

# Developing a digital application for quality assurance of assessment programmes in higher education

Received 8 March 2022  
Revised 2 August 2022  
8 September 2022  
Accepted 25 September 2022

Lonneke H. Schellekens

*Department of Population Health Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands and Department of Education, Faculty of Social Sciences, Utrecht University, Utrecht, The Netherlands*

Marieke F. van der Schaaf

*Utrecht Centre for Research and Development of Health Professions Education, University Medical Centre Utrecht, Utrecht, The Netherlands and Department of Education, Utrecht University, Utrecht, The Netherlands*

Cees P.M. van der Vleuten

*Department of Educational Development and Research, Faculty of Health Medicine and Life Sciences, Maastricht University, Maastricht, The Netherlands*

Frans J. Prins

*Department of Education, Faculty of Social Sciences, Utrecht University, Utrecht, The Netherlands*

Saskia Wools

*Department of Research and Innovation, Cito Institute of Test Development Foundation, Arnhem, The Netherlands, and*

Harold G.J. Bok

*Department of Population Health Sciences, Faculty of Veterinary Medicine, Utrecht University, Utrecht, The Netherlands*

## Abstract

**Purpose** – This study aims to report the design, development and evaluation of a digital quality assurance application aimed at improving and ensuring the quality of assessment programmes in higher education.

**Design/methodology/approach** – The application was developed using a design-based research (DBR) methodology. The application's design was informed by a literature search and needs assessment of quality assurance stakeholders to ensure compliance with daily practices and accreditation requirements. Stakeholders from three study programmes evaluated the application.



**Findings** – As part of the development of the application, module- and programme-level dashboards were created to provide an overview of the programme’s outcomes, assessment methods, assessment metrics, self-evaluated quality indicators and assessment documents. The application was evaluated by stakeholders at the module and programme levels. Overall, the results indicated that the dashboards aided them in gaining insight into the assessment programme and its alignment with underlying assessments.

**Practical implications** – Visualisation of the assessment programme’s structure and content identifies gaps and opportunities for improvement, which can be used to initiate a dialogue and further actions to improve assessment quality.

**Originality/value** – The application developed facilitates a cyclical and transparent assessment quality assurance procedure that is continuously available to various stakeholders in quality assurance.

**Keywords** Quality assurance application, Assessment programme, Improving assessment quality, Ensuring assessment quality, Higher education

**Paper type** Research paper

## Introduction

An assessment programme can be regarded as a combination of various assessment activities that are purposefully designed and tailored to the aims, content and structure of a curriculum (Baartman *et al.*, 2017; Van der Vleuten *et al.*, 2012). The quality of an assessment programme is vital in its ability to optimise student support throughout their learning and development, as well as to validate student diplomas and certification degrees. Quality can be defined in various ways, for example as “exceptional”, “fitness for purpose” or “value for money” (Harvey and Stensaker, 2008, p. 433). In this study, quality is viewed as a “transformation, a process of change that adds value to students through their learning experience” (Harvey and Stensaker, 2008, p. 433). It highlights the possibility of qualitative change, for example by developing and enhancing students’ competencies (Kleijnen *et al.*, 2013). In daily practice, higher education study programmes often encounter difficulties in monitoring and ensuring assessment quality. They frequently struggle to clearly demonstrate alignment, as indicated by the existence of gaps between module and programme outcomes and a lack of confirmation that programme outcomes are efficiently incorporated within the curriculum. Similarly, assessment procedures may lack transparency and can thus be unclear to stakeholders (Hulpiau and Waeytens, 2003; Sridharan *et al.*, 2015). When the quality of an assessment programme cannot be ensured, both the quality of the education and students’ learning processes can be negatively affected (Lucander and Christersson, 2020).

Quality assurance in higher education refers to measures taken to determine whether the educational quality is guaranteed (Sluijsmans *et al.*, 2015) and is defined as “an ongoing, continuous process of evaluating (...) the quality of a higher education system, institutions, or programmes” (Vlăsceanu *et al.*, 2004, p. 74). With regard to assessment, quality assurance generally serves two purposes: improvement and accountability (EUA, 2006; Hulpiau and Waeytens, 2003). Concerning improvement itself, quality assurance is often tailored toward internal procedures in which programme stakeholders share the responsibility of enhancing assessment quality (Dill, 2007). From the viewpoint of accountability, quality assurance can be directed to external procedures in which evidence is provided to an auditing committee to demonstrate that the assessment programme meets established standards (Dill, 2007; Hobson *et al.*, 2008). The overall aim of assessment quality assurance in higher education in Europe is to foster a quality culture in which multiple stakeholders are continuously engaged and encouraged to improve assessment quality to ultimately meet internal and external accreditation requirements (EUA, 2006).

Quality assurance within an assessment programme requires a comprehensive, integrative approach, because it involves a complex evaluation of programme outcomes that

represent both the educational programme's philosophy and the complexity of outcomes that are not typically covered in a 12-week module (Jessop *et al.*, 2012). Furthermore, it is essential that the various assessment practices are structured in such a way as to facilitate continuous student learning (Bok *et al.*, 2013; Van der Vleuten *et al.*, 2012). To ensure the programme's overall quality, multiple stakeholders should be challenged to regularly review both the quality of the individual assessment practices and the alignment between individual assessments throughout the programme (Boud and Falchikov, 2006; Norcini *et al.*, 2018). This requires that stakeholders have an overview of the study programme, and that they are familiar with the procedures and quality indicators used to evaluate assessment quality (Baartman *et al.*, 2017; Sluijsmans *et al.*, 2015).

Various examples of the procedures used to ensure assessment quality can be found in the literature (Baartman *et al.*, 2007, 2011; Lucander and Christersson, 2020; Marciniak, 2018). For example, Baartman *et al.* (2007, 2011) developed and validated a self-evaluation procedure to assess the quality of assessment programmes in competence-based education by evaluating programme-level quality criteria (e.g. authenticity, fairness, transparency, educational consequence, reproducibility). The self-evaluation procedure comprises two phases: an individual Web-questionnaire and a subsequent group interview. The confrontation of participants' opinions in the group interview stimulate discussion about and reflection on assessment quality. The self-evaluation procedure was positively evaluated by its users. However, the procedure mainly focuses on assessment quality at a single point in time and does not provide stakeholders with a clear overview of the assessment programme. Lucander and Christersson (2020) extended the procedure of Baartman *et al.* (2007) by using the same criteria to develop a process for quality assurance of assessment (PQAA) for entire educational programmes. A survey was developed to identify stakeholders' perceptions of the assessments in the programme, which functioned as a starting point for internal analysis and evaluation to develop plans and realise change. In addition to the survey, curriculum maps and assessment plans were created to visualise the assessments in the various modules of the programme. As a result, the PQAA provides a framework for quality improvement by supporting a more continuous procedure and by visualising the assessment programme. However, a possible gap is that the continuous engagement of various stakeholders may be restricted. By using a survey as a starting point, the PQAA is not part of the day-to-day practices of the various stakeholders involved. In addition, the process was directed by a development team that was not embedded within the educational programme itself.

