no effect on the dependent variables. We took the type I error into consideration due to multiple comparisons and considered the findings with $p < 0.001$ as statistically significant.

5 Results

5.1 Support from learning environment and learning outcomes in courses based on teacher-led and SCAN methods

The students' perceptions of the support they received in their courses were positive overall (mean values: 3.3–4.2; min. 1, max. 5, Table 2). Despite the different methods of assessing problem solving, the two courses showed no considerable differences in the students' perceptions of such support. For example, students felt they received constructive feedback on problem solving (3.9 and 3.8).

Students' perceptions of their learning outcomes were also high (mean values: 3.2–4.0, Table 3). A small difference between the courses was observed in terms of students' improved skills in presenting reasons for physical phenomena in the SCAN course (4.0) compared to the teacher-led course (3.7; Table 3). The highest ratings were given to domain-specific outcomes, such as enhanced problem solving skills (3.9, 3.7), ability to apply theoretical concepts and models (3.8, 3.9) and ability to explain physical phenomena (3.7, 4.0).

Table 2. Students' perceptions of the support received by the learning environment in two courses.

Received support	Course 1 (teacher-led)	Course 2 (SCAN)	*p*-value
Support from other students	4.2 (0.9)	4.1 (0.8)	0.5
Constructive feedback related to problem solving	3.9 (0.8)	3.8 (0.9)	0.5
Enjoyment of studying	3.7 (1.2)	3.6 (1.2)	0.6
Teaching for understanding	3.5 (0.9)	3.5 (0.8)	1.0
Understanding the demands of problem solving	3.3 (1.1)	3.4 (1.1)	0.7

Table 3. Students' perceptions of their learning outcomes in two courses.

Learning outcomes	Course 1 (teacher-led)	Course 2 (SCAN)	*p*-value
My ability to solve problems was enhanced during the course.	3.9 (0.9)	3.7 (1.0)	0.4
I learned to apply theoretical concepts and models during this course.	3.8 (0.7)	3.9 (0.7)	0.5
I learned to present the reasons for physical phenomena.	3.7 (0.9)	4.0 (0.7)	0.03
I learned to analyse physical phenomena from different points of view.	3.7 (0.8)	3.9 (0.7)	0.3
My collaborative skills were honed during the course.	3.6 (0.9)	3.6 (0.9)	1.0
I learned to analyse and specify concepts and models.	3.5 (0.9)	3.5 (0.9)	1.0
I learned to evaluate solutions critically.	3.4 (0.9)	3.7 (0.8)	0.2
My skills in presenting my opinions during interactive situations were fortified during the course.	3.3 (1.1)	3.3 (0.9)	0.9
I was able to refine my ability to develop new ideas.	3.2 (1.0)	3.2 (0.7)	0.9

5.2 Perceptions of the SCAN method and their associations with learning experiences and learning outcomes

We identified differences in terms of how students perceived the SCAN method as beneficial (factor 1) and reliable (factor 2, Figure 4). These differences became visible when the students compared the two courses with different methods for assessing problem solving (Table 4). First, the students who perceived the SCAN method more beneficial agreed more strongly that they had a deeper approach to learning in Course 2 (SCAN method) than in Course 1 (teacher-led method). They also perceived that the SCAN method provided better support to their learning than the teacher-led method. The students who perceived the SCAN as a more unreliable method correspondingly considered the assessment in Course 1 more reliable than in Course 2. The students experienced that they relied on teachers' assessment more than their own assessment, and the perceptions of the SCAN as a beneficial or reliable assessment method were only weakly associated with these experiences (adjusted R^2 = 0.08, see Table 4).

Perceived utility and reliability were interrelated by a correlation factor of 0.64.

