Studies in Educational Evaluation 63 (2019) 72–82

Contents lists available at ScienceDirect

Studies in Educational Evaluation

journal homepage: www.elsevier.com/locate/stueduc

Students' perceptions of assessment quality related to their learning approaches and learning outcomes

Karin J. Gerritsen-van Leeuwenkamp[a,b,*], Desirée Joosten-ten Brinke[a,c], Liesbeth Kester[d]

[a] Open University of The Netherlands, PO Box 2960, 6401 DL Heerlen, The Netherlands
[b] Saxion, University of Applied Sciences, PO Box 70.000, 7500 KB Enschede, The Netherlands
[c] Fontys University of Applied Sciences, PO Box 347, 5600 AH Eindhoven, The Netherlands
[d] Utrecht University, PO Box 80125, 3508 TC Utrecht, The Netherlands

ARTICLE INFO

Keywords:
Assessment quality
Student perception
Learning approach
Learning outcome
Student evaluation

ABSTRACT

This study aims to provide more insight into how students' perceptions of assessment quality are related to their learning approaches and learning outcomes. Six variables associated with the construct of students' perceptions of assessment quality are distinguished: 1) effects of assessment on learning, 2) fairness of assessment, 3) conditions of assessment, 4) interpretation of test scores, 5) authenticity of assessment, and 6) credibility of assessment. 204 higher education students completed the Students' Perceptions of Assessment Quality Questionnaire (SPAQQ), and the Approaches to Learning and Studying Inventory (ALSI), and the students' learning outcomes (grades) were obtained. Firstly, results indicate that the students' perceptions of the effects of assessment on learning are positively related to the students' deep learning approach and the strategic learning approach and negatively related to the surface learning approach. Secondly, the students' perceptions of the conditions of assessment are positively related to their learning outcomes of the assessments.

1. Introduction

In higher education, assessment is one of the most important factors in a learning environment; it drives students' learning by providing them with insight into their learning progress as well as information about the intended learning goals and how those can be achieved (ARG, 2002; Boud and Associates, 2010; Watkins, Dahlin, & Ekholm, 2005). Moreover, students are selected and certified based on assessment (Stobart, 2008). Given the major influence of assessment on learning, the success of educational programmes depends not only on assessment itself but, even more, on the quality of assessment (Van der Vleuten, Sluijsmans, & Joosten-ten Brinke, 2017). *Assessment quality* refers to more than the quality of a single assessment; it encompasses the quality of all the evaluation practices' elements (i.e. the assessment, the test questions, the assignments, the criteria, the score reports, the procedures, the feedback, the programmes, and the policies) (Gerritsen-van Leeuwenkamp, Joosten-ten Brinke, & Kester, 2017).

Inferior assessment quality has major negative implications for students. For example, when assessment is not aligned with the goals of the learning process, validity is jeopardised, which can hinder students' learning (Baartman, Bastiaens, Kirschner, & Van der Vleuten, 2006; Martin, 1997). When students experience these implications of assessment quality, they form perceptions of assessment. Considering that the students themselves determine how to prepare for and participate in assessment and what they will do with the feedback and the outcomes (Cowie, 2009; Segers, Dochy, Gijbels, & Struyven, 2009), it is important that educational organisations understand how students perceive the quality of these assessment practices. In their efforts towards attaining high-quality assessment practices to enhance students' learning, it is necessary for educational organisations to deliberately consider these perceptions.

However, currently, little is known about how students' perceptions of assessment quality are related to their learning approaches (i.e. their intentions, motives, and processes they use to learn) and their learning outcomes (i.e. grades). Therefore, the aim of this research is to provide insight into the relationships between students' perceptions of assessment quality, their learning approaches, and their learning outcomes in the context of higher education. In the next paragraphs, the main topics of interest of this research will be further explored.

* Corresponding author. Saxion University of Applied Sciences, PO Box 70.000, 7500 KB Enschede, The Netherlands.
E-mail addresses: k.j.vanleeuwenkamp@saxion.nl (K.J. Gerritsen-van Leeuwenkamp), desiree.joosten-tenbrinke@ou.nl (D. Joosten-ten Brinke), l.kester@uu.nl (L. Kester).

https://doi.org/10.1016/j.stueduc.2019.07.005
Received 23 July 2018; Received in revised form 7 July 2019; Accepted 8 July 2019
Available online 05 August 2019
0191-491X/

1.1. Learning approaches and learning outcomes in higher education

Graduates of higher education should be competent in their profession and leveraging advanced knowledge (Bologna Working Group, 2005). Students vary their learning approaches (i.e. their intentions, motives, and learning processes to study) to achieve these learning goals. They 'tend to be very much aware of a range of ways of learning and the various kinds of information that they need to be able to acquire in today's societies' (Birenbaum et al., 2006, p. 63). The effectiveness of a learning approach depends on the intended learning goals and will lead to certain learning outcomes (Byrne, Flood, & Willis, 2002).

Firstly, students may intend to understand the subject matter; these students' learning processes are focused on making sense of new ideas and using evidence to refine their prior knowledge. Thus, they are engaged in elaboration and induction, which are learning processes related to a deep learning approach. Comprehension is the expected learning outcome of this approach (Byrne et al., 2002; Entwistle & McCune, 2004; Entwistle & Ramsden, 1982; Marton & Säljö, 2005; Mayer, 2002). A deep learning approach is related to personal interest, motivation, and self-regulation (Entwistle & McCune, 2004).

Secondly, students may intend to reproduce the learning content; these students' learning processes are focused on memorising and rehearsing the information that needs to be learned without understanding. Such learning processes are related to a surface learning approach, and reproduction is the expected learning outcome (Byrne et al., 2002; Entwistle & McCune, 2004; Entwistle & Ramsden, 1982; Marton & Säljö, 2005; Mayer, 2002).

Thirdly, students may intend to study strategically, their learning processes are then focused on managing their time and effort and organising how they study. Such learning processes are related to a strategic learning approach. Achieving the highest possible grades on assessment within the available time is the expected learning outcome (Byrne et al., 2002; Entwistle & McCune, 2004; Entwistle & Ramsden, 1982).

Although, the outcomes of a deep learning approach meet the demands that knowledge and information societies impose on graduates of higher education (Birenbaum et al., 2006), a deep learning approach does not per definition lead to better assessment scores (Gijbels, Watering, Dochy, & Bossche, 2005; Minbashian, Huon, & Bird, 2004). For example, previous research showed that the strategic learning approach is positively related to final year students' study success and this relation was not found for the deep learning approach (Asikainen, Parpala, Lindblom-Ylänne, Vanthournout, & Coertjens, 2014). For study success, a deep approach only might not be enough (Asikainen et al., 2014). It might require a combination with a strategic learning approach (Asikainen, Parpala, Virtanen, & Lindblom-Ylänne, 2013; Entwistle, 2000, November). Then, students start on time with their learning, invest enough time and effort in processes and techniques that are focused on understanding (e.g. retrieval techniques) and this foster deeper learning and might lead to study success (Asikainen et al., 2013; Roediger & Butler, 2011).

1.2. Relationships between assessment and learning

Assessment affects learning because it provides students with insight into their learning progress as well as information about the intended learning goals and how those can be achieved (ARG, 2002; Boud and Associates, 2010; Watkins et al., 2005). Moreover, students are qualified based on assessment (Stobart, 2008). Students choose the most effective learning approach – not necessarily their preferred approach – in certain situations to adapt to the learning environment (strategic reasons) (Lindblom-Ylänne & Lonka, 1998; Lindblom, 2017). Assessment's effectiveness as a driving force for learning is hampered when there is a lack of constructive alignment between curriculum objectives, learning/teaching activities, and assessment (Biggs, 1996). When assessment does not match the intended learning goals, teaching, and learning, this can jeopardise students' learning (Segers et al., 2009). For example, regarding students' learning approaches, Scouller (1998) found that students are more likely to use a surface learning approach when they are studying for multiple-choice exams. In contrast, they use a deeper learning approach when they are preparing for essays. Furthermore, incidental testing may lead to a situation in which students study intensively just before the test instead of distributing their learning over time, which is more beneficial for long-term retention (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). Moreover, regarding students' learning outcomes, Van der Kleij, Feskens and Eggen (2015) found that the effect was greater when students received feedback with an explanation on a computer-based test item than when they were only given the right answer or information about the correctness of an answer. Furthermore, practice tests have a positive influence on students' performance (Dunlosky et al., 2013). Hence, the constructive alignment of learning goals, learning/teaching activities, and assessment are crucial (Biggs, 1996; S. A. Cohen, 1987). Because students' perceptions of assessment might influence the relations between assessment and learning (Boud, 1990; Heeneman, Oudkerk Pool, Schuwirth, Van der Vleuten, & Driessen, 2015; Scouller, 1998), they should also be considered in efforts towards achieving constructive alignment.