Having a cyclical and transparent assessment quality assurance procedure that is embedded within the educational programme is a prerequisite for structurally improving and ensuring assessment quality (Inspectie van het Onderwijs, 2018). The aim of this study was therefore to design, develop and evaluate an assessment quality assurance application that can be used for both improvement and accountability purposes. To ensure that the application is continuously available to various stakeholder groups, an online web application was developed. The digital application should support stakeholders in conducting a structural review of the assessment programme to improve assessment quality, as well as provide insight into the criteria commonly used to evaluate assessment quality. Additionally, the application should be user-friendly and adhere to stakeholders' standard quality assurance procedures and activities to avoid creating additional work for these stakeholders. The application was evaluated on users' effort and performance expectancy because these theoretical concepts are directly related to stakeholders' intentions to use newly developed systems (Venkatesh *et al.*, 2003). Effort expectancy is defined as "the degree of ease associated with the use of the system", and performance expectancy as "the degree to which the user expects that using the system will help him or her to attain gains in

job performance” (Venkatesh *et al.*, 2003, p. 450, p. 447). The following research questions were formulated to guide the research:

- RQ1a.* Which quality assurance procedures are commonly seen in higher education educational programmes?
- RQ1b.* Which quality criteria are commonly used to evaluate assessment quality?
- RQ2.* What design principles guide the development of a user-friendly digital assessment quality assurance application that complies with improvement and accountability purposes?
- RQ3.* What are the characteristics of a digital quality assurance application for assessment programmes in higher education?
- RQ4.* How do stakeholders perceive the developed application in terms of effort and performance expectancy?

## Methods

### *Research design*

The application was developed using a design-based research (DBR) methodology characterised by a systematic and flexible approach through iterative phases of design, development, evaluation and redesign, based on collaboration among stakeholders (Bakker and Van Eerde, 2015). DBR advances both theory and practice by using literature and theories to guide the application’s design principles, and by evaluating these principles in practice (Dolmans and Tigelaar, 2012; Wang and Hamafin, 2005). The development of the application was directed by a project team from Utrecht University with extensive expertise in assessment and quality assurance. The application was developed in four iterative phases:

Design of the application:

- (1) *Phase 1:* Analysis; Scientific literature search and needs assessment to examine commonly used quality assurance procedures, quality criteria and stakeholders’ current practices (*RQ1*).
- (2) *Phase 2:* Design guidelines; formulation of design principles that guide the development of a user-friendly digital assessment quality assurance application that complies with improvement and accountability purposes (*RQ2*).

Development of the application:

- (3) *Phase 3:* Development; creating electronic dashboards for the quality assurance of the assessment programme and its alignment with individual assessment practices (*RQ3*).

Evaluation of the application:

- (4) *Phase 4:* Evaluation; pilot study to examine perceived stakeholder efforts and performance expectancy (*RQ4*).

The results of the evaluation will be used to redesign the application. This phase is not covered in this study but will be discussed in the conclusion section.

### *Participants*

Two stakeholder groups were distinguished in the DBR-study based on their roles and positions in assessment quality assurance in The Netherlands (Schellekens *et al.*, 2021):

- (1) Educational programme stakeholders:
  - Programme directors or coordinators, who are chiefly responsible for assessment policy, the quality of the assessment programme and the implementation thereof.
  - Board of examiners, an active and independent committee that monitors the utilisation of quality-assurance procedures and the assessment of learning outcomes. The board gives annual reports to the programme director about the activities they have performed.
- (2) Module stakeholders: Teacher(s) that coordinate the module who are responsible for individual assessment practices within modules.

In Phase 1, stakeholder groups participated in a needs assessment to inform the design of the application. Stakeholders were interviewed to gain insight into how quality assurance procedures were performed and implemented, as well as to determine what needs stakeholders have in ensuring quality. In Phase 4, stakeholder groups participated in a pilot study to evaluate the application. During Phases 2 and 3, no stakeholder groups were consulted. In Phase 1, a total of 22 module stakeholders and 14 programme stakeholders from 5 study programmes participated in a needs assessment, as illustrated in [Table 1](#). The study programmes were selected to reflect the range of disciplines offered at Utrecht University and were based upon the inclusion of various types of assessment methods. Programme directors of five study programmes were asked to randomly nominate four or five module stakeholders with varying years of assessment expertise to participate in the needs assessment. Participants were approached via email, informed about the research project and asked if they were both interested in and had the time available to participate (convenience sample). All participants were willing to cooperate.

In Phase 4, The application was piloted and evaluated in three study programmes by both module and programme stakeholders ([Table 1](#)). Module stakeholders were invited if they taught a module between November and May during the academic year 2019–2020. The pilot included 16 module stakeholders (out of a total of 23 invited). A total of 9 of the 11 programme-level stakeholders who were invited to participate in the pilot study did so. Reasons for non-participation were lack of time, illness and the impact of the COVID-19 pandemic that affected higher education in The Netherlands in March 2020.

Faculty	Participant needs assessment (Phase 1)		Participant pilot study (Phase 4)	
	Module stakeholders	Programme stakeholders	Module stakeholders	Programme stakeholders
Faculty of Geosciences		4	7	4
Faculty of Humanities	3			
Faculty of Law, Economics, and Governance	4			
Faculty of Medicine		3		
Faculty of Science	5			
Faculty of Social and Behavioural Sciences	5	2	4	2
Faculty of Veterinary Medicine	5	5	5	3
Total	<i>N</i> = 22	<i>N</i> = 14	<i>N</i> = 16	<i>N</i> = 9

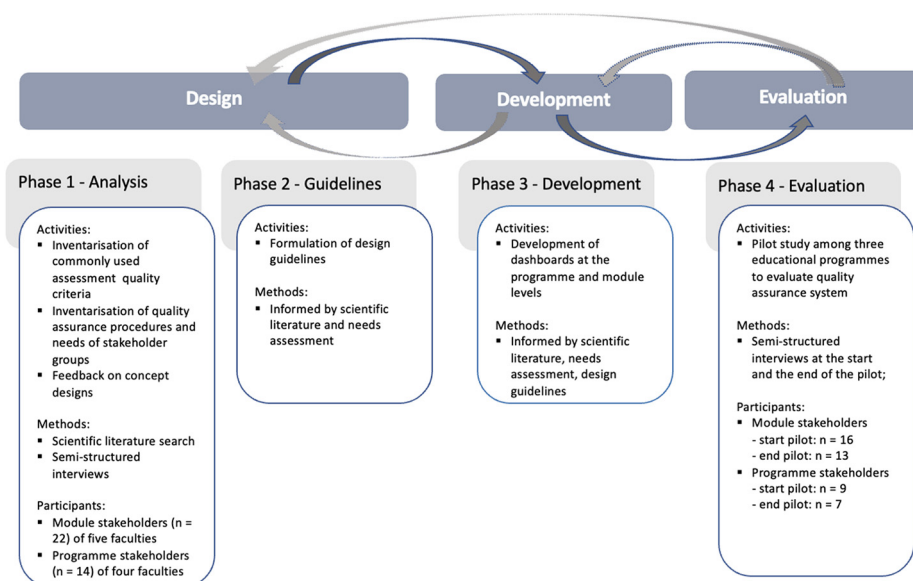
**Table 1.**  
Overview of  
participants in  
Phases 1 and 4

*Procedure*

Figure 1 depicts the study’s research design in chronological order, including the activities performed, methods used and the participants involved in each phase. The arrows in the figure emphasise the study’s iterative nature. The application was developed in collaboration with designers, programmers and assessment experts from Cito Foundation, an institute for educational measurement in The Netherlands, who develop assessments and educational technology for all educational sectors.

In Phase 1, a scientific literature search was conducted to examine commonly used internal and external quality assurance procedures in higher educational programmes, as well as the criteria used to evaluate assessment quality. Furthermore, a needs assessment was carried out among both module and programme stakeholders of various study programmes. Interviews were conducted to gain insight into how quality assurance procedures were performed and implemented, as well as to determine what needs stakeholders have in ensuring quality. Various concept designs were piloted during the interviews, both on paper and in online environments. In multiple sessions, the project team evaluated the findings of the literature search and needs assessment and used them to inform the design of the application. In Phase 2, the design guidelines that directed the development were formulated. In Phase 3, both module- and programme-level dashboards were created, resulting in an assessment quality assurance application that was available online.