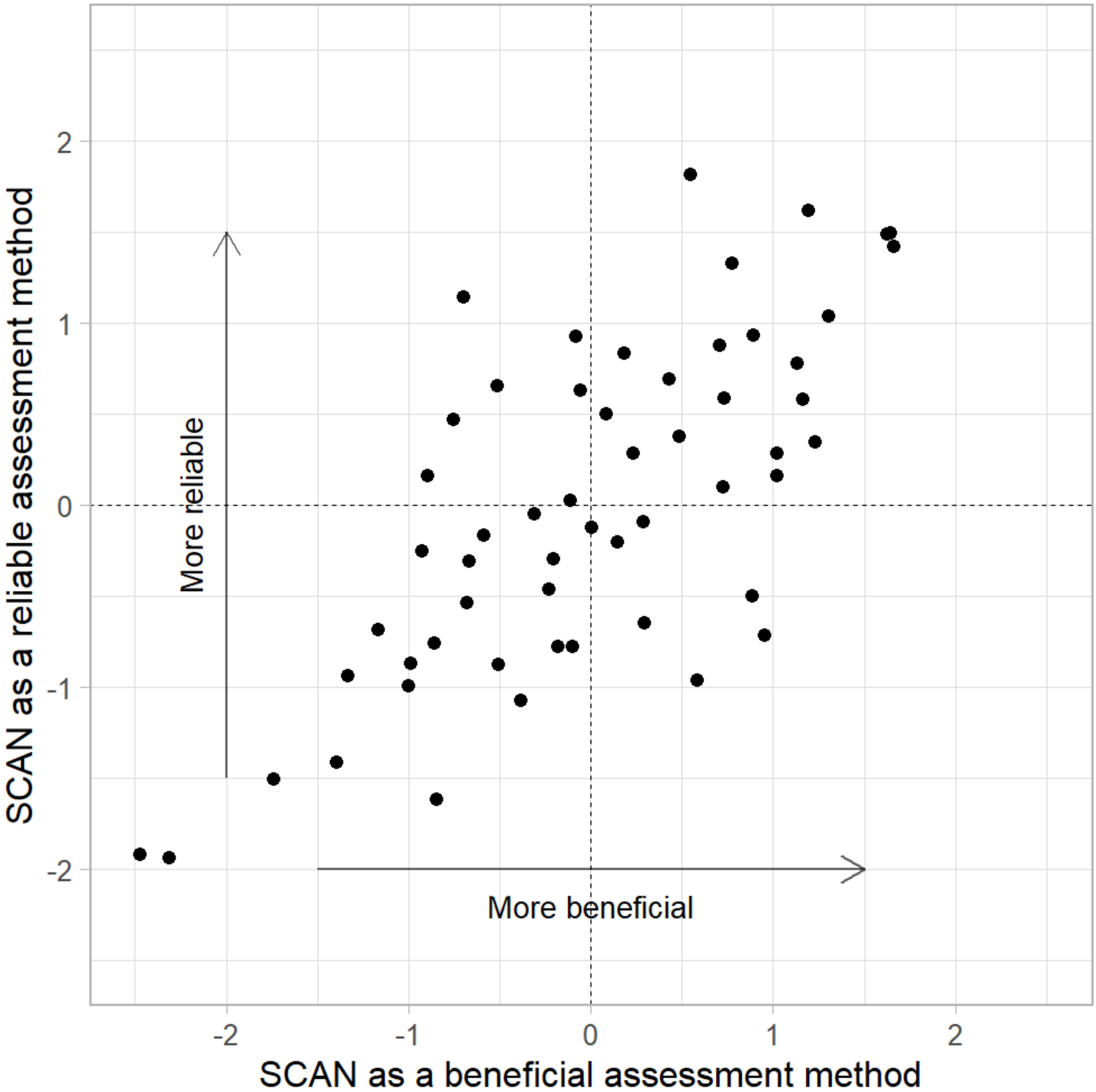


Figure 4. Distribution of the scores of factor 1 (SCAN as a beneficial method) and factor 2 (SCAN as a reliable method). Dashed lines show the mean values of the factor scores of the whole sample (N = 56).

Students who perceived SCAN as a more beneficial method experienced more support from the learning environment when considering constructive feedback related to problem solving, teaching for understanding and enjoyment of studying (Table 5). Moreover, these students also perceived more support from other students even though the students' perceptions of the SCAN method explained less variation in this variable compared to the former three items (adjusted $R^2 = 0.21$). The students who perceived the SCAN as a more reliable assessment method experienced more support for understanding the demands of problem solving ($b = 0.46$, s.e. = 0.19, $p = 0.02$).