1.3. Relations between students' perceptions of assessment quality and learning

Students form perceptions, whether consciously or not; they elaborate and assign meaning (Zimbardo, Weber, & Johnson, 2009) to all the characteristics of assessment they experience as they study, such as its form, demands, and importance. For example, students perceive mistakes made by examiners while assessing (e.g. grading inconsistencies due to biases) (Holmes & Smith, 2003), and these mistakes can lead to lower grades, undeserved failure, and thus, obstruction of the students' learning progress. In addition, students are aware of ineffective feedback, which leads to a misunderstanding of their strengths and improvement points (Holmes & Smith, 2003; Segers, Gijbels, & Thurlings, 2008). Moreover, students are the ones doing the learning (Cowie, 2009), therefore, it is important that teachers and employers collaborate with students to achieve high-quality assessment practices (Van der Vleuten et al., 2017). As Boud (1995) stated: 'The perceptions and interactions of a student are more important to learning than what staff take for granted as the "reality" of the assessment' (p. 36).

Regarding the relations between students' perceptions about assessment *in general* and their learning approaches, Lizzio, Wilson and Simons (2002) found that students adopt a deeper learning approach when they perceive the assessment as being appropriate – meaning that the assessment includes feedback and is focused on the students' understanding of the content rather than on their memorisation of it. Furthermore, Scouller (1998) found that students who preferred writing an essay as an assessment method over multiple-choice exams were more likely to pass the assignment essay than students who preferred multiple-choice exams. Many students exert less effort when they perceive an assessment as low-stakes; thus, their performance does not match their ability (Wise & DeMars, 2005). Regarding the relations between students' perceptions about assessment in general and learning outcomes, De Kleijn, Mainhard, Meijer, Pilot and Brekelmans (2012) found a positive relation between students' perceptions of the degree of affiliation (interpersonal proximity) from their supervisors and their learning outcomes (final grades). Furthermore, Brown and Hirschfeld (2008) found that students' perceptions of assessment are related to their learning outcomes. For example, they found that students perform better and get higher grades when they perceive assessment as an instrument to take responsibility for their learning. These examples show not only that students vary their intentions, motives, and learning processes (learning approaches) to align with their perceptions of the

evaluation practices but also that students' perceptions are related to their learning outcomes.

Though the literature does look at students' perceptions of assessment in general, there is less focus on how their perceptions of assessment *quality* affect their learning approaches and outcomes. Gulikers, Kester, Kirschner, and Bastiaens (2008) showed that, according to students, when the assessment task and context meets the quality criterion of authenticity (i.e. the degree of compliance of assessment with professional practice), students used a deeper learning approach. Further, Segers et al. (2008) showed a negative relationship between a surface learning approach and students' perceptions of the transparency of the demands of a portfolio assessment. Gibbs (2010) described how students' perceptions of assessment quality influence their learning outcomes: 'In the National Student Survey in the UK [United Kingdom], the universities with the lowest ratings for assessment and feedback have the worst student performance and the highest drop-out' (p. 164).

Previous research by Gerritsen-van Leeuwenkamp, Joosten-ten Brinke and Kester (2018) showed that students' perceptions of assessment quality consist of six variables (see for a detailed description Appendix A): 1) effects of assessment on learning, which refers to the alignment between assessment and learning (e.g. motivation and feedback), 2) fairness of assessment, which refers to the reasonability and feasibility of the test requirements for students (e.g. validity and comparability), 3) conditions of assessment, which refers to assessment circumstances that cannot be controlled by students but that will influence them (e.g. manageability and soundness), 4) interpretation of test scores, which refers to the meaning of the students' test scores (e.g. construct validity and generalisability), 5) authenticity of assessment, which refers to the alignment of assessment with one's intended professional life (e.g. transparency and authenticity), and 6) credibility of assessment, which refers to the degree to which the students trust assessment (e.g. consistency and trustworthiness).

However, little is known about the extent to which students' perceptions of assessment quality and these six variables are related to their learning approaches and outcomes. This is important to understand because assessment quality affects students' learning. Furthermore, students' perceptions matter for students' learning, therefore, it is not sufficient to only consider objective measures of assessment quality. The outcome of this research might provide leverage points for educational organisations to optimise students' learning. Towards that end, the study posed two research questions:

1 What is the relation between students' perceptions of assessment quality, including the six variables, and their learning approaches?
2 What is the relation between students' perceptions of assessment quality, including the six variables, and their learning outcomes?

2. Method

2.1. Participants

Students (n = 204) from a university of applied sciences in the Netherlands (183 females, 21 males, M_{age} = 23.53, SD_{age} = 9.39, age range: 17–60 years) participated in the study. Students in the first and second year of the Bachelor's Podiatry (full-time), Nursing (full-time and part-time), and Art and Technology (full-time) programmes participated in the study.

2.2. Instruments

2.2.1. Students' perceptions of assessment quality questionnaire (SPAQQ)

The SPAQQ (Appendix A) was used to measure the participants' perceptions of assessment quality (Gerritsen-van Leeuwenkamp et al., 2018). This questionnaire is based on a literature review and is validated by a sample of students in higher education (Gerritsen-van Leeuwenkamp et al., 2018). The SPAQQ consists of 39 closed items covering the six assessment quality variables: 1) effects of assessment on learning, 2) fairness of assessment, 3) conditions of assessment, 4) interpretation of test scores, 5) authenticity of assessment, and 6) credibility of assessment. Students were asked to fill in their general perceptions of assessment quality 'at this point in time'. To make sure they refer to a comparable context and to make students alert for the same aspects in the context (Zimbardo et al., 2009), a description of assessment quality was given on each page of the questionnaire. This description explained that assessment quality refers to the quality of test items, assignments, assessment instruments (e.g. a knowledge test, an assignment, a performance assessment, or a portfolio assessment), test organisation, the professionality of assessors, and the quality of the assessment programme of their year of study. A seven-point Likert scale was used to register the students' responses to the items, ranging from: 1 (completely disagree) to 2 (disagree), 3 (somewhat disagree), 4 (neutral), 5 (somewhat agree), 6 (agree), and 7 (completely agree). The reliability of the scales ranged between .75 and .89 (see Table 1).

Table 1
Reliability of the SPAQQ.

Scale	Number of items	α	Mean	*SD*
1. Effects of assessment on learning	11	.89	4.93	0.97
2. Fairness of assessment	5	.78	5.22	1.06
3. Conditions of assessment	8	.81	4.96	1.01
4. Interpretation of test scores	4	.75	4.94	1.01
5. Authenticity of assessment	5	.76	4.89	1.00
6. Credibility of assessment	6	.81	4.67	1.14

2.2.2. Approaches to learning and studying inventory (ALSI)

A shortened version of the ALSI (Appendix B) was used to measure the students' approaches to learning and studying. It contained 18 items divided into five scales. The first three scales operationalise the following learning approaches: 1) deep learning approach, 2) surface learning approach, and 3) monitoring studying (referring to the metacognitive aspects of learning). The fourth and fifth scales operationalise the strategic learning approach: 4) organised studying, and 5) effort management (Entwistle & McCune, 2004; Entwistle, McCune, & Hounsell, 2002; ETL project, 2002). The translation of the shorter ALSI (18 items) was based on a validated Dutch translation of the longer version of the ALSI (36 items) by Struyven, Dochy, Janssens, and Gielen (2006). A five-point Likert scale was used to register the students' responses to the items, ranging from: 1 (disagree) to 2 (somewhat disagree), 3 (neutral), 4 (somewhat agree), and 5 (agree). The reliability of the scales ranged between .30 and .72 (see Table 2). The reliability coefficients indicated an insufficient degree of reliability for some of the scales. Therefore, the underlying factor structure was explored using principal axis factoring (PAF).

Oblique rotation, in the form of direct oblimin rotation, was chosen as the factor rotation method because correlations between the factors were expected (Field, 2013). Before executing the PAF, factorability was examined. The Bartlett's sphericity test result was significant, which indicates that the data were probably factorable, $p < .001$ (Brace, Kemp, & Snelgar, 2016). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was good (.79), which indicates the amount

Table 2
Reliability of the ALSI.