In Phase 4, the application was piloted in three study programmes among module and programme stakeholders. For each study programme, a beta version of the application was programmed with authentic assessment information that was delivered by the programme director (e.g. the programme’s curriculum, learning outcomes, assessment methods, etc). Participants in the pilot study were interviewed twice. The first interview was conducted prior to the start of the pilot to inform participants about the study and become acquainted with participants’ knowledge and the methods used to ensure assessment quality. All



**Figure 1.** Research design, activities, methods and participants

participants were given a handout with instructions on how to use the application independently. The second interview with module stakeholders was conducted after they completed their module and entered the required information into the application. The second interview with programme stakeholders was conducted at the end of the pilot study. The purpose of the second interview was to evaluate stakeholders' experiences with the application concerning the perceived effort and performance expectancy. Semi-structured interviews were held during Phases 1 and 4. In Phase 1, interviews were guided by both the first author and a research assistant. In Phase 4, all interviews were conducted by the first author. During Phase 4, six participants were unable to fit the interview into their schedule and instead answered the questions on paper. After reading the information letter, all participants signed a letter of informed consent prior to the start of their interview.

*Analysis*

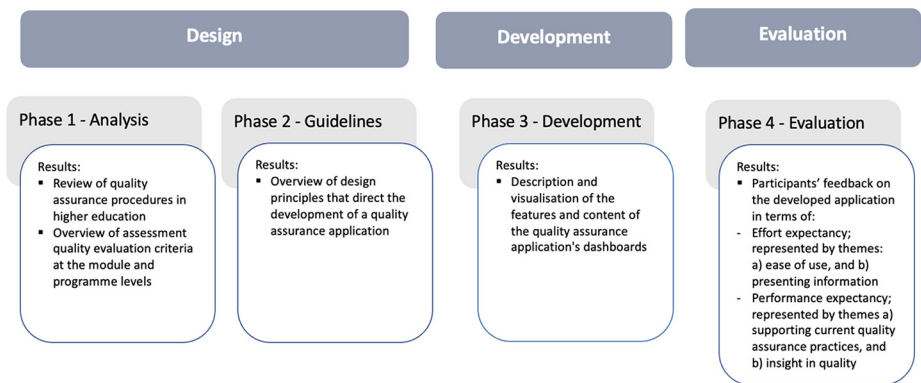
All interviews were recorded on a voice recorder and transcribed verbatim. The first author reread the transcriptions of the interviews and written answers several times to ensure they understood their content and context. All data were imported into NVivo (version 12) and a content analysis was carried out (Cohen *et al.*, 2018). The analysis of the data proceeded in stages, starting with an exploration stage to assign descriptive codes to topics that related to the research questions. Then, in the specification stage, the codes were constantly compared to group them within a structure of categories. Finally, in the reduction stage, the first and last authors examined the codes for coherence on content, frequencies and stakeholder groups to summarise the data (Baarda *et al.*, 2001; Cohen *et al.*, 2018). In an iterative process, the emerging themes of the content analysis were discussed until a consensus was reached.

*Ethical considerations*

Ethical approval was given by the ethical board of the Faculty of Social and Behaviour Sciences of Utrecht University and registered under nr. UU-FETC18-126 (phase 1) and nr. UU-FETC19-022 (phase 4).

**Results**

This section will summarise the research findings in relation to the research questions that emerged from the four phases of the DBR study. Figure 2 illustrates a summary of the results for each phase. First, quality assurance procedures in higher education study



**Figure 2.** Summary of the results for each phase in the research design

programmes will be described, in addition to the criteria commonly used to evaluate assessment quality. Second, the design guidelines that steered the development of the application will be presented. Third, examples of dashboards designed to characterise a quality assurance application for assessment programmes in higher education are provided. Fourth, findings regarding stakeholders' perceptions of the application developed in terms of effort and performance expectancy are presented.

### *Quality assurance procedures in higher education*

*Internal quality assurance procedures.* Within higher education in Europe, it is an explicit aim of accreditation bodies to further develop and improve educational programmes (EUA, 2006). Ideally, internal quality procedures in which programme stakeholders share the responsibility to enhance assessment quality serve as a balance to the requirements of external accountability (Dill, 2007; EUA, 2006). Internal quality assurance procedures are used to review assessment practices and the assessment programme on a regular basis. To enable stakeholders' reflection on current assessment practices, internal quality assurance procedures refer to a method of self-evaluation (Baartman *et al.*, 2007; Hobson *et al.*, 2008; Sluijsmans *et al.*, 2015). Self-evaluation can be an effective approach to both ensure external quality assurance and increase internal quality (Sluijsmans *et al.*, 2015). Moreover, self-evaluation can result in concrete points for improvement and can raise awareness about the quality of the assessments (Dijkstra and Baartman, 2011). Another method to facilitate reflection on assessment practices is the inclusion of a plan, do, check, act (PDCA) procedure (Baartman *et al.*, 2007; Hobson *et al.*, 2008). Within higher education, a PDCA procedure to ensure assessment quality is often organised around the stages of the "assessment cycle" (Sluijsmans *et al.*, 2015; Van Berkel *et al.*, 2017). The assessment cycle prescribes several steps to realise the quality of individual assessment practices within a module. The assessment cycle generally comprises four stages:

- (1) the design stage, for alignment between module learning outcomes, as well as the purpose, content and method of the assessment;
- (2) the administration stage, for the proper administration and assessment of the test;
- (3) the evaluation stage, for analysing and evaluating both the test scores and the test itself; and
- (4) the action stage, for formulating actions for improvement as the final stage to close the PDCA (Bijkerk, 2015).

For each subsequent stage, procedures can be performed to ensure high quality assessment.

*External quality assurance procedures and standards.* Assessment is an integral part of external quality assurance procedures. Although national frameworks for quality assurance vary from country to country, they generally follow the model of an external quality audit process by a committee that assesses the educational programme's performance against governmental requirements (Dill, 2007). For example, in The Netherlands, once every six years the study programme is visited by independent experts who assess the programme's quality according to the level of its discipline (Inspectie van het Onderwijs, 2018). In preparation for this visit, the programme's management board writes a self-evaluation report and presents additional documents. The documents provide insight into how assessment quality is ensured at both the programme level (e.g. assessment policy plan, curriculum evaluations) and the module level (e.g. module assessment plans, rubrics, module evaluations, etc). In relation to assessment quality, the following two standards are assessed



throughout a Dutch accreditation (NVAO, 2016, 2018). Higher education study programmes must provide evidence that the programme:

- has an adequate system of assessment, that is, the individual assessment practices are valid, reliable, the assessment requirements are clear to students and assessments should support students' learning processes; and
- has covered programme learning outcomes, that is, it should be demonstrated at the programme level that the intended learning outcomes have been achieved.

#### *Criteria for evaluating the quality of assessments*

For both internal and external quality assurance purposes, quality criteria are used to evaluate assessment quality. The results of review studies generally distinguish between four assessment quality criteria (Baartman *et al.*, 2007; Gerritsen-van Leeuwenkamp *et al.*, 2017; Maassen *et al.*, 2015; Van der Vleuten and Schuwirth, 2005):

- (1) validity, which refers to whether an assessment (programme) is appropriate for achieving the intended learning outcomes;
- (2) reliability, which refers to whether an assessment (programme) is administered accurately and consistently;
- (3) transparency, which refers to whether the assessment (programme) has clear procedures and criteria for judging performance; and
- (4) educational learning impact, which refers to whether an assessment (programme) optimally supports and enhances students' learning processes and development.