Table 4. The regression coefficients (*b*) with their standard errors (s.e.) and adjusted R-squared when the perceptions of the SCAN method were used to predict students' comparisons of Course 1 (teacher-led method) and Course 2 (SCAN method).

Statement	Intercept	SCAN as a beneficial method			SCAN as a reliable method			R^2	$F(2, 52)^a$; *p*
		b	s.e.	*p*	*b*	s.e.	*p*		
I adopted a deeper approach to learning in Course 2 than in Course 1 due to the SCAN method.	2.83	**0.90**	**0.19**	**< 0.001**	−0.04	0.19	0.83	0.45	22.95; < 0.001
The SCAN method in Course 2 supported my learning better than the teacher-led method in Course 1.	3.22	**0.70**	**0.16**	**< 0.001**	0.20	0.17	0.23	0.51	29.39; < 0.001
The assessment of problem solving was more reliable in Course 1 than in Course 2.	3.31	0.28	0.19	0.14	**−0.78**	**0.19**	**< 0.001**	0.26	10.66; < 0.001
I rely more on the teacher's assessment than my self-assessment.	3.78	0.12	0.17	0.49	−0.40	0.18	0.03	0.08	3.50; 0.04

[b] One student did not answer all the statements (*N* = 55).

Table 5. The regression coefficients (*b*) with their standard errors (s.e.) and adjusted R-squared when the perceptions of the SCAN method were used to predict the perceived support provided by the learning environment in Course 2 (SCAN method).

Received support	Intercept	SCAN as a beneficial method			SCAN as a reliable method			R^2	$F(2, 51)^a$; *p*
		b	s.e.	*p*	*b*	s.e.	*p*		
Support from other students	4.11	**0.53**	**0.14**	**< 0.001**	−0.25	0.15	0.10	0.21	7.84; 0.001
Constructive feedback related to problem solving	3.81	**0.80**	**0.13**	**< 0.001**	−0.15	0.14	0.27	0.52	29.68; < 0.001
Enjoyment of studying	3.57	**0.87**	**0.19**	**< 0.001**	−0.11	0.19	0.59	0.42	20.00; < 0.001
Teaching for understanding	3.52	**0.41**	**0.10**	**< 0.001**	0.16	0.11	0.15	0.51	28.79; < 0.001
Understanding the demands of problem solving	3.38	0.11	0.18	0.53	0.46	0.19	0.02	0.25	9.63; < 0.001

[a] Two students did not answer all the statements (*N* = 54).

Students' perceptions of the SCAN explained weakly the variation in the students' learning outcomes (Table 6; adjusted $R^2 < 0.21$ and $p > 0.001$ in all the items). However, the students who perceived the SCAN more beneficial method had a slightly better problem-solving ability. Regarding the perceived learning outcomes (see also Table 3), these students strongly experienced that their ability to solve problems was enhanced during the course. Moreover, the students who found SCAN to be more beneficial also perceived that they learned to evaluate solutions critically, apply theoretical concepts and models, and present the reasons for physical phenomena to a greater extent. These students also perceived more improvement in their collaborative skills.

Table 6. The regression coefficients (*b*) with their standard errors (s.e.) and adjusted R-squared when the perceptions of the SCAN method were used to predict students' problem-solving ability (aggregated gradings of the weekly problem-solving solutions) and perceived learning outcomes in Course 2 (SCAN method).

Learning outcomes	Intercept	SCAN as a beneficial method			SCAN as a reliable method			R^2	$F(2,51)^a$; *p*
		b	s.e.	*p*	*b*	s.e.	*p*		
Problem solving ability	3.33	0.45	0.19	0.018	−0.17	0.19	0.37	0.096	3.8; 0.028
My ability to solve problems was enhanced during the course.	3.81	0.45	0.15	0.0037	−0.22	0.15	0.16	0.14	5.3; 0.008
I learned to apply theoretical concepts and models during this course.	3.85	0.48	0.14	0.0011	−0.16	0.14	0.25	0.21	8.0; < 0.001
I learned to evaluate solutions critically.	3.61	0.43	0.14	0.0038	−0.13	0.14	0.39	0.17	6.5; 0.003
I learned to analyse and specify concepts and models.	3.60	0.18	0.15	0.22	0.26	0.15	0.084	0.21	8.2; < 0.001
I learned to analyse physical phenomena from different points of view.	3.85	0.31	0.15	0.052	−0.084	0.16	0.60	0.067	2.9; 0.064
I was able to refine my ability to develop new ideas.	3.30	0.15	0.17	0.88	−0.021	0.17	0.91	0.00	0.68; 0.51
I learned to present the reasons for physical phenomena.	3.94	0.35	0.16	0.038	−0.11	0.17	0.52	0.076	3.19; 0.050
My collaborative skills were honed during the course.	3.61	0.48	0.18	0.010	−0.20	0.19	0.29	0.11	4.4; 0.017
My skills in presenting my opinions during the interactive situations were fortified during the course.	3.39	0.57	0.19	0.0042	−0.26	0.20	0.18	0.14	5.3; 0.008