Scale	Number of items	α	Mean	*SD*
1. Deep learning approach	6	.67	3.95	0.53
2. Surface learning approach	4	.63	2.73	0.82
3. Monitoring studying	4	.45	3.87	0.51
4. Effort management	2	.30	3.86	0.78
5. Organised studying	2	.72	3.82	0.97

of variance within the data that can be explained by the factors (Brace et al., 2016). The Measure of Sampling Adequacy (MSA) for each variable indicated that all the items should be retained (the MSA values ranged between .60 and .88) (Brace et al., 2016). The PAF found five factors. Factor 1 combines items of the ALSI scale deep learning approach and of monitoring studying with the focus on understanding. Factor 2 contains only three of the items in the surface learning approach scale found in the ALSI. Factor 3 combines all the items in the organised studying and effort management scales in the ALSI. Factor 4 combines items from the deep learning approach and monitoring studying ALSI scales with a focus on overview. Factor 5 contains one item from the surface learning approach scale.

Based on the PAF, the following choices were made regarding the scale structure of the ALSI used in this study. Firstly, the deep learning approach scale contains all the items found in the ALSI. Secondly, the monitoring studying scale was dropped (4 items). Similar to previous findings of Entwistle and McCune (2004), the results of the PAF, showed relations between the items of deep learning approach and monitoring studying scales, so one might expect the integration of both scales to be in line with the PAF outcomes. However, Entwistle and McCune (2004) argued that these two scales encompass different concepts; the deep learning approach refers to understanding, relating ideas, and the use of evidence while monitoring studying refers to the metacognitive aspects of learning. Therefore, it is not plausible to integrate these two concepts. Secondly, the surface learning approach scale contains all four items found in the ALSI. Thirdly, the effort management and organised studying scales are clustered. This is plausible because these two scales were originally one factor, the strategic learning approach (Entwistle & McCune, 2004). This is in line with the findings reported in previous research by Mattick, Dennis and Bligh (2004); it is also in line with the Experiences of Teaching and Learning Questionnaire scoring key of which the ALSI is a component (ETL project, n.d.).

In summary, the ALSI used in this study contains the deep learning approach scale, the surface learning approach scale, and the strategic learning approach scale. The reliability of the scales ranged between .63 and .71 (see Table 3). These reliability coefficients indicate a sufficient degree of reliability for each scale. The documented reliability of these three scales ranged between .64 and .70 (Mattick et al., 2004).

2.2.3. Learning outcomes

Students' grades of assessment were used to measure students' learning outcomes. According to the assessment policy framework of the University of Applied Sciences (which the bachelor's programmes of podiatry, nursing, and art and technology should meet), the two main purposes of assessment are enhancing students' learning and selection. The focus is on an alignment between assessment and learning. For example, there should be an alignment between the training profile and the assessment programme. Furthermore, in those assessment programmes, a range of assessment types is used that can be categorised into two main groups: 1) knowledge tests, in which only knowledge is measured (low-order cognitive knowledge, or 'knows' and 'knows how' in Miller's pyramid (Miller, 1990) and 2) other assessments (e.g. presentations, products, papers, performance assessments, portfolio assessments, or internships) in which knowledge, skills, and attitudes were integrated (higher-order cognitive knowledge or 'knows how', 'shows', and 'does' in Miller's pyramid (Miller, 1990). The learning outcomes ranges from 1–10.

Table 3
Reliability of the ALSI after the PAF.

Scale	Number of items	α	Mean	*SD*
1. Deep learning approach	6	.67	3.95	0.53
2. Surface learning approach	4	.63	2.73	0.82
3. Strategic learning approach	4	.71	3.84	0.77

2.3. Data collection procedure

Two weeks before the students received an e-mail invitation to complete the questionnaires, a total of 710 students were sent an e-mail informing them about the research study. When feasible, the students were allotted time in class to fill in the questionnaires, and absent students received a reminder to complete the questionnaires at home. After the first invitation, the students received three reminders at weekly intervals. A consent form was included in the questionnaires. An e-mail and a letter were sent to the guardians or parents of under-aged students requesting permission for the student to participate in the study. In all, 213 (30.1%) of the 710 students that were initially contacted responded to the invitation to participate in the study. Of those 213 students, nine were excluded (i.e. five under-aged students were excluded because no consent form was received from their guardian or parents; two students were excluded because they noted that they occasionally incorrectly interpreted the answer scale; and two students were excluded because they noted that they did not fill in the questionnaire seriously). To compensate students for participating in the study, a raffle was held and two 20€ gift vouchers were awarded.

The students granted permission for their learning outcomes to be retrieved from the grade registration system at the university of applied sciences. A mean grade was calculated based on all the assessments' grades. The number of assessments differed between the respondents (M_{amount} = 6.31, *SD* = 2.62, range: 1–12 assessments) due to differences in the educational programme, year of study, the number of resits, and the study progress.

Furthermore, to obtain more detailed insight into the relations between students' perceptions of assessment quality and their learning outcomes, the outcomes were divided in two learning outcome categories: 1) knowledge tests, in which only knowledge was measured and 2) other assessments, in which knowledge, skills, and attitudes were integrated (for example, presentations, papers, performance assessments, portfolio assessments, or internships). Because not all of the students have learning outcomes in both categories, the cases were pairwise (correlation analyses) or listwise deleted (regression analyses). For the knowledge tests, this resulted in 170 respondents; for the other assessments, this resulted in 192 respondents.

2.4. Data analysis procedure

Preliminary assumptions were checked. A Shapiro-Wilk test revealed that the data was not normally distributed (all six of the SPAQQ scales have a $p < .001$; for the three scales of the ALSI used in the study, the deep learning approach scale has a $p = .005$, the surface learning approach has a $p = .007$, and the strategic learning approach has a $p < .001$). However, based on the central limit theorem, it was expected that the sampling distribution would approximate normal because the sample size was large enough ($n = 204$) (Field, 2013). Therefore, parametric tests were performed. A total of 15 outliers were identified. Although these outliers might influence the power of significance tests, they were not removed because they represent real scores.

Firstly, correlation analyses were performed in RStudio (RStudio Team, 2016) to determine which variables had to be entered as predictors in the multiple regression analyses. Pearson's Product-Moment correlation coefficient was applied to determine the strengths and direction of the relationships between the average scores of the six variables in the SPAQQ and the average scores of the three ALSI scales and the two categories of students' learning outcomes. Confidence intervals were calculated to provide an estimation of the value of the Pearson Product-Moment correlation coefficient in the population. Effect sizes were reported (J. Cohen, 1988). Secondly, bivariate regression analyses were performed in IBM SPSS Statistics 24 (IBM Corp.,

2016) on the students' overall perceptions of assessment quality (mean score of all six scales of the questionnaire SPAQQ), and the criterion variable learning approaches (all three scales of the ALSI separately), and the criterion variable learning outcomes (learning outcomes in general, and learning outcomes of knowledge tests and other assessments separately). Thirdly, it was then determined which set of predictor variables (the six variables of the SPAQQ) provides the best prediction of the score on the criterion variable of learning approaches (all three scales of the ALSI separately) and the criterion variable of learning outcomes (learning outcomes, in general, and of knowledge tests and other assessments separately). Confidence intervals were calculated to provide an estimation of the value of the regression coefficient in the population. Multiple testing was performed in this study. Therefore *p*-values (*p*) were adjusted (p^{BH}) using the Benjamini and Hochberg False Discovery Rate (FDR in RStudio) (Benjamini & Hochberg, 1995). The consequence of using this approach, is an increased risk of making type II errors (Feise, 2002). Since this is an explorative study of the relation of six variables associated with students' perceptions of assessment quality with their learning approaches and learning outcomes, both *p*-values will be presented in the next section and the implications will be discussed.

3. Results

3.1. Pre-analyses

Correlation analyses were performed to determine which variables had to be entered as predictors in the multiple regression analyses. All the variables of the students' perceptions of assessment quality were found to have significant but (very) weak correlations with at least one type of learning approach; they were positive for the deep learning and strategic learning approaches and negative for the surface learning approach (Table 4). The variables of the students' perceptions of assessment quality had a significant but (very) weak positive correlation with the students' learning outcomes (Table 5). A medium effect size ($r = > .30$) was found for the variable 'effects of assessment on learning' in relation to students' deep learning approach. For all other significant variables, a small effect size was found ($r = > .10$). Based on the correlation analyses, the following six variables were analysed using multiple regression in IBM SPSS Statistics 24: 1) effects of assessment on learning, 2) fairness of assessment, 3) conditions of assessment, 4) interpretation of test scores, 5) authenticity of assessment, and 6) credibility of assessment.

3.2. The relation of students' perceptions of assessment quality and their learning approaches

The first research question sought to determine the relation between students' perceptions of assessment quality, including the six variables, and their learning approaches.