The latter criterion emphasises that the assessment programme is part of students' overall learning process (Van der Vleuten and Schuwirth, 2005). For example, by directing student participation in assessment activities and by thinking systematically about students' development throughout the programme (Boud and Falchikov, 2006). The examples in Table 2 demonstrate how these criteria can be operationalised at the module and programme levels. The examples are drawn from the scientific literature search and current assessment practices of the needs assessment participants.

#### *Design guidelines for the digital assessment quality assurance application*

The guidelines were developed with the goal of creating an application that:

- can be used for improvement, by embedding the application as an integral part of quality assurance procedures;
- can be used for accountability, by providing insight into the quality of the assessment programme and individual assessment practices; and
- is user friendly, by being applicable and easy to use for all users.

Table 3 summarises the design guidelines for each goal.

#### *Characteristics of the assessment quality assurance application developed*

Based on these guidelines, learning dashboards were developed to visualise data in a variety of ways to facilitate both self-monitoring and administrative monitoring (Schwendimann *et al.*, 2016). Learning dashboards are defined as "a single display that aggregates different indicators about learner(s), learning process(es) and/or learning context(s) into one or

Assessment quality criteria	Assessment programme
<p><i>Validity</i></p> <p>There is evidence that the assessment is aligned with learning activities and module learning outcomes, for example by:</p> <ul style="list-style-type: none"> <li>• Presenting an overview of how the module learning outcomes relate to the assessment (i.e., assessment plan)</li> <li>• Presenting an overview of how specific assessment items relate to the content and associated mastery levels (i.e., assessment blueprint)</li> </ul> <p><i>Reliability</i></p> <p>There is evidence that the scoring of the assessment performance is reliable, and independent of the rater, for example by:</p> <ul style="list-style-type: none"> <li>• Utilising the four-eyes principle during the assessment design and related scoring forms</li> <li>• Predetermining the assessment criteria and assessment method</li> <li>• Scheduling a calibration session to align the assessment criteria and their application with the assessors</li> </ul> <p><i>Transparency</i></p> <p>There is evidence that the student knows what to expect from the assessment and how it is scored, for example by:</p> <ul style="list-style-type: none"> <li>• Sharing and discussing assessment criteria with students</li> <li>• Providing students with opportunities to practise</li> <li>• Organising perusing sessions in which students can verify the scoring and grading</li> </ul> <p><i>Educational learning impact</i></p> <p>There is evidence that the student is actively involved in learning and assessment tasks, for example by:</p> <ul style="list-style-type: none"> <li>• Giving students feedback on their performance where one stands and how to progress</li> <li>• Developing students' peer- and self-assessment skills and providing opportunities to practise these skills</li> <li>• Examining student evaluations of the perceived usefulness of the feedback they have received</li> </ul>	<p>There is evidence that the assessment programme is aligned with programme outcomes and the assessment methods used, for example by:</p> <ul style="list-style-type: none"> <li>• Availability of a curriculum matrix that gives insight into how each module contributes to the assessment programme</li> <li>• Clear description and operationalisation of programme outcomes</li> </ul> <p>There is evidence that the assessment programme provides sufficient information to make reliable decisions, for example by:</p> <ul style="list-style-type: none"> <li>• Presenting an overview of how the variation of assessment practices is included in the assessment programme</li> <li>• Providing assessment training and professional development to teachers and examination board members</li> </ul> <p>There is evidence that the assessment programme is clearly structured and understood by all stakeholders, for example by:</p> <ul style="list-style-type: none"> <li>• Presenting an overview of how the assessment programme is built up in complexity</li> <li>• Availability of clear assessment procedures and guidelines for teachers</li> </ul> <p>There is evidence that the assessment programme optimally supports and enhances students' learning process towards accomplishing the defined learning outcomes, for example by:</p> <ul style="list-style-type: none"> <li>• Visibility of learning trajectories</li> <li>• Balancing formative and summative assessment tasks</li> <li>• Continuous feedback loops throughout the programme</li> </ul>

**Table 2.**  
Examples of operationalisation of assessment quality criteria at the module and programme levels

**Table 3.**  
Overview design  
guidelines for digital  
assessment quality  
assurance  
application

Goals	Design guidelines
Improvement	<p>This requires that the application:</p> <ul style="list-style-type: none"> <li>• involves various assessment stakeholders</li> <li>• is available for day-to-day practice</li> <li>• facilitates a PDCA to continuously monitor and reflect on improvements</li> <li>• includes information that supports stakeholders in creating an overview of individual assessment practices</li> <li>• includes information that supports stakeholders in creating an overview of the assessment programme</li> <li>• is informative about assessment quality (giving insight into its potential weaknesses and strengths) instead of normative (assessing the quality as sufficient or insufficient)</li> </ul>
Accountability	<p>This requires that the application:</p> <ul style="list-style-type: none"> <li>• includes information on how individual assessment practices are realised</li> <li>• includes information on how the assessment programme is structured</li> <li>• makes use of assessment quality criteria that are commonly used for accreditation purposes</li> <li>• stores and visualises information (evidence) that is requested for accreditation purposes</li> </ul>
User-friendly	<p>This requires that the application:</p> <ul style="list-style-type: none"> <li>• complies with stakeholders' current quality assurance procedures and practices</li> <li>• is transparent (safe to use for all stakeholders)</li> <li>• is adaptive/flexible to the user's organisational role (different rights and obligations)</li> <li>• has the possibility to input data at different times</li> <li>• enables intuitive use</li> </ul>

multiple visualisations” (Schwendimann *et al.*, 2016, p. 37). In Figure 3, a use case diagram is displayed to make the dashboards easier to interpret by illustrating the user’s possible actions with the application (Gemino and Parker, 2009). Figures 4–6 show various dashboard visualisations at the module and programme levels.

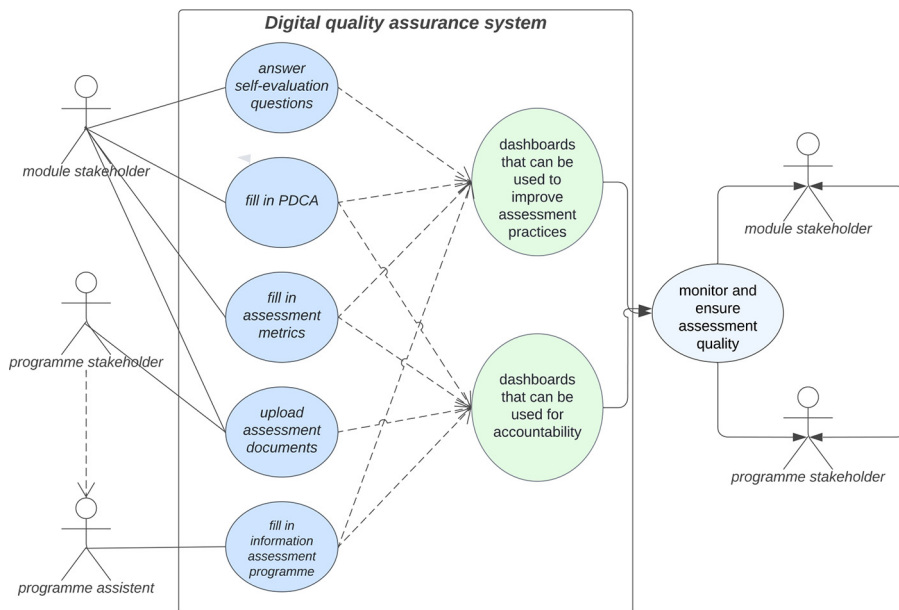
*Dashboards at the module level.* Figures 4 and 5 depict dashboards displaying module-level assessment information. Specifically, Figure 4 presents the dashboard that was created to collect data on individual assessment practices, whereas Figure 5 illustrates an overview dashboard, aggregating data from multiple assessment practices into a single view. This dashboard was created to visualise how module stakeholders view various aspects of assessment quality. Module-level information is accessible to the module’s owner and programme stakeholders. The following features and content are included in dashboards. The numbers in brackets denote features in Figures 4 and 5:

- A PDCA cycle.