[a] Two students did not answer all the statements ($N = 54$).

6 Discussion

In this study, we proposed a novel self-assessment method called the SCAN method of assessing problem solving that included both formative and summative elements. We studied students' learning experiences in two courses that had different methods of assessing problem solving—the teacher-led method based on teacher assessment and the SCAN method (RQ1, section 5.1). We also studied associations between these learning experiences and students' perceptions of the SCAN method (RQ2, section 5.2). On average, the students perceived high support by the learning environment and achieved good learning outcomes regardless of the method of assessing problem solving. We also found that following the completion of the two courses, the students reported better reliance on the teacher's assessment than the self-assessment. Although summative self-assessment (self-grading) has rarely been implemented in higher education (Häsä et al., 2021), its combination with formative self-assessment can change the students' perceptions of self-assessment over time and, ultimately, foster their lifelong learning skills (Nieminen et al., 2021). After all, the students were first-year students, and according to the studies (e.g. Virtanen et al., 2009), students' skills to assess their own learning develop with practice.

Even though students had different perceptions of the beneficialness and reliability of the self-assessment (see similar findings in Kangaslampi et al., 2022), their problem-solving ability seemed to be only weakly, but positively related to the perceived utility of the self-assessment; we found no association between the problem-solving ability and perceived reliability of the self-assessment. The students who perceived SCAN as a beneficial method agreed more strongly that they adopted a deeper approach to learning and received more support for their learning than they did in the course using the teacher-led assessment method. This finding supports the previous finding according to which some students may vary their approaches to learning with changing learning environment characteristics (Kangaslampi et al., 2022; Lindblom-Ylänne et al., 2004). At the same time, the students who perceived SCAN as an unreliable method perceived less support for understanding the demands of problem solving. This finding illustrates the importance of SCAN's negotiation stage, where teachers and students can discuss the utility and reliability of the different assessment methods (Yan et al., 2021; Yan & Carless, 2021). The negotiation stage also allows supporting students to assess their own learning and raising the awareness of the alignment between teaching and assessment methods, and the intended learning outcomes.

Although our study deepens the understanding of students' perceptions of the self-assessment in their learning experiences, some limitations must be considered. First, the sample size of our study was small and includes only higher education physics students. Future studies should examine whether our findings generalise to other contexts. Second, even though students' perceptions of their own learning outcomes were similar in both courses, actual learning gains may be higher than the self-report gains in the presence of active learning strategies such as the SCAN (Deslauriers et al., 2019). Moreover, certain items in the questionnaire (e.g., "I adopted a deeper approach to learning in Course 2 than in Course 1 due to the SCAN method", see Table 4) could potentially introduce bias by tempting respondents to prefer choices that they find desirable (Paulhus, 2002). Third, although similar and comparable, the two courses still had different contents, which may influence the students' perceptions of the learning outcomes and support provided by the learning environment. Using the data from these two courses enabled focusing on the perceptions of the same sample of students. Fourth, although problem solving paths may vary, the use of the SCAN method requires that the problems have well-defined model solutions. While real-life problems rarely have pre-determined solutions (Randles et al., 2018), students may benefit from the SCAN method and related critical analysis of the chosen problem-solving actions.