The following results were found for the deep learning approach. Firstly, the model with the students' overall perceptions of assessment quality ($F(1, 202) = 21.22$, $p < .001$) explains 9.1% of the variance in the students' deep learning approach (adjusted $R^2 = .091$), $b = .19$, 95% CI [.11, .27]. This relationship appears to be positive. Secondly, when all six variables associated with the construct of students' perceptions of assessment quality were analysed using multiple regression, a significant model emerged: $F(6, 197) = 4.33$, $p < .001$. The model (Table 6) explains 9.0% of the variance in the students' deep learning approach (adjusted $R^2 = .090$). There appears to be only one significant predictor (also after *p*-value adjustment) that has a positive relationship to the students' deep learning approach: the students' perceptions of the effects of assessment on learning. Thirdly, the model that only tested the effects of assessment on learning ($F(1, 202) = 25.79$, $p < .001$) explains 10.9% of the variance in the students' deep learning approach (adjusted $R^2 = .109$), $b = .18$, 95% CI [.11, .25]. This relationship appears to be positive.

The following results were found for the surface learning approach.

Table 4
Correlations between the Students' Perceptions of Assessment Quality and their Learning Approaches.

Variable	ALSI: Deep learning approach ($n = 204$)				ALSI: Surface learning approach ($n = 204$)				ALSI: Strategic learning approach ($n = 204$)			
	r	*p*	p^{BH}	95% CI	*r*	*p*	p^{BH}	95% CI	*r*	*p*	p^{BH}	95% CI
SPAQQ: Effects of assessment on learning	.34	< .001	.000	[.21, .45]	−.24	.001	.004	[-.37, -.11]	.23	.002	.006	[.10, .36]
SPAQQ: Fairness of assessment	.20	.007	.014	[.07, .33]	−.10	.206	.232	[-.23, .04]	.07	.356	.356	[-.07, .21]
SPAQQ: Conditions of assessment	.24	.001	.004	[.11, .37]	−.18	.020	.033	[-.31, -.04]	.15	.047	.065	[.01, .28]
SPAQQ: Interpretation of test scores	.22	.004	.010	[.09, .35]	−.09	.255	.270	[-.22, .05]	.20	.007	.014	[.07, .33]
SPAQQ: Authenticity of assessment	.24	.001	.004	[.11, .37]	−.11	.138	.177	[-.25, .02]	.19	.011	.020	[.06, .32]
SPAQQ: Credibility of assessment	.24	.001	.004	[.11, .37]	−.16	.034	.051	[-.29, -.02]	.10	.178	.214	[-.03, .24]

Note. p-values are two-tailed. p^{BH}: Benjamini-Hochberg adjusted p-values.

Table 5
Correlations between the Students' Perceptions of Assessment Quality and their Learning Outcomes.

Variable	Learning outcomes general ($n = 204$)				Learning outcomes knowledge tests ($n = 170$)				Learning outcomes other assessments ($n = 192$)			
	r	*p*	p^{BH}	95% CI	*r*	*p*	p^{BH}	95% CI	*r*	*p*	p^{BH}	95% CI
SPAQQ: Effects of assessment on learning	.24	.001	.005	[.10, .36]	.06	.489	.677	[-.09, .21]	.18	.020	.045	[.04, .32]
SPAQQ: Fairness of assessment	.24	.001	.005	[.11, .37]	.00	.974	.974	[-.15, .15]	.23	.004	.014	[.09, .36]
SPAQQ: Conditions of assessment	.29	< .001	.000	[.16, .41]	.04	.703	.844	[-.12, .18]	.28	< .001	.000	[.14, .40]
SPAQQ: Interpretation of test scores	.19	.010	.026	[.06, .32]	.05	.581	.747	[-.10, .20]	.13	.091	.136	[-.01, .27]
SPAQQ: Authenticity of assessment	.16	.038	.068	[.02, .29]	−.03	.779	.876	[-.18, .12]	.14	.066	.108	[.00, .28]
SPAQQ: Credibility of assessment	.20	.007	.021	[.07, .33]	−.02	.860	.911	[-.17, .13]	.17	.031	.062	[.03, .30]

Note. *p*-values are two-tailed. p^{BH}: Benjamini-Hochberg adjusted *p*-values.

Table 6
Relations between the Students' Perceptions of Assessment Quality and their Deep Learning Approach.

Predictor	*b*	SE B	β	*t*	*p*	p^{BH}	95% CI
1. Effects of assessment on learning	.19	.07	.35	2.83	.005	.030	[.06, .32]
2. Fairness of assessment	.02	.05	.04	.38	.708	.953	[-.08, .12]
3. Conditions of assessment	.00	.06	.01	.06	.953	.953	[-.11, .12]
4. Interpretation of test scores	−.01	.05	−.02	−.26	.798	.953	[-.11, .09]
5. Authenticity of assessment	.02	.05	.05	.49	.625	.953	[-.07, .12]
6. Credibility of assessment	−.03	.06	−.07	−.58	.560	.953	[-.14, .08]

Note. *p*-values are two-tailed. p^{BH}: Benjamini-Hochberg adjusted *p*-values.

Firstly, the model with the students' overall perceptions of assessment quality ($F(1, 202) = 7.90$, $p = .005$) explains 3.3% of the variance in the students' surface learning approach (adjusted $R^2 = .033$), $b = -.18$, 95% CI [-.31, -.06]. This relationship appears to be negative. Secondly, when all six variables associated with the construct of students' perceptions of assessment quality were analysed using multiple regression, a significant model emerged: $F(6, 197) = 2.62$, $p = .018$. This model (Table 7) explains 4.6% of the variance in the students' surface learning approach (adjusted $R^2 = .046$). The students' perceptions of the effects of assessment on learning were the only significant predictor (also after *p*-value adjustment). Thirdly, the model that only tested the effects of assessment on learning ($F(1, 202) = 12.57$, $p < .001$) explains 5.4% of the variance in the students' surface learning approach (adjusted $R^2 = .054$), $b = -.21$, 95% CI [-.32, -.09]. This relationship appears to be negative.

The following results were found for the strategic learning approach. Firstly, the model with the students' overall perceptions of assessment quality ($F(1, 202) = 7.88$, $p = .005$) explains 3.3% of the variance in the students' strategic learning approach (adjusted $R^2 = .033$), $b = .17$, 95% CI [.05, .30]. This relationship appears to be positive. Secondly, when all six variables associated with the construct of students' perceptions of assessment quality were analysed using multiple regression, a significant model emerged: $F(6, 197) = 2.87$, $p = .011$. The model (Table 8) explains 5.2% of the variance in the students' strategic learning approach (adjusted $R^2 = .052$). In line with the deep learning approach and the surface learning approach, there appears to be only one significant predictor with a relationship to the students' strategic learning approach: the students' perceptions of the effects of assessment on learning. After *p*-value adjustment (Table 8), the students' perceptions of the effects of assessment on learning were no longer a significant predictor of the students' strategic learning approach. Thirdly, the model that only tested the effects of assessment on learning ($F(1, 202) = 11.44$, $p < .001$) explains 4.9% of the variance in the students' strategic learning approach (adjusted $R^2 = .049$), $b = .18$, 95% CI [.08, .29].

Thus, in response to the first research question, the study results show that the students' perceptions of assessment quality have a significant relation with their learning approaches. However, only one significant predictor was found: the students' perceptions of the effects of assessment on learning.

3.3. The relation of students' perceptions of assessment quality and their learning outcomes

The second research question sought to determine the relation between students' perceptions of assessment quality, including the six variables, and their learning outcomes.

The following results were found for learning outcomes. Firstly, the model with the students' overall perceptions of assessment quality ($F(1, 202) = 15.89$, $p < .001$) explains 6.8% of the variance in the students' learning outcomes (adjusted $R^2 = .068$), $b = .33$, 95% CI [.17, .49]. This relationship appears to be positive. Secondly, when all six variables associated with the construct of students' perceptions of assessment quality were analysed using multiple regression, a significant model emerged: $F(6, 197) = 3.52$, $p = .002$. The model (Table 9) explains 6.9% of the variance in the students' learning outcomes (adjusted $R^2 = .069$). No significant predictors were found.

To obtain more detailed insight into the extent to which the students' perceptions of assessment quality are related to their learning outcomes, the model was also separately tested for students' learning outcomes of knowledge tests and for students' learning outcomes of other assessments. The following results were found for the learning outcomes of the knowledge tests ($n = 170$). Firstly, the model with the students' overall perceptions of assessment quality: $F(1, 168) = .11$, $p = .739$, was nonsignificant. Secondly, when all six variables associated with the construct of students' perceptions of assessment quality were analysed using multiple regression, a nonsignificant model emerged (Table 10): $F(6, 163) = .57$, $p = .750$.