The PDCA serves as a guide for teachers’ assessment practices to monitor and improve assessment quality. The PDCA cycle is comprised of four stages that correspond to teachers’ daily practices: design stage (1a); administration stage (1 b); evaluation stage (1c); and action stage (1d).

- Self-evaluation questions.

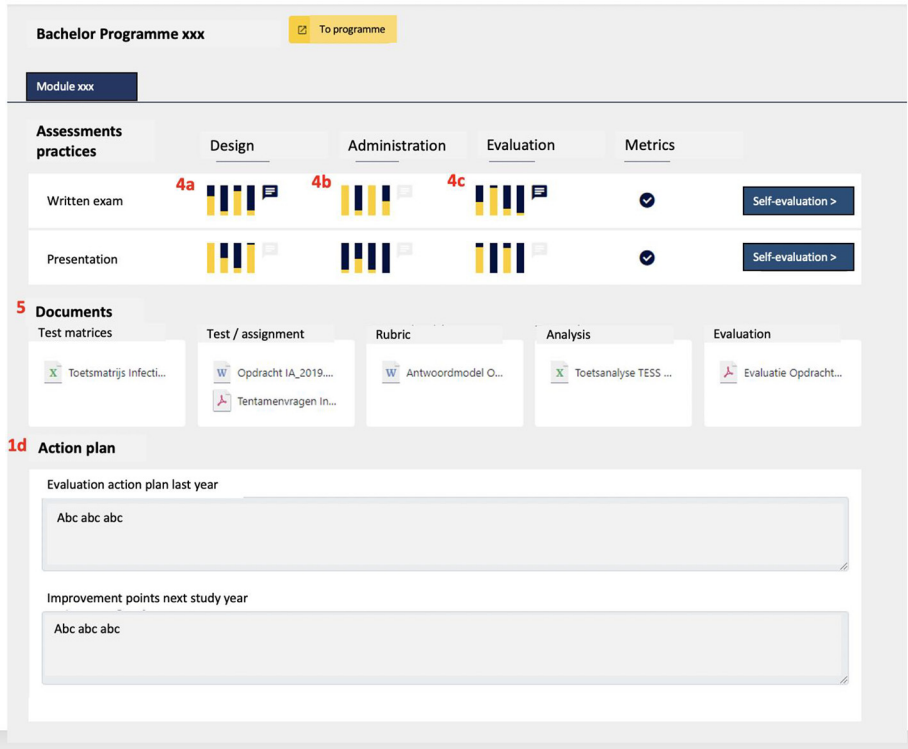
The questions evaluate individual assessment practices in Stages 1–3 of the PDCA. Four questions can be answered during each stage. The content of these questions covers four quality criteria utilised to indicate assessment quality: validity (2a); reliability (2b); transparency (2c); and educational learning impact (2d). In Figure 5,



**Figure 3.** Use case diagram to illustrate the digital quality assurance application

Name module	Assessment practice: Written exam	submit
<b>1a</b> Design	Design stage	
<b>1b</b> Administration	<b>2d</b> Question 1 – educational learning impact	<b>2e</b> <input type="checkbox"/> yes <input type="checkbox"/> no <b>2f</b> <input type="checkbox"/> yes <input type="checkbox"/> no <b>2g</b> Are you satisfied? <input type="range" value="50"/> 100
<b>1c</b> Evaluation	<b>2d</b> Question 2 – educational learning impact	<b>2e</b> <input type="checkbox"/> yes <input type="checkbox"/> no <b>2f</b> <input type="checkbox"/> yes <input type="checkbox"/> no <b>2g</b> Are you satisfied? <input type="range" value="20"/> 100
<b>3</b> Assessment metrics	<b>2a</b> Question 3 – validity	<input type="checkbox"/> yes <input type="checkbox"/> no <b>2g</b> Are you satisfied? <input type="range" value="50"/> 100
	<b>2b</b> Question 4 – reliability	<input type="checkbox"/> yes <input type="checkbox"/> no <b>2g</b> Are you satisfied? <input type="range" value="50"/> 100
	<b>2c</b> Question 5 – reliability	<input type="checkbox"/> yes <input type="checkbox"/> no <b>2g</b> Are you satisfied? <input type="range" value="50"/> 100
<b>2h</b>	Rationale	

**Figure 4.** Dashboard module level – data entry



**Figure 5.**  
Dashboard module  
level – overview

questions are displayed for the design stage. The stakeholder indicates whether (2e) or not (2f) the quality criterion is applicable to the assessment practice. Furthermore, the stakeholder indicates their level of satisfaction with each question (2g). The study assumed that less satisfaction could indicate a point for attention or improvement. For each stage, a text field was presented to substantiate choices and answers (2h).

- Assessment metrics.

For each assessment practice, metrics can be included such as grade average, pass/fail rate and reliability measures (i.e. Cronbach alpha's), etc.

- Aggregated data of individual assessment practices

For each assessment practice, an overview is presented of aggregated self-evaluation scores (level of satisfaction) for the PDCA stages of design (4a), administration (4b) and evaluation (4c).

- Storage of assessment documents.

For each assessment practice, documents can be uploaded that are relevant to store for accreditation purposes (i.e. test blueprints, assignments, rubrics, test analysis, evaluations).

*Dashboards at the programme level.* Dashboards at the programme level were created to visualise the structure and content of the assessment programme. The dashboard depicted

in Figure 6 was created to assist users in visualising the alignment of module and programme outcomes (dashboard A). The data in dashboards A and B were compiled using assessment information provided by the programme director (e.g. by means of an assessment policy plan). This information was previously entered and can be added by a programme director's assistant. Dashboards C, D and E present aggregated data derived from module-level information sources. All programme dashboards were accessible for all stakeholders involved. Dashboards include the following features and content. The numbers in brackets denote features in Figure 6.

- Interactive information filters.

All information at the programme level can be filtered (1), indicating that selections can be made per year (1a) and per module (1b) as needed. This enables stakeholders to examine the differences between modules and/or years and to select specific learning trajectories.

- Assessment programme policy and procedures (2).

Assessment documents that are relevant to the entire programme can be stored and are available for all stakeholders involved.

- Dashboard programme outcomes (A).

This dashboard presents an overview of the alignment between module (a1) and programme outcomes (a2). It displays in which module (a1) and how often (a3) the programme outcomes are assessed. The programme outcomes appear in a textbox when scrolling with a mouse (a4). When a programme outcome is assessed, the cell turns grey (a5).

- Dashboard assessment methods (B).

This dashboard presents an overview of the alignment between the assessment methods used (i.e. written exam, presentation, paper, etc) and programme outcomes.

- Dashboard assessment metrics (C).

This dashboard presents an overview of the assessment metrics that were entered for individual assessment practices (i.e. grade average, pass/fail percentages, reliability measures).