Despite these limitations, our study has several implications. First, the SCAN method is an easy and ready-to-use assessment practice involving handy technical realisation. Such practices are needed to reduce the barriers to implementing active learning (see Børte et al., 2023) as a part of various instructional strategies, including distance learning and teaching. The SCAN method is also flexible. It can be used for peer assessment, and the teacher can customise the assessment rubric to emphasise different aspects of problem solving (creativity, argumentation, correct answers, organised process and correcting own mistakes) and the unique characteristics of the different STEM fields. Second, the SCAN method may foster reflective study practices (Lindblom-Ylänne et al., 2019) and help redirect the teaching resources to the interaction between teacher and students while reducing teachers' straightforward grading duties (Leung et al., 2017). At the same time, we argue that successful implementation of self-assessment-based practices for problem solving requires that teachers are aware of the different perceptions that students have towards these practices. In this way, the support provided by the learning environment can be personalized so that all the students can practice and acquire the necessary problem-solving and self-assessment skills needed in lifelong learning.

Since the first-year university students would preferably trust teacher assessment because of the perceived low reliability of self-grading, future research could focus on the development of students' perceptions of self-assessment and the associated skills. Future studies could also examine the learning processes that occur in the different stages of the SCAN method and investigate differences between first-year students and more experienced students. Future research on the development of problem-solving and self-assessment skills among university students is important, since these features are key skills in contemporary society.

7 Conclusions

Central goals of higher education in STEM fields include learning problem-solving and self-assessment skills. In these fields, which frequently involve problems with well-defined solutions, self-assessment methods can be effective for teaching students problem-solving procedures, critical analysis, and self-assessment skills (Panadero et al., 2017; Randles et al., 2018; Yan et al., 2021). In this study, we increased understanding of undergraduate students' perceptions of self-assessment when it included both formative and summative elements. We found that the perceived reliability of self-assessment was associated with its perceived utility. The students' problem-solving abilities were only weakly related to their perceptions of the utility of self-assessment. Our findings also showed that the students relied more on teachers' assessments than on their own self-assessments. These results can be used to promote the successful and sustainable implementation of innovative self-assessment-based practices for problem solving by increasing teachers' awareness of the various perceptions that students may have towards these practices. In particular, SCAN's negotiation stage provides possibilities for teachers and students to discuss the utility and reliability of the different assessment methods. The negotiation stage also allows supporting students to assess their own learning and raising the awareness of the alignment between teaching and assessment methods, and the intended learning outcomes.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Acknowledgements

This work was supported by the Academy of Finland under Grant number 318905; and the Finnish Ministry of Education and Culture under Grant number OKM/280/523/2017.

References

Andrade, H. L. (2019). A Critical Review of Research on Student Self-Assessment. *Frontiers in Education*, *4*(87), 1–13. https://doi.org/10.3389/feduc.2019.00087

Andrade, H. L., & Du, Y. (2007). Student responses to criteria-referenced self-assessment. *Assessment & Evaluation in Higher Education*, *32*(2), 159–181. https://doi.org/10.1080/02602930600801928

Andrade, H. L., & Valtcheva, A. (2009). Promoting Learning and Achievement Through Self-Assessment. *Theory Into Practice*, *48*(1), 12–19. https://doi.org/10.1080/00405840802577544

Ashenafi, M. M. (2017). Peer-assessment in higher education – twenty-first century practices, challenges and the way forward. *Assessment & Evaluation in Higher Education*, *42*(2), 226–251. https://doi.org/10.1080/02602938.2015.1100711

Biggs, J. B., & Tang, C. S. (2011). *Teaching for quality learning at university: What the student does* (4th edition). McGraw-Hill, Society for Research into Higher Education & Open University Press.

Børte, K., Nesje, K., & Lillejord, S. (2023). Barriers to student active learning in higher education. *Teaching in Higher Education*, *28*(3), 597–615. https://doi.org/10.1080/13562517.2020.1839746

Boud, D., & Falchikov, N. (1989). Quantitative Studies of Student Self-Assessment in Higher Education: A Critical Analysis of Findings. *Higher Education*, *18*(5), 529–549. https://doi.org/10.1007/BF00138746

Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment & Evaluation in Higher Education*, *31*(4), 399–413. https://doi.org/10.1080/02602930600679050