The following results were found for the learning outcomes of the assessments ($n = 192$). Firstly, the model with the students' overall perceptions of assessment quality ($F(1, 190) = 11.34$, $p < .001$) explains 5.1% of the variance in the students' learning outcomes of the assessments (adjusted $R^2 = .051$), $b = .41$, 95% CI [.17, .65]. This relationship appears to be positive. Secondly, when all six variables associated with the construct of students' perceptions of assessment quality were analysed using multiple regression, a significant model emerged: $F(6, 185) = 2.89$, $p = .010$. The model (Table 11) explains 5.6% of the variance in the students' learning outcomes of the assessments (adjusted $R^2 = .056$). The students' perceptions of the conditions of assessment is the only significant predictor ($p = .019$) of the students' learning outcomes of the assessments. The other predictors were nonsignificant. After *p*-value adjustment (Table 11), the students' perceptions of the conditions of

Table 7
Relations between the Students' Perceptions of Assessment Quality and their Surface Learning Approach.

Predictor	*b*	SE B	β	*t*	*p*	p^{BH}	95% CI
1. Effects of assessment on learning	−.30	.11	−.35	−2.76	.006	.036	[-.51, -.09]
2. Fairness of assessment	.03	.08	.04	.39	.694	.833	[-.13, .19]
3. Conditions of assessment	−.06	.09	−.07	−.67	.506	.833	[-.24, .12]
4. Interpretation of test scores	.12	.08	.14	1.47	.144	.432	[-.04, .28]
5. Authenticity of assessment	.04	.08	.04	.46	.648	.833	[-.12, .20]
6. Credibility of assessment	.01	.09	.02	.13	.898	.898	[-.16, .19]

Note. *p*-values are two-tailed. p^{BH}: Benjamini-Hochberg adjusted *p*-values.

Table 8
Relations between the Students' Perceptions of Assessment Quality and their Strategic Learning Approach.

Predictor	*b*	SE B	β	*t*	*p*	p^{BH}	95% CI
1. Effects of assessment on learning	.21	.10	.26	2.06	.041	.246	[.01, .41]
2. Fairness of assessment	−.05	.07	−.07	−.67	.506	.607	[-.20, .10]
3. Conditions of assessment	.01	.09	.01	.10	.917	.917	[-.16, .18]
4. Interpretation of test scores	.08	.08	.10	1.02	.307	.460	[-.07, .23]
5. Authenticity of assessment	.09	.08	.12	1.18	.238	.460	[-.06, .24]
6. Credibility of assessment	−.13	.08	−.19	−1.52	.130	.390	[-.29, .04]

Note. *p*-values are two-tailed. p^{BH}: Benjamini-Hochberg adjusted *p*-values.

assessment were no longer a significant predictor of the students' learning outcomes of the assessments. Thirdly, the model that only tested the conditions of assessment ($F(1, 190) = 15.80$, $p < .001$) explains 7.2% of the variance in the students' learning outcomes of the assessments (adjusted $R^2 = .072$), $b = -.40$, 95% CI [.20, .60].

Thus, in response to the second research question, the results showed that the students' perceptions of assessment quality have a significant positive relation with their learning outcomes for the assessments. Only one significant predictor was found: the students' perceptions of the conditions of assessment.

4. Discussion

This study aimed to provide more insight in the relation between students' perceptions of assessment quality and their learning approaches and learning outcomes. Six variables associated with the construct of students' perceptions of assessment quality were investigated: 1) effects of assessment on learning, 2) fairness of assessment, 3) conditions of assessment, 4) interpretation of test scores, 5) authenticity of assessment, and 6) credibility of assessment.

Regarding the first research question, the study's results showed that the students' overall perceptions are related to their learning approaches. The students' positive overall perceptions are related to a deeper and a more strategic learning approach, and their negative overall perceptions are related to a more surface learning approach. This is due to one variable: the effects of assessment on learning; if students perceive assessment as challenging and motivating and see its added value, they are more likely to deepen their learning approach. This finding supports the need to provide students with explicit information about the objectives and merits of assessment. Moreover, the study's results are in line with previous research conducted by Segers et al. (2008) that found negative correlations between a surface learning approach and students' perceptions of the quality of feedback, the motivating and stimulating nature of the portfolio assessments for their learning, and the transparency of the requirements of the portfolio assessments. The students' perceptions explain only a small part of the variance in their learning approaches. This is not remarkable; previous studies have already showed several influences on students' learning approaches (e.g. the type of assessment, teaching methods, and classroom environment) (Biggs, 1993; Struyven, Dochy, & Janssens, 2005).

Regarding the second research question, the study's results showed that the students' overall perceptions of assessment quality are positively related to the learning outcomes of the assessments. This is due to one variable: the conditions of assessment. An explanation for this result might be that some of the conditions of assessment are sources of measurement error (for example, disturbing external factors). Measurement errors influence the students' assessment scores (Reynolds, Livingston, & Willson, 2010). This raises the question of whether the conditions of assessment differ from the students' perceptions of those conditions. Therefore, further research is needed to investigate the correlation between students' perceptions of assessment quality and the objective evaluation of assessment quality. As with the learning approaches results, the students' perceptions explain only a small part of the variance in their learning outcomes. This is also to be expected; previous studies have already showed several influences on students' learning outcomes (e.g. students' genetics) (Krapohl et al., 2014).

None of the other four variables associated with the construct of students' perceptions of assessment quality – fairness of assessment, interpretation of test scores, authenticity of assessment, and credibility of assessment – are related to the students' learning approaches or outcomes. However, previous research has shown that students form some perceptions about these characteristics of evaluation practices (e.g. Holmes & Smith, 2003; Sambell, McDowell, & Brown, 1997; Tata, 1999). As Dijksterhuis and Bargh (2001) noted: 'Humans are flexible and they can override direct effects of perception on behavior (p. 28)'. One explanation for this could be that the students might not consider some of the assessment quality criteria to be important; therefore, their perceptions do not lead to behaviour. Previous research has shown that not all assessment quality criteria are equally important for ensuring students' satisfaction (Gerritsen-van Leeuwenkamp & Joosten-ten Brinke, 2017). In addition, differences in the importance of assessment quality criteria can be found for different groups of students. Gulikers et al. (2008) found differences in students' perceptions of authenticity based on their experience with professional practice. Lizzio and Wilson (2004) found that students' perceptions of their capabilities varied based on their expectations. It is unclear why students do not perceive the variables of their perceptions of assessment quality to be equally important. Perhaps they cannot relate to some of the variables, or perhaps some of the variables are outside the zone of impact for the students (i.e. teaching adjustments based on test results or the professionalism of teachers). To obtain a deeper understanding of the

Table 9
Relations between the Students' Perceptions of Assessment Quality and their Learning Outcomes in General.

Predictor	*b*	SE B	β	*t*	*p*	p^{BH}	95% CI
1. Effects of assessment on learning	.13	.14	.12	.93	.355	.491	[-.14, .40]
2. Fairness of assessment	.15	.10	.15	1.49	.138	.414	[-.05, .35]
3. Conditions of assessment	.23	.12	.22	1.97	.050	.300	[.00, .46]
4. Interpretation of test scores	.02	.10	.02	.23	.817	.817	[-.18, .22]
5. Authenticity of assessment	−.08	.10	−.08	−.83	.409	.491	[-.29, .12]
6. Credibility of assessment	−.10	.11	−.11	−.87	.386	.491	[-.32, .13]

Note. *p*-values are two-tailed. p^{BH}: Benjamini-Hochberg adjusted *p*-values.

Table 10
Relations between the Students' Perceptions of Assessment Quality and their Learning Outcomes of the Knowledge Tests.

Predictor	*b*	SE B	β	*t*	*p*	p^{BH}	95% CI
1. Effects of assessment on learning	.18	.14	.19	1.29	.200	.624	[-.10, .45]
2. Fairness of assessment	.02	.11	.03	.22	.828	.828	[-.20, .24]
3. Conditions of assessment	.03	.12	.04	.30	.765	.828	[-.19, .26]
4. Interpretation of test scores	.05	.10	.05	.45	.654	.828	[-.16, .25]
5. Authenticity of assessment	−.11	.11	−.12	−1.01	.312	.624	[-.32, .10]
6. Credibility of assessment	−.13	.12	−.16	−1.08	.281	.624	[-.38, .11]

Note. *p*-values are two-tailed. p^{BH}: Benjamini-Hochberg adjusted *p*-values.

Table 11
Relations between the Students' Perceptions of Assessment Quality and their Learning Outcomes of the Other Assessments.

Predictor	*b*	SE B	β	*t*	*p*	p^{BH}	95% CI
1. Effects of assessment on learning	.05	.18	.03	.25	.804	.833	[-.32, .41]
2. Fairness of assessment	.17	.13	.13	1.29	.199	.597	[-.09, .44]
3. Conditions of assessment	.38	.16	.26	2.37	.019	.114	[.06, .69]
4. Interpretation of test scores	−.03	.14	−.02	−.21	.833	.833	[-.30, .24]
5. Authenticity of assessment	−.04	.14	−.03	−.28	.784	.833	[-.31, .23]
6. Credibility of assessment	−.11	.15	−.08	−.71	.481	.833	[-.41, .19]

Note. *p*-values are two-tailed. p^{BH}: Benjamini-Hochberg adjusted *p*-values.

differences in students' perceptions of assessment quality criteria, further research should focus on how students form perceptions of assessment quality based on, for example, their expectations and previous experiences.