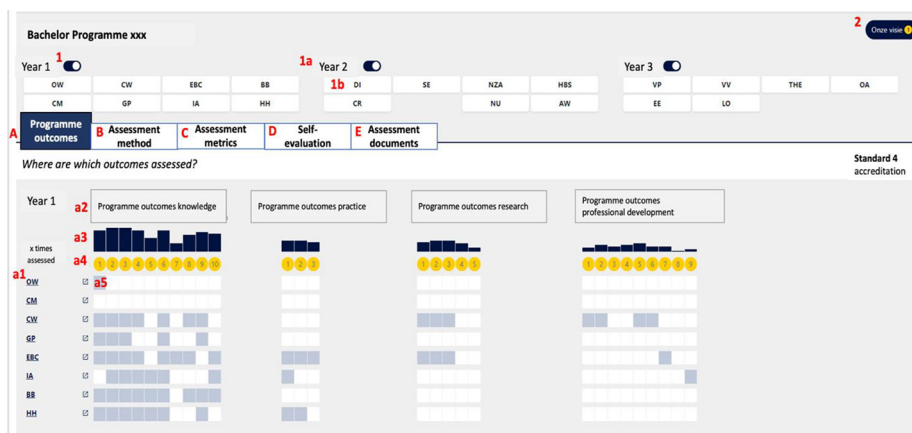


Figure 6. Dashboard programme level – programme outcomes

- Dashboard self-evaluation (D).

This dashboard presents an overview of the self-assessment questions that were answered by the module stakeholder. Data is aggregated for each module and visualised by categorising it according to one of the four assessment quality criteria.

- Dashboard assessment documents (E).

This dashboard presents an overview of the assessment documents that were uploaded by the module stakeholder.

#### *Stakeholders' perceptions of effort and performance expectancy*

This section presents stakeholders' perceptions of the developed application in terms of its effort and performance expectancy.

*Participants' feedback on effort expectancy.* We asked stakeholders about the perceived degree of effort associated with the use of the application. An analysis of their answers indicated the existence of two main themes:

- (1) ease of use; and
- (2) presenting information.

*Ease of Use.* Most of the participants found the application to be user friendly and straightforward to use. They were thus able to easily navigate between the dashboards:

It was actually very gallant, also because of the design and the ease with which you can navigate the data. It remains clear. You can quickly return to the main screen. And you don't get bogged down in all sorts of sub-screens. (Programme stakeholder 2, programme 1).

Some bugs and lay-out problems were reported. For example, when using a specific device, certain dashboards and functionalities within the application were not properly laid out.

*Presenting information.* While the visualisation and simple design of the application meant that it was generally perceived as being easy to use, some participants had difficulty processing the information displayed. Similarly, some programme dashboards were perceived as having too much detail and thus being difficult to interpret:

There's a lot of data here. So, to obtain a good image, you must take in a lot of data. (Programme stakeholder 2, programme 2)

To avoid information overload, participants recommended to include a legend explaining the use of colours, numbers and symbols. Furthermore, various stakeholders did not fully comprehend all the information included. In the self-evaluation questions, for example, assessment concepts such as test blueprint, rubric and formative assessment were used. Additionally, participants mentioned that the programme learning outcomes were formulated in such a way that they were too abstract to comprehend. Participants also recommended that the assessment include additional information recourses (e.g. hyperlinks or a glossary) to facilitate a shared framework of assessment concepts and to support the standardisation of assessment concepts that are commonly used in the study programme.

*Participants' feedback on performance expectancy.* Stakeholders were asked about their perceived performance expectancy, or the degree to which the user expects that utilising the application will help them to improve quality assurance practices. From the interviews, two themes emerged:

- (1) supporting current quality assurance practices; and
- (2) insight in quality.

---

*Supporting current quality assurance practices.* In general, the application complied with stakeholders' current methods to ensure assessment quality. In terms of additional workload, module stakeholders said that filling out the application required minimal extra time because the necessary information was already provided. At the module level, the four stages identified in the PDCA aligned with the practices of module stakeholders to ensure assessment quality (even though in practice stakeholders often failed to consciously go through these stages when designing assessments). The self-evaluation questions assisted stakeholders in becoming aware of the various aspects related to assessment quality and facilitated them in reflecting on their assessment practices. The questions also prompted some module stakeholders to try out new things:

I have two major goals for next year. A rubric is the first of these, that I will be able to better assess the module's final report, that the way I assess becomes a little more visible to the students. [...] And the second is that I might want to work with the students for the review report assessment. That they look over each other's reports. I had considered it before, but now that I've completed it [the self-evaluation questions], I realised: okay, I could involve the students themselves a bit more in that. (Module stakeholder 4, programme 1)

At the programme level, the application provided stakeholders with information on how the assessment was structured and how each module contributed to the programme's outcomes. This information helped stakeholders to reflect on the programme and provided them with insight into areas for improvement:

You see that it [programme learning outcomes] is distributed very unevenly and you also see that some modules cover all the programme outcomes. And, thinking with all your common sense, I think "but that is just not possible". We therefore need to take a good look at how this works. (Programme stakeholder 2, programme 2).

Programme stakeholders valued the fact that information was easily retrievable and up to date. This saved them time in requesting information for module analyses and accreditation. For this reason, some module stakeholders stated that the application would benefit programme stakeholders rather than individual module stakeholders. Suggestions for further compliance with current quality assurance procedures were the creation of a dashboard with a PDCA at the programme level and the possibility of automatically uploading information from other systems (e.g. student evaluations, grading systems, etc).

*Insight in quality.* The module overview allowed stakeholders to quickly see how satisfied they were with their assessment practices. However, most of the participants reported that the aggregated data of self-evaluation questions (i.e. whether module stakeholders were satisfied or not with various quality aspects of their assessment methods) was not directly related to assessment quality. Due to the subjective evaluation of the module stakeholder self, this overview did not equip stakeholders with the knowledge needed to improve assessment quality. Participants preferred that assessment quality questions be compared to (objective) quality guidelines or feedback from third parties:

Is it necessary that I am satisfied with quality aspects of the assessment? It's not just about whether I am satisfied as an individual. It's also about if my colleagues, the department, the students, or anybody else are satisfied as well. That is less measured here. (Module stakeholder 5, programme 3).

Dashboards with aggregated assessment metrics and assessment documents at the programme level provided stakeholders with a better understanding of assessment quality because these indicators were based on measurable, objective data and included the ability to compare modules. Furthermore, programme-level dashboards enabled module stakeholders to "peek into another's kitchen". Module stakeholders valued the opportunity



---

to gain insight into how the various assessment practices were constructed and applied at the programme level:

You are frequently on your own as a teacher. This [application] is, in my opinion, a very good method to question how we are performing as a study programme. I believe it is also critical to be able to compare your own module to what others are doing. It also serves as a starting point for discussion: how do we do this? Are you going to change something in your module, or am I going to do that? (Module stakeholder 4, programme 3).

---

To improve further insight into assessment quality, it was recommended that self-evaluation questions be answered and compared within a module-team and that deviations from the norm would be visualised and highlighted, rather than simple data aggregation.

### Conclusion and discussion

In this DBR study, a digital quality assurance application for higher education assessment programmes was iteratively designed (Phases 1 and 2), developed (Phase 3) and evaluated (Phase 4). The application's design was informed by a literature search and needs assessment of the programme and the module assessment stakeholders. Design guidelines were formulated to develop an application that is easy to use in daily practice and serves the dual purposes of improvement and accountability. As part of the development of the application, module-level dashboards were created to assist module stakeholders in a PDCA cycle of designing, administering and evaluating assessment practices using self-evaluation questions. Programme-level dashboards were developed to provide a comprehensive overview of the outcomes assessed at each level, in what way and how often. The application was evaluated for perceived effort and performance expectancy by stakeholders at the module and programme levels because these concepts directly link to stakeholders' intentions to use the developed application (Venkatesh *et al.*, 2003). Regarding effort expectancy, participants generally perceived the application as user friendly. In terms of performance expectancy, the application generally supported participants in their quality assurance practices. For example, module stakeholders perceived the included PDCA as supporting in becoming aware of the various methods assessment quality can be ensured.