Bourke, R. (2018). Self-assessment to incite learning in higher education: Developing ontological awareness. *Assessment & Evaluation in Higher Education*, *43*(5), 827–839. https://doi.org/10.1080/02602938.2017.1411881

Broadbent, J., Panadero, E., & Boud, D. (2018). Implementing summative assessment with a formative flavour: A case study in a large class. *Assessment & Evaluation in Higher Education*, *43*(2), 307–322. https://doi.org/10.1080/02602938.2017.1343455

Candel, C., Vidal-Abarca, E., Cerdán, R., Lippmann, M., & Narciss, S. (2020). Effects of timing of formative feedback in computer-assisted learning environments. *Journal of Computer Assisted Learning*, *36*(5), 718–728. https://doi.org/10.1111/jcal.12439

Carroll, D. (2020). Observations of student accuracy in criteria-based self-assessment. *Assessment & Evaluation in Higher Education*, *45*(8), 1088–1105. https://doi.org/10.1080/02602938.2020.1727411

de Jong, T. (2019). Moving towards engaged learning in STEM domains; there is no simple answer, but clearly a road ahead. *Journal of Computer Assisted Learning*, *35*(2), 153–167. https://doi.org/10.1111/jcal.12337

Deslauriers, L., McCarty, L. S., Miller, K., Callaghan, K., & Kestin, G. (2019). Measuring actual learning versus feeling of learning in response to being actively engaged in the classroom.

Proceedings of the National Academy of Sciences, *116*(39), 19251–19257. https://doi.org/10.1073/pnas.1821936116

Docktor, J. L., Dornfeld, J., Frodermann, E., Heller, K., Hsu, L., Jackson, K. A., Mason, A., Ryan, Q. X., & Yang, J. (2016). Assessing student written problem solutions: A problem-solving rubric with application to introductory physics. *Physical Review Physics Education Research*, *12*(1), 010130. https://doi.org/10.1103/PhysRevPhysEducRes.12.010130

Entwistle, N., & McCune, V. (2004). The Conceptual Bases of Study Strategy Inventories. *Educational Psychology Review*, *16*(4), 325–345. https://doi.org/10.1007/s10648-004-0003-0

Entwistle, N., McCune, V., & Hounsell, J. (2003). Investigating ways of enhancing university teaching-learning environments: Measuring students' approaches to studying and perceptions of teaching. In *Powerful learning environments: Unravelling basic components and dimensions* (pp. 89–108). Elsevier Science Limited. https://www.research.ed.ac.uk/en/publications/investigating-ways-of-enhancing-university-teaching-learning-envi

Fryer, L. K., & Vermunt, J. D. (2018). Regulating approaches to learning: Testing learning strategy convergences across a year at university. *British Journal of Educational Psychology*, *88*(1), 21–41. https://doi.org/10.1111/bjep.12169

Häsä, J., Rämö, J., & Nieminen, J. H. (2021). Supporting quality of learning by letting students give their own grades: An innovative self-assessment model in university mathematics. In Z. Yan & L. Yang (Eds.), *Assessment as Learning: Maximising Opportunities for Student Learning and Achievement* (1st edition, pp. 232–243). Routledge.

Jax, J., Ahn, J. N., & Lin-Siegler, X. (2019). Using contrasting cases to improve self-assessment in physics learning. *Educational Psychology*, *39*(6), 815–838. https://doi.org/10.1080/01443410.2019.1577360

Kangaslampi, R., Asikainen, H., & Virtanen, V. (2022). Students' perceptions of self-assessment and their approaches to learning in university mathematics. *LUMAT: International Journal on Math, Science and Technology Education*, *10*(1), Article 1. https://doi.org/10.31129/LUMAT.10.1.1604

Kearney, S. (2013). Improving engagement: The use of 'Authentic self-and peer-assessment for learning' to enhance the student learning experience. *Assessment & Evaluation in Higher Education*, *38*(7), 875–891. https://doi.org/10.1080/02602938.2012.751963

Koskinen, P. (2012). Tietokone avuksi pienryhmien ripeään muodostamiseen (engl. Computer assistance for the rapid formation of small groups). *Yliopistopedagogiikka*, *19*(1), 26–27. https://lehti.yliopistopedagogiikka.fi/2012/05/03/tietokone-avuksi-pienryhmien-ripeaan-muodostamiseen/

Koskinen, P., & Lämsä, J. (2019). Ruotiminen: Toimintamalli harjoitustehtävien läpikäyntiin. *Yliopistopedagogiikka*, *26*(2), 52–55.