Another explanation for why the other four variables associated with the construct of students' perceptions of assessment quality are not related to their learning approaches might be that a change in behaviour does not support the students' own goals (Dijksterhuis & Bargh, 2001). When students have already reached their own goals, for example passing tests, there is no merit in changing their behaviour afterwards based on their perceptions of, for example, the fairness of assessment or the interpretation of test scores. Similar results were found regarding the use of feedback. If students feel that feedback is no longer useful or is irrelevant (it is too late), they will not pay attention to it, so it will not influence subsequent learning (Gibbs & Simpson, 2004). Further research is required to investigate the role of achievement goals as a moderator between students' perceptions of assessment quality and their learning, because achievement goals impact how students favour and prioritise their goals.

The methodology used in this study has some limitations. One limitation is that none of the variables were manipulated. Further experimental research with a larger sample size is required to determine whether these relationships are causal. Such additional research could also focus on investigating a structural model with latent variables in which all the relevant (bidirectional) relationships between the variables of this study are further explored (Hair, Black, Babin, & Anderson, 2014). For example, reverse relationships between students' study behaviour and their perceptions (Richardson, 2006) or between their learning outcomes and their perceptions. Another limitation is the generalisability of the results. The data were all collected from three different bachelor's degree programmes at one university of applied sciences in the Netherlands, each of which conform to the same assessment policy framework. The students were asked to give their perceptions of assessment quality of their year of study, and the same definition of assessment quality was provided to all the students. Due to differences in the assessment programmes' number of (types of) assessment, there is some variance in the sample. These extra variables could not be included into the analyses due to the sample-size in relation to the complexity of the regression model. Therefore, additional studies could focus on analysing whether there are differences between subgroups (e.g. age, gender, and discipline). Furthermore, the students involved were primarily in higher health care education programmes. Therefore, in line with the population of higher health care students in the Netherlands (CBS, 2017), more females than males were represented in the sample. Samples of students from other universities, disciplines, and other countries are necessary to generalise the findings of this study to other contexts. Furthermore, the present study focused on students' perceptions of assessment quality in general; further validation could focus on using the questionnaires in more specified contexts, i.e. by using the SPAQQ and ALSI in one specific assessment programme.

Several implications arise from the findings. This study found some relationships between the students' perceptions of assessment quality and their learning approaches and outcomes. The perceptions of the effects of assessment on learning and of the conditions of assessment are especially important. This implies for educational practice, that if teachers or coaches can generate more positive perceptions of students on these aspects in their evaluation practices, the learning approach may become deeper and more strategic, and better learning outcomes could be realised. This can be done by, for example, more explicitly discuss manners in which assessment can have a positive effect on students' learning or students' motivation to learn or on the way assessment prepares students for future learning activities. Besides, teachers can improve the actual assessment quality by spreading tests over time and careful test construction.

Furthermore, educational organisations should deliberately consider at least two variables in their efforts towards attaining assessment quality: students' perceptions of the 'effects of assessment on learning' and of the 'conditions of assessment'. Although students' perspectives of assessment quality may not always be perfect reflections of reality, this does not have to be a problem; it means that students' perspectives regarding these two variables must be compared with the 'objective' quality in an assessment practice (Gerritsen-van Leeuwenkamp, 2019). However, because there are multiple interpretations of the concept of assessment quality in theory and practice, this is difficult (Gerritsen-van Leeuwenkamp et al., 2017). Therefore, triangulation, in which evidence is collected from different sources, data, and methods of data collection (Creswell, 2008), should be used to judge assessment quality. This

process requires a combination of all the collected quantitative and qualitative data, for example, of conversations or surveys with relevant stakeholders. Then an explicit dialogue, discussion, and negotiation should take place with all the relevant stakeholders (i.e. students, staff, experts) to reach a consensus about what assessment quality means and how it could be improved to enhance students' learning (Gerritsen-van Leeuwenkamp, 2019).

This study showed that not all variables associated with the construct of students' perceptions of assessment quality are related to their learning. Furthermore, the variables that are related to students' learning explain only a small part of the variance. This does not imply that students' perceptions of these variables are less interesting or valuable for educational organisations. Incorporating students' perceptions of all the components of assessment quality remains essential for the quality assurance, monitoring, and improvement of assessment. Students still have the right to participate and voice their opinions (Levin, 1998; Svensson & Wood, 2007). For example, students can be encouraged to participate in the evaluations, serve on advisory boards, or be co-designers in development groups (Healey, Flint, & Harrington, 2014). This would enhance the students' ownership in evaluation practices.

In summary, the results of this study provide a foundation for further (experimental) research to contribute more insight into students' perceptions of assessment quality and how students' perceptions relate to 'objective quality'. It encourages educational organisations to deliberately take students' perspectives into consideration, in order to improve and guarantee the overall assessment quality in higher education.

Appendix A

SPAQQ

In Table A1 the SPAQQ is given.

Table A1
SPAQQ (Gerritsen-van Leeuwenkamp et al., 2018).

Factor	Code item	Number item	Item SPAQQ: In general, at this moment I perceive:
1. Effects of assessment on learning	V03	34	testing and assessment have a positive effect on my learning.
11 items	V07	33	testing and assessment add value to the time I have spent on the work done.
	V10	37	testing and assessment are valuable instances of learning in their own right.
	T05	38	testing and assessment motivate me to continue learning.
	T08	39	testing and assessment help me to navigate my own learning process.
	T09	40	testing and assessment are geared towards the retention of my competencies in the longer run.
	T10	41	testing and assessment prepare me well for future learning activities.
	T11	36	testing and assessment give me the confidence to continue learning.
	O02	13	the tests are challenging.
	O03	30	when I get feedback on tests it shows clearly what I have not yet mastered.
	O04	31	when I get feedback on tests it shows clearly what I have already mastered.
2. Fairness of assessment	V02	1	the tests correspond with the learning targets.
5 items	T01	5	testing and assessment are the same for all students in my year.
	T02	7	testing and assessment are fair.
	T17	12	testing and assessment can be done in the time given.
	O01	4	the difficulty of testing and assessment concur with the level of my education.
3. Conditions of assessment	V11	19	the tests and assessments are organised well.
8 items	T18	21	tests have been spread out evenly during the periods set for testing in the year of study.
	R04	32	when I get feedback on my tests, I will receive it in time.
	R05	20	the team of teachers in my educational programme are accomplished in testing and assessment.
	R06	15	all tests feature correct language.
	R07	18	during testing and assessments there are no disturbing external factors, such as fraudulent behaviour.
	R08	24	whether I pass or fail is based correctly on the score of a test I have taken.
	R09	16	tests have been constructed with care.
4. Interpretation of test scores	V05	25	my scores on tests reflect the extent to which I have mastered the subject.
4 items	V06	26	my scores on various tests on the same topic are comparable.
	V08	27	I would score the same for a test if different questions or tasks about the same subject were presented to me.
	R03	28	I would get more or less the same score on a test if I took the test for a second time (supposing my understanding of the subject matter has remained the same).
5. Authenticity of assessment	T06	2	testing and assessment correspond with the activities I will have to perform in my future occupation.
5 items	T07	8	I understand testing and assessment.
	T13	9	the circumstances in which I am tested or assessed are similar to the working conditions of my future profession.
	T14	10	testing and assessment unveil my thinking processes, for instance when I am asked to underpin certain choices.
	T16	11	I need the competences I require to pass my tests in other (professional) situations as well.
6. Credibility of assessment	V01	17	I agree with the manner in which I am examined.
6 items	V04	35	the teachers use the results of the tests and assessments to adjust the teaching.
	R01	22	judgements are made independently of the persons who rate me.
	R02	23	assessments are made independently of the situations I am assessed in.
	T03	42	I trust testing and assessment in my educational programme to be of good quality.
	T04	43	I get actively involved in testing and assessment in my educational programme.

Note. The items were originally presented in Dutch, and they have been translated into English.

Appendix B

ALSI

In Table B1 the ALSI is given.

Table B1
ALSI (Entwistle & McCune, 2004; Entwistle et al., 2002; ETL project, n.d.; Mattick et al., 2004).