By digitising assessment quality assurance procedures, the application enhances the value of existing procedures (Baartman, 2007, 2011; Lucander and Christersson, 2020). A primary advantage of an online application is that it is always accessible, so multiple assessment stakeholder groups can work on it independently. For example, in the current study was found that stakeholders were motivated to use the application because they gained control over their own quality assurance procedures. Most participants perceived the application as user friendly and were able to begin using it after reading the manual. Hence, in contrast to the developed procedures outlined in the introduction (Baartman *et al.*, 2007, 2017; Lucander and Christersson, 2020), there is no need to schedule group interviews or hire an external team to work on quality assurance. In this way, the application can support educational programmes in achieving their aim of integrating assessment quality into daily activities (Emil and Cress, 2014), thus preserving stakeholders' commitment and motivation (EUA, 2010).

A second advantage of using an online application to assess quality assurance is that the developed dashboards, at the module and programme levels, allow the user to visualise assessment quality. Depending on the question being investigated, or the level of quality that is at stake, different overviews can be created. These overviews contribute to the transparency of the assessment programme by enabling stakeholders to comprehend how their practices contribute to the assessment programme and to gain insight into how their peers conduct assessments and identify potential gaps. To ensure transparency, all

programme dashboards are accessible to all stakeholders. The majority of the participants appreciated this because it enabled them to view assessment holistically. This may encourage discussions of the assessment programme's structure and can facilitate module stakeholders' collaboration across modules (Medland, 2016). The visualisation of assessment quality was also useful for accountability purposes. For example, the application provided information about how programme outcomes were realised and gave stakeholders insight into the aspects that relate to assessment quality (i.e. validity, reliability, transparency and learning impact). Programme stakeholders found information sources such as assessment documents and metrics helpful for analysing and ensuring assessment quality. However, for most participants, the aggregated data based on module stakeholders' self-ratings did not directly relate to assessment quality. These measures were deemed too subjective to consider because they were not linked to predefined quality standards or independently assessed. This is in line with Baartman *et al.* (2007), who stated that when self-evaluation is used for accountability purposes, issues of reliability become critical. Additional research is required to determine the extent to which self-evaluations benefit both the improvement and accountability purposes.

In this study, we focused on the design, development and evaluation of a digital application for assessment quality assurance. The application's design is unique in that it was produced through cocreation. As a result, the developed prototype had a good fit with the users' daily quality assurance procedures and working with the application was generally not experienced as an extra workload. Regarding future application development, the feasibility of creating a programme-level PDCA should be examined because this functionality was designed only at the module level. In addition, users should be guided to process the information on the dashboards, for example, by including legends or hyperlinks. The results of the evaluation will be used to redesign the application, which is the final phase of a DBR study. This final phase was not included in the current study; as a result, the impact of the application on participant behaviour and changes in assessment practices were not examined. This is a limitation of the current study. However, in the pilot study was found that, by using the application, participants became more familiar with the procedures and quality indicators used to evaluate assessment quality, thereby enhancing their assessment literacy (Price *et al.*, 2011). The results of a follow-up study conducted by Lucander and Christersson (2020), which examined the effects of implementing a quality assurance procedure, showed that the development of assessment literacy did change practices because it resulted "in an enhanced assessment structure and curriculum reform" (p. 148).

When using a digital application to assess quality assurance, an issue that requires further consideration is how to best use available assessment data. Possible impediments to effective information use include information being accessed but not acted on and actions not resulting in significant improvements (Kuh *et al.*, 2015). New research can focus on, for example, which actions stakeholders take on the basis of the information (EUA, 2010). Furthermore, when assessment information will be stored for multiple years, it will be essential to focus on data collection limitations to maintain efficiency (EUA, 2010). In this regard, the use of learning analytics techniques is worth investigating in relation to how it can stimulate workplace learning (Van der Schaaf *et al.*, 2017).

We found that the digital application developed provides a comprehensive picture of the assessment programme's quality and supports internal and external quality assurance procedures. Although the application was evaluated at a single institution in The Netherlands, the different study programmes that participated adhere to European quality concepts and standards (Lucander and Christersson, 2020). We also acknowledge that continuously improving and ensuring assessment quality entails more than simply adding a

new application (Dolmans and Tigelaar, 2012); it also necessitates a cultural shift, as evidenced by a shift away from quality control (with an emphasis on accountability) and towards increased autonomy based on the experiences and expertise of the stakeholders involved (Bendermacher *et al.*, 2017; EUA, 2006). To foster such a learning culture, the introduction and implementation of the application are critical. Expectations must be justified for successful embedding of the application. All stakeholders must understand why the application is necessary to use and what is expected of them. It is also important that all stakeholders have a shared understanding of programme outcomes, assessment quality and the factors that influence them (Price *et al.*, 2011; Russell and Markle, 2017). By facilitating ongoing discussions about assessment (programme) quality as represented in the application, the application can help stakeholders initiate productive dialogue and think about assessment quality in a more sophisticated manner, thereby contributing to the development of a high-quality culture of learning (Bendermacher *et al.*, 2017).

*Funding:* The digital assessment quality assurance application was developed with funding from the Utrecht Education Incentive Fund from Utrecht University and funding of Cito Foundation, Institute for Educational Measurement in The Netherlands.

*Disclosure statement:* The authors of this study certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The digital application is open-source software and is published under license of Microsoft Public License and is approved by the Open-Source Initiative: <https://opensource.org/licenses/MS-PL>. The code can be found at <https://github.com/Citolab/equality>. The application will be further developed in collaboration with Utrecht University.

## References

- Baarda, D.B., de Goede, M.P.M. and Teunissen, J. (2001), "Basisboek kwalitatief onderzoek: praktische handleiding voor het opzetten en uitvoeren van kwalitatief onderzoek", Stenfert Kroese.
- Baartman, L., Kloppenburg, R. and Prins, F. (2017), "Kwaliteit van toetsprogramma's", *Toetsen in Het Hoger Onderwijs*, Bohn Stafleu van Loghum, Houten, pp. 37-49.
- Baartman, L.K.J., Prins, F.J., Kirschner, P.A. and Van der Vleuten, C.P.M. (2007), "Determining the quality of assessment programs: a self-evaluation procedure", *Studies in Educational Evaluation*, Vol. 33 Nos 3/4, pp. 258-281.
- Baartman, L.K., Prins, F.J., Kirschner, P.A. and van der Vleuten, C.P. (2011), "Self-evaluation of assessment programs: a cross-case analysis", *Evaluation and Program Planning*, Vol. 34 No. 3, pp. 206-216.
- Bakker, A. and Van Eerde, D. (2015), "An introduction to design-based research with an example from statistics education", *Approaches to Qualitative Research in Mathematics Education*, Springer, Dordrecht, pp. 429-466.
- Bendermacher, G., Oude Egbrink, M.G., Wolfhagen, I. and Dolmans, D.H. (2017), "Unravelling quality culture in higher education: a realist review", *Higher Education*, Vol. 73 No. 1, pp. 39-60.
- Bijkerk, L. (2015), *Basis Kwalificatie Examinering in Het Hoger Beroepsonderwijs*, Springer media, Houten.
- Bok, H.G., Teunissen, P.W., Favier, R.P., Rietbroek, N.J., Theyse, L.F., Brommer, H., Haarhuis, J.C., van Beukelen, P., van der Vleuten, C.P. and Jaarsma, D.A. (2013), "Programmatic assessment of competency-based workplace learning: when theory meets practice", *BMC Medical Education*, Vol. 13 No. 1, p. 123.