Koskinen, P., Lämsä, J., Maunuksela, J., Hämäläinen, R., & Viiri, J. (2018). Primetime learning: Collaborative and technology-enhanced studying with genuine teacher presence. *International Journal of STEM Education*, *5*(20), 1–13. https://doi.org/10.1186/s40594-018-0113-8

Leung, A. C. K., Hashemi Pour, B., Reynolds, D., & Jerzak, S. (2017). New assessment process in an introductory undergraduate physics laboratory: An exploration on collaborative learning. *Assessment & Evaluation in Higher Education*, *42*(2), 169–181. https://doi.org/10.1080/02602938.2015.1089977

Lindblom-Ylänne, S., Parla, A., & Postareff, L. (2004). Challenges in analysing change in students' approaches to learning. In D. Gijbels, V. Donche, J. T. E. Richardson, & J. D. Vermunt (Eds.), *Learning Patterns in Higher Education* (pp. 232–248). Routledge.

Lindblom-Ylänne, S., Parpala, A., & Postareff, L. (2019). What constitutes the surface approach to learning in the light of new empirical evidence? *Studies in Higher Education*, *44*(12), 2183–2195. https://doi.org/10.1080/03075079.2018.1482267

Marton, F., & Säljö, R. (1979). On qualitative differences in learning: I—Outcome and process. *British Journal of Educational Psychology*, *46*(1), 4–11. https://doi.org/10.1111/j.2044-8279.1976.tb02980.x
McCune, V., & Entwistle, N. (2011). Cultivating the disposition to understand in 21st century university education. *Learning and Individual Differences*, *21*(3), 303–310. https://doi.org/10.1016/j.lindif.2010.11.017
McNeish, D., & Wolf, M. G. (2020). Thinking twice about sum scores. *Behavior Research Methods*, *52*(6), 2287–2305. https://doi.org/10.3758/s13428-020-01398-0
Nieminen, J. H., Asikainen, H., & Rämö, J. (2021). Promoting deep approach to learning and self-efficacy by changing the purpose of self-assessment: A comparison of summative and formative models. *Studies in Higher Education*, *46*(7), 1296–1311. https://doi.org/10.1080/03075079.2019.1688282
Panadero, E., Brown, G. T. L., & Strijbos, J.-W. (2016). The Future of Student Self-Assessment: A Review of Known Unknowns and Potential Directions. *Educational Psychology Review*, *28*(4), 803–830. https://doi.org/10.1007/s10648-015-9350-2
Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. https://doi.org/10.1016/j.edurev.2017.08.004
Parpala, A., Lindblom-Ylänne, S., Komulainen, E., & Entwistle, N. (2013). Assessing students' experiences of teaching–learning environments and approaches to learning: Validation of a questionnaire in different countries and varying contexts. *Learning Environments Research*, *16*(2), 201–215. https://doi.org/10.1007/s10984-013-9128-8
Parpala, A., Lindblom-Ylänne, S., Komulainen, E., Litmanen, T., & Hirsto, L. (2010). Students' approaches to learning and their experiences of the teaching–learning environment in different disciplines. *British Journal of Educational Psychology*, *80*(2), 269–282. https://doi.org/10.1348/000709909X476946
Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In *The role of constructs in psychological and educational measurement* (pp. 49–69). Lawrence Erlbaum Associates Publishers.
Postareff, L., Parpala, A., & Lindblom-Ylänne, S. (2015). Factors contributing to changes in a deep approach to learning in different learning environments. *Learning Environments Research*, *18*(3), 315–333. https://doi.org/10.1007/s10984-015-9186-1
Randles, C., Overton, T., Galloway, R., & Wallace, M. (2018). How do approaches to solving open-ended problems vary within the science disciplines? *International Journal of Science Education*, *40*(11), 1367–1390. https://doi.org/10.1080/09500693.2018.1503432
Struyven, K., Dochy, F., Janssens, S., & Gielen, S. (2006). On the dynamics of students' approaches to learning: The effects of the teaching/learning environment. *Learning and Instruction*, *16*(4), 279–294. https://doi.org/10.1016/j.learninstruc.2006.07.001
Taras, M. (2016). Situating power potentials and dynamics of learners and tutors within self-assessment models. *Journal of Further and Higher Education*, *40*(6), 846–863. https://doi.org/10.1080/0309877X.2014.1000283
Tremblay, K., Lalancette, D., & Roseveare, D. (2012). *Assessment of Higher Education Learning Outcomes—AHELO. Feasibility study report. Volume 1: Design and Implementation.* OECD. https://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume1.pdf
Tynjälä, P., Virtanen, A., Klemola, U., Kostiainen, E., & Rasku-Puttonen, H. (2016). Developing social competence and other generic skills in teacher education: Applying the model of integrative pedagogy. *European Journal of Teacher Education*, *39*(3), 368–387. https://doi.org/10.1080/02619768.2016.1171314
Utriainen, J., Tynjälä, P., Kallio, E., & Marttunen, M. (2018). Validation of a modified version of the Experiences of Teaching and Learning Questionnaire. *Studies in Educational Evaluation*, *56*, 133–143. https://doi.org/10.1016/j.stueduc.2017.12.007