Factor	Code item	Item
1. Deep learning approach (DA)	DA01	I have usually set out to understand for myself the meaning of what we had to learn.
6 items	DA02	In making sense of new ideas, I have often related them to practical or real-life contexts.
	DA03	Ideas I have come across in my academic reading often set me off on long chains of thought.
	DA04	I have looked at evidence carefully to reach my own conclusion about what I am studying.
	DA05	It has been important for me to follow the argument, or to see the reasons behind things.
	DA06	In reading for this course unit, I have tried to find out for myself exactly what the author means.
2. Surface learning approach (SA)	SA01	I have often had trouble in making sense of the things I have to remember.
4 items	SA02	Much of what I have learned seems no more than lots of unrelated bits and pieces in my mind.
	SA03	I have tended to take what we have been taught at face value without questioning it much.
	SA04	I have just been going through the motions of studying without seeing where I am going.
3. Monitoring Studying (MS) 4 items	MS01	I have been over the work I have done to check my reasoning and see that it makes sense.
	MS02	When I have been communicating ideas, I have thought over how well I have got my points across.
	MS03	I have tried to find better ways of tracking down relevant information in this subject.
	MS04	If I have not understood things well enough when studying, I have tried a different approach.
4. Strategic learning approach (effort management/	EM01	I have generally put a lot of effort into my studying.
organised studying; EMOS)	EM02	Concentration has not usually been a problem for me, unless I have been really tired.
4 items	OS01	On the whole, I have been quite systematic and organised in my studying.
	OS02	I have organised my study time carefully to make the best use of it.

Note. The items were originally presented in Dutch.

References

ARG (2002). *Assessment for learning: 10 principles.* Retrieved from https://www.aaia.org.uk/content/uploads/2010/06/Assessment-for-Learning-10-principles.pdf.

Asikainen, H., Parpala, A., Lindblom-Ylänne, S., Vanthournout, G., & Coertjens, L. (2014). The development of approaches to learning and perceptions of the teaching-learning environment during bachelor level studies and their relation to study success. *Higher Education Studies, 4*(4), 24–36. https://doi.org/10.5539/hes.v4n4p24.

Asikainen, H., Parpala, A., Virtanen, V., & Lindblom-Ylänne, S. (2013). The relationship between student learning process, study success and the nature of assessment: A qualitative study. *Studies in Educational Evaluation, 39*, 211–217. https://doi.org/10.1016/j.stueduc.2013.10.008.

Baartman, L. K. J., Bastiaens, T. J., Kirschner, P. A., & Van der Vleuten, C. P. M. (2006). The wheel of competency assessment: Presenting quality criteria for competency assessment programs. *Studies in Educational Evaluation, 32*, 153–170. https://doi.org/10.1016/j.stueduc.2006.04.006.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological), 57*, 289–300 Retrieved from https://www.jstor.org/stable/2346101.

Biggs, J. (1993). What do inventories of students' learning processes really measure? A theoretical review and clarification. *The British Journal of Educational Psychology, 63*, 3–19. https://doi.org/10.1111/j.2044-8279.1993.tb01038.x.

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*, 347–364. https://doi.org/10.1007/BF00138871.

Birenbaum, M., Breuer, K., Cascallar, E., Dochy, F., Dori, Y., Ridgway, J., ... Nickmans, G. (2006). A learning integrated assessment system. *Educational Research Review, 1*, 61–67. https://doi.org/10.1016/j. edurev.2006.01.001.

Bologna Working Group (2005). *A framework for qualifications of the European higher education area.* Retrieved from European Consortium for Accreditation website: http://ecahe.eu/w/index.php/File:Overarching_qualifications_framework_of_EHEA.jpg.

Boud and Associates (2010). *Assessment 2020: Seven propositions for assessment reform in higher education.* Retrieved from University of Technology Sydney website: https://www.uts.edu.au/sites/default/files/Assessment-2020_propositions_final.pdf.

Boud, D. (1990). Assessment and the promotion of academic values. *Studies in Higher Education, 15*, 101–111. https://doi.org/10.1080/03075079012331377621.

Boud, D. (1995). Assessment and learning: Contradictory or complementary? In P. T. Knight (Ed.). *Assessment for learning in higher education* (pp. 35–48). London, England: Kogan.

Brace, N., Kemp, R., & Snelgar, R. (2016). *SPSS for psychologists and everybody else.* New York, NY: Palgrave.

Brown, G. T. L., & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: Links to outcomes. *Assessment in Education: Principles, Policy & Practice, 15*, 3–17. https://doi.org/10.1080/09695940701876003.

Byrne, M., Flood, B., & Willis, P. (2002). The relationship between learning approaches and learning outcomes: A study of Irish accounting students. *Accounting Education, 11*, 27–42. https://doi.org/10.1080/09639280210153254.

CBS (2017). *StatLine: HBO en WO; gediplomeerden, studierichting, herkomst 1995/'96–2015/'16 [StatLine: Higher education; the number of graduates in Nursing, origin 1995/'96–2015/'16].* Retrieved from https://opendata.cbs.nl/statline/#/CBS/nl/dataset/71054ned/table?ts=1532353674257.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* New York, NY: Academic Press.

Cohen, S. A. (1987). Instructional alignment: Searching for a magic bullet. *Educational Researcher, 16*(8), 16–20. https://doi.org/10.3102/0013189X016008016.

Cowie, B. (2009). My teacher and my friends helped me learn. Student perceptions and experiences of classroom assessment. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.). *Student perspectives on assessment: What students can tell us about assessment for learning* (pp. 85–105). Charlotte, NC: Information Age Publishing, INC.

Creswell, J. W. (2008). *Educational research, planning, conducting, and evaluating quantitative and qualitative research.* Upper Saddle River, NJ: Pearson Education, Inc.

De Kleijn, R. A. M., Mainhard, M. T., Meijer, P. C., Pilot, A., & Brekelmans, M. (2012). Master's thesis supervision: Relations between perceptions of the supervisor–Student relationship, final grade, perceived supervisor contribution to learning and student satisfaction. *Studies in Higher Education, 37*, 925–939. https://doi.org/10.1080/03075079.2011.556717.

Dijksterhuis, A., & Bargh, J. A. (2001). The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology, 33*, 1–40. https://doi.org/10.1016/S0065- 2601(01)80003-4.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4–58. https://doi.org/10.1177/1529100612453266.

Entwistle, N. J. (2000, November). Promoting deep learning through teaching and assessment: Conceptual frameworks and educational contexts. *Paper presented at the ESRC Teaching and Learning Research Programme Conference* Retrieved from http://www.leeds.ac.uk/educol/documents/00003220.htm.

Entwistle, N. J., & McCune, V. (2004). The conceptual bases of study strategy inventories. *Educational Psychology Review, 16*, 325–345. https://doi.org/10.1007/s10648-004-0003-0.

Entwistle, N. J., McCune, V., & Hounsell, J. (2002). *Occasional report 1. Approaches to studying and perceptions of university teaching-learning environments: Concepts, measures and preliminary findings* Retrieved from University of Edinburgh, School of Education website: www.etl.tla.ed.ac.uk/docs/ETLreport1.pdf.

Entwistle, N. J., & Ramsden, P. (1982). *Understanding student learning.* New York, NY: Nichols Publishing Company.

ETL project (2002). *Experiences of teaching & learning questionnaire (ETLQ).* Retrieved from http://www.etl.tla.ed.ac.uk/questionnaires/ETLQ.pdf.

ETL project. (n.d.). *Learning and studying questionnaire (LSQ). Experiences of teaching and learning questionnaire (ETLQ).* Retrieved from http://www.etl.tla.ed.ac.uk/questionnaires/scoringkey.pdf.

Feise, R. J. (2002). Do multiple outcome measures require p-value adjustment? *BMC*

Medical Research Methodology, 2(8), https://doi.org/10.1186/1471-2288-2-8.
Field, A. (2013). *Discovering statistics using SPSS.* London, England: SAGE publications.
Gerritsen-van Leeuwenkamp, K. J. (2019). *Assessment quality in higher education: Students' perspectives in relation to their learning* (Doctoral thesis). Open University of the Netherlands, Heerlen, the Netherlands.
Gerritsen-van Leeuwenkamp, K. J., & Joosten-ten Brinke, D. (2017). Increasing student satisfaction by improving assessment quality. *Manuscript submitted for publication.*
Gerritsen-van Leeuwenkamp, K. J., Joosten-ten Brinke, D., & Kester, L. (2017). Assessment quality in tertiary education: An integrative literature review. *Studies in Educational Evaluation, 55*, 94–116. https://doi.org/10.1016/j. stueduc.2017.08.001.
Gerritsen-van Leeuwenkamp, K. J., Joosten-ten Brinke, D., & Kester, L. (2018). Developing questionnaires to measure students' expectations and perceptions of assessment quality. *Cogent Education, 5*(1464425), https://doi.org/10.1080/2331186X.2018.1464425.
Gibbs, G. (2010). Does assessment in open learning support students? *Open Learning, 25*, 163–166. https://doi.org/10.1080/02680511003787495.
Gibbs, G., & Simpson, C. (2004). Conditions under which assessment supports students' learning. *Learning and Teaching in Higher Education, 5*(1), 3–31 Retrieved from http://eprints.glos.ac.uk/id/eprint/3609.
Gijbels, D., Watering, G., Dochy, F., & Bossche, P. (2005). The relationship between students' approaches to learning and the assessment of learning outcomes. *European Journal of Psychology of Education, 20*, 327–341. https://doi.org/10.1007/bf03173560.
Gulikers, J., Kester, L., Kirschner, P. A., & Bastiaens, T. J. (2008). The effect of practical experience on perceptions of assessment authenticity, study approach, and learning outcomes. *Learning and Instruction, 18*, 172–186. https://doi.org/10.1016/j.learninstruc.2007.02.012.
Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis.* Essex, England: Pearson Education Limited.
Healey, M., Flint, A., & Harrington, K. (2014). *Engagement through partnership: Students as partners in learning and teaching in higher education.* Retrieved from Higher Education Academy website: https://www.heacademy.ac.uk/system/files/resources/engagement_through_partnership.pdf.
Heeneman, S., Oudkerk Pool, A., Schuwirth, L. W. T., Van der Vleuten, C. P. M., & Driessen, E. W. (2015). The impact of programmatic assessment on student learning: Theory versus practice. *Medical Education, 49*, 487–498. https://doi.org/10.1111/medu.12645.
Holmes, L. E., & Smith, L. J. (2003). Student evaluations of faculty grading methods. *The Journal of Education for Business, 78*, 318–323. https://doi.org/10.1080/08832320309598620.
IBM Corp (2016). *IBM SPSS statistics for Windows, version 24.0 [Computer software].* Armonk, NY: Author.
Krapohl, E., Rimfeld, K., Shakeshaft, N. G., Trzaskowski, M., McMillan, A., Pingault, J. B., ... Plomin, R. (2014). The high heritability of educational achievement reflects many genetically influenced traits, not just intelligence. *Proceedings of the National Academy of Sciences of the United States of America, 111*, 15273–15278. https://doi.org/10.1073/pnas.1408777111.
Levin, B. (1998). The educational requirement for democracy. *Curriculum Inquiry, 28*, 57–79. https://doi.org/10.1111/0362-6784.00075.
Lindblom-Ylänne, S., & Lonka, K. (1998). Individual ways of interacting with the learning environment — are they related to study success? *Learning and Instruction, 9*, 1–18. https://doi.org/10.1016/S0959-4752(98)00025-5.
Lindblom, S. (2017). Student approaches to learning - a critical analysis in the light of recent empirical evidence. *Paper presented at the EARLI 2017* Retrieved from https://youtu.be/P9NclXaFpts.
Lizzio, A., & Wilson, K. (2004). First-year students' perception of capability. *Studies in Higher Education, 29*, 109–128. https://doi.org/10.1080/1234567032000164903.
Lizzio, A., Wilson, K., & Simons, R. (2002). University students' perceptions of the learning environment and academic outcomes: Implications for theory and practice. *Studies in Higher Education, 27*, 27–52. https://doi.org/10.1080/03075070120099359.
Martin, S. (1997). Two models of educational assessment: A response from initial teacher education: If the cap fits … *Assessment and Evaluation in Higher Education, 22*, 337–343. https://doi.org/10.1080/0260293970220307.
Marton, F., & Säljö, R. (2005). *Approaches to learning.* Retrieved from http://www.docs.hss.ed.ac.uk/iad/Learning_teaching/Academic_teaching/Resources/Experience_of_learning/EoLChapter3.pdf.
Mattick, K., Dennis, I., & Bligh, J. (2004). Approaches to learning and studying in medical students: Validation of a revised inventory and its relation to student characteristics and performance. *Medical Education, 38*, 535–543. https://doi.org/10.1111/j.1365-2929.2004.01836.x.
Mayer, R. E. (2002). Rote versus meaningful learning. *Theory Into Practice, 41*, 226–232. https://doi.org/10.1207/s15430421tip4104_4.
Miller, G. E. (1990). The assessment of clinical skills/competence/performance. *Academic Medicine, 65*, S63–S67. https://doi.org/10.1097/00001888-199009000-00045.
Minbashian, A., Huon, G. F., & Bird, K. D. (2004). Approaches to studying and academic performance in short-essay exams. *Higher Education, 47*, 161–176. https://doi.org/10.1023/B:HIGH.0000016443.43594.d1.
Reynolds, C. R., Livingston, R. B., & Willson, V. (2010). Reliability for teachers. In L. Reinkober (Ed.). *Measurement and assessment in education* (pp. 90–122). Upper Saddle River, NJ: Pearson Education.
Richardson, J. T. E. (2006). Investigating the relationship between variations in students' perceptions of their academic environment and variations in study behaviour in distance education. *The British Journal of Educational Psychology, 76*, 867–893. https://doi.org/10.1348/000709905X69690.
Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences, 15*, 20–27. https://doi.org/10.1016/j.tics.2010.09.003.
RStudio Team (2016). *RStudio: Integrated development for R (1.0.136) [Computer Program].* Boston, MA: Author.
Sambell, K., McDowell, L., & Brown, S. (1997). "But is it fair?": An exploratory study of student perceptions of the consequential validity of assessment. *Studies in Educational Evaluation, 23*, 349–371. https://doi.org/10.1016/S0191-491X(97)86215-3.
Scouller, K. (1998). The influence of assessment method on students' learning approaches: Multiple choice question examination versus assignment essay. *Higher Education, 35*, 453–472. https://doi.org/10.1023/A:1003196224280.
Segers, M., Dochy, F., Gijbels, D., & Struyven, K. (2009). Changing insights in the domain of assessment in higher education: Novel assessments and their pre-, post- and pure effects on student learning. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.). *Student perspectives on assessment: What students can tell us about assessment for learning* (pp. 297–319). Charlotte, NC: Information Age Publishing, INC.
Segers, M., Gijbels, D., & Thurlings, M. (2008). The relationship between students' perceptions of portfolio assessment practice and their approaches to learning. *Educational Studies, 34*, 35–44. https://doi.org/10.1080/03055690701785269.
Stobart, G. (2008). *Testing times. The uses and abuses of assessment.* Oxon, England: Routledge.
Struyven, K., Dochy, F., & Janssens, S. (2005). Students' perceptions about evaluation and assessment in higher education: A review. *Assessment and Evaluation in Higher Education, 30*, 325–341. https://doi.org/10.1080/02602930500099102.
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*, 279–294. https://doi.org/10.1016/j.learninstruc.2006.07.001.
Svensson, G., & Wood, G. (2007). Are university students really customers? When illusion may lead to delusion for all! *International Journal of Educational Management, 21*, 17–28. https://doi.org/10.1108/09513540710716795.
Tata, J. (1999). Grade distributions, grading procedures, and students' evaluations of instructors: A justice perspective. *Journal of Psychology, 133*, 263–271. https://doi.org/10.1080/00223989909599739.
Van der Kleij, F. M., Feskens, R. C. W., & Eggen, T. J. H. M. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes. *Review of Educational Research, 85*, 475–511. https://doi.org/10.3102/0034654314564881.
Van der Vleuten, C. P. M., Sluijsmans, D., & Joosten-ten Brinke, D. (2017). Competence assessment as learner support in education. In M. Mulder (Ed.). *Competence-based vocational and professional education. Bridging the worlds of work and education* (pp. 607–630). Cham, Switzerland: Springer International Publishing Switzerland.
Watkins, D., Dahlin, B., & Ekholm, M. (2005). Awareness of the backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lecturers. *Instructional Science, 33*, 283–309. https://doi.org/10.1007/s11251-005-3002-8.
Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment, 10*, 1–17. https://doi.org/10.1207/s15326977ea1001_1.
Zimbardo, P. G., Weber, A. L., & Johnson, R. L. (2009). Sensation and perception. In S. Frail, J. Swasey, D. Hanlon, & A. Pickard (Eds.). *Psychology core concepts* (pp. 287–333). Boston, MA: Pearson Education.

Karin J. Gerritsen-van Leeuwenkamp is educational advisor at Saxion, University of Applied Sciences.

Desirée Joosten-ten Brinke is an associate professor in testing and assessment at the Welten institute of the Open University of the Netherlands and an associate professor in testing and assessment at Fontys, University of Applied Sciences for teacher trainers.

Liesbeth Kester is full professor Educational Sciences at the Department of Education & Pedagogy, Utrecht University.