van Merrienboer, J. J. G., Kirschner, P. A., & Kester, L. (2003). Taking the Load Off a Learner's Mind: Instructional Design for Complex Learning. *Educational Psychologist*, *38*(1), 5–13. https://doi.org/10.1207/S15326985EP3801_2

Vermunt, J. D., & Donche, V. (2017). A Learning Patterns Perspective on Student Learning in Higher Education: State of the Art and Moving Forward. *Educational Psychology Review*, *29*(2), 269–299. https://doi.org/10.1007/s10648-017-9414-6

Virtanen, A., & Tynjälä, P. (2019). Factors explaining the learning of generic skills: A study of university students' experiences. *Teaching in Higher Education*, *24*(7), 880–894. https://doi.org/10.1080/13562517.2018.1515195

Virtanen, A., Tynjälä, P., & Collin, K. (2009). Characteristics of Workplace Learning Among Finnish Vocational Students. *Vocations and Learning*, *2*(3), 153–175. https://doi.org/10.1007/s12186-009-9022-9

Wancham, K., & Tangdhanakanond, K. (2022). Effects of Feedback Types and Opportunities to Change Answers on Achievement and Ability to Solve Physics Problems. *Research in Science Education*, *52*(2), 427–444. https://doi.org/10.1007/s11165-020-09956-4

Yan, Z., & Carless, D. (2021). Self-assessment is about more than self: The enabling role of feedback literacy. *Assessment & Evaluation in Higher Education*, 1–13. https://doi.org/10.1080/02602938.2021.2001431

Yan, Z., Wang, X., Boud, D., & Lao, H. (2021). The effect of self-assessment on academic performance and the role of explicitness: A meta-analysis. *Assessment & Evaluation in Higher Education*, *0*(0), 1–15. https://doi.org/10.1080/02602938.2021.2012644

Yan, Z., Wang, X., Boud, D., & Lao, H. (2023). The effect of self-assessment on academic performance and the role of explicitness: A meta-analysis. *Assessment & Evaluation in Higher Education*, *48*(1), 1–15. https://doi.org/10.1080/02602938.2021.2012644

Appendix

Table 7. Items that were used to measure three aspects of the support provided by the learning environment: support from other students; constructive feedback related to problem solving; teaching for understanding.

Aggregated scale	Items
Support from other students	Talking to other students helped me understand things better. Working with other students was natural for me. If necessary, I also received support from other students outside of small group meetings and "primetime" meetings (meetings between the teacher and a small group). I received support from other students during small group meetings and "primetime" meetings.
Constructive feedback related to problem solving	Reviewing the problem-solving solutions helped improve my learning and study. Reviewing the problem-solving solutions helped clarify things I had not fully understood.
Teaching for understanding	Teaching helped me to find reasons for different perspectives and solutions. The course inspired me to combine new concepts and models with those previously learned. Reviewing the problem-solving solutions was meaningful and appropriate.