- 
- Boud, D. and Falchikov, N. (2006), "Aligning assessment with long-term learning", *Assessment and Evaluation in Higher Education*, Vol. 31 No. 4, pp. 399-413.
- Cohen, L., Manion, L. and Morrison, K. (2018), *Research Methods in Education*, 8th ed., Abingdon, Oxon.
- Dijkstra, A. and Baartman, L. (2011), "Zelfevaluatie van de kwaliteit van assessment", *Tijdschrift OnderwijsInnovatie*, Vol. 1, pp. 17-25.
- Dill, D. (2007), *Quality Assurance in Higher Education: Practices and Issues*, Elsevier Publications, Chapel Hill, NC.
- Dolmans, D.H. and Tigelaar, D. (2012), "Building bridges between theory and practice in medical education using a design-based research approach: AMEE guide no. 60", *Medical Teacher*, Vol. 34 No. 1, pp. 1-10.
- Emil, S. and Cress, C. (2014), "Faculty perspectives on programme curricular assessment: individual and institutional characteristics that influence participation engagement", *Assessment and Evaluation in Higher Education*, Vol. 39 No. 5, pp. 531-552.
- EUA (2006), "Quality culture in European Universities: a bottom-up approach: report on the three rounds of the quality culture project 2002-2006 (9081069837)".
- EUA (2010), "Examining quality culture: part 1-quality assurance processes in higher education institutions (9078997214)".
- Gemino, A. and Parker, D. (2009), "Use case diagrams in support of use case modeling: deriving understanding from the picture", *Journal of Database Management*, Vol. 20 No. 1, pp. 1-24.
- Gerritsen-van Leeuwenkamp, K.J., Joosten-ten Brinke, D. and Kester, L. (2017), "Assessment quality in tertiary education: an integrative literature review", *Studies in Educational Evaluation*, Vol. 55, pp. 94-116.
- Harvey, L. and Stensaker, B. (2008), "Quality culture: understandings, boundaries and linkages", *European Journal of Education*, Vol. 43 No. 4, pp. 427-442.
- Hobson, R., Rolland, S., Rotgans, J., Schoonheim-Klein, M., Best, H., Chomyszyn-Gajewska, M., Dymock, D., Essop, R., Hupp, J. and Kundzina, R. (2008), "Quality assurance, benchmarking, assessment and mutual international recognition of qualifications", *European Journal of Dental Education*, Vol. 12 No. s1, pp. 92-100.
- Hulpiau, V. and Waeytens, K. (2003), "Improving quality of education: what makes it actually work? A case study", *Realizing Qualitative Research into Higher Education*, Routledge, Abingdon, Oxfordshire, pp. 145-169.
- Inspectie van het Onderwijs (2018), "De kwaliteit van het Nederlandse accreditatiestelsel hoger onderwijs".
- Jessop, T., McNab, N. and Gubby, L. (2012), "Mind the gap: an analysis of how quality assurance processes influence programme assessment patterns", *Active Learning in Higher Education*, Vol. 13 No. 2, pp. 143-154.
- Kleijnen, J., Dolmans, D., Willems, J. and Hout, H.V. (2013), "Teachers' conceptions of quality and organisational values in higher education: compliance or enhancement?", *Assessment and Evaluation in Higher Education*, Vol. 38 No. 2, pp. 152-166.
- Kuh, G.D., Ikenberry, S.O., Jankowski, N.A., Cain, T.R., Ewell, P.T., Hutchings, P. and Kinzie, J. (2015), *Using Evidence of Student Learning to Improve Higher Education*, John Wiley and Sons, Oxford.
- Lucander, H. and Christersson, C. (2020), "Engagement for quality development in higher education: a process for quality assurance of assessment", *Quality in Higher Education*, Vol. 26 No. 2, pp. 135-155.
- Maassen, N.A.M., den Otter, D., Wools, S. and Hemker, B. (2015), "Kwaliteit van toetsen binnen handbereik: reviewstudie van onderzoek en onderzoeksresultaten naar de kwaliteit van toetsen", *Pedagogische Studiën*, Vol. 92 No. 6, pp. 380-393.
- Marciniak, R. (2018), "Quality assurance for online higher education programmes: design and validation of an integrative assessment model applicable to Spanish universities", *International Review of Research in Open and Distributed Learning*, Vol. 19 No. 2, pp. 126-154.
- Medland, E. (2016), "Assessment in higher education: drivers, barriers and directions for change in the UK", *Assessment and Evaluation in Higher Education*, Vol. 41 No. 1, pp. 81-96.

- Norcini, J., Anderson, M.B., Bollela, V., Burch, V., Costa, M.J., Duvivier, R., Hays, R., Palacios Mackay, M.F., Roberts, T. and Swanson, D. (2018), "2018 Consensus framework for good assessment", *Medical Teacher*, Vol. 40 No. 11, pp. 1102-1109.
- NVAO (2016), "Assessment and demonstration of achieved learning outcomes: recommendations and good practices".
- NVAO (2018), "Beoordelingskader accreditatiestelsel hoger onderwijs Nederland".
- Price, M., Carroll, J., O'Donovan, B. and Rust, C. (2011), "If I was going there I wouldn't start from here: a critical commentary on current assessment practice", *Assessment and Evaluation in Higher Education*, Vol. 36 No. 4, pp. 479-492.
- Russell, J. and Markle, R. (2017), "Continuing a culture of evidence: assessment for improvement", *ETS Research Report Series*, Vol. 2017 No. 1, pp. 1-10.
- Schellekens, L.H., Slof, B. and Bok, H.G. (2021), "Quality assurance of an assessment program", in Malau-Aduli, B., Hays, R. and Van der Vleuten, C. (Eds), *Understanding Assessment in Medical Education*, McGraw Hill, New York City.
- Schwendimann, B.A., Rodriguez-Triana, M.J., Vozniuk, A., Prieto, L.P., Boroujeni, M.S., Holzer, A., Gillet, D. and Dillenbourg, P. (2016), "Perceiving learning at a glance: a systematic literature review of learning dashboard research", *IEEE Transactions on Learning Technologies*, Vol. 10 No. 1, pp. 30-41.
- Sluijsmans, D., Joosten-ten Brinke, D. and van Schilt-Mol, T. (2015), "Kwaliteit van toetsing onder de loep. Handvatten om de kwaliteit van toetsing in het hoger onderwijs te analyseren, verbeteren en borgen", Maklu.
- Sridharan, B., Leitch, S. and Watty, K. (2015), "Evidencing learning outcomes: a multi-level, multi-dimensional course alignment model", *Quality in Higher Education*, Vol. 21 No. 2, pp. 171-188.
- Van Berkel, H., Bax, A. and Joosten-ten Brinke, D. (2017), "Toetsen in het hoger onderwijs", Bohn stafleu van Loghum.
- Van der Schaaf, M., Donkers, J., Slof, B., Moonen-van Loon, J., van Tartwijk, J., Driessen, E., Badii, A., Serban, O. and Ten Cate, O. (2017), "Improving workplace-based assessment and feedback by an e-portfolio enhanced with learning analytics", *Educational Technology Research and Development*, Vol. 65 No. 2, pp. 359-380.
- Van der Vleuten, C.P. and Schuwirth, L.W. (2005), "Assessing professional competence: from methods to programmes", *Medical Education*, Vol. 39 No. 3, pp. 309-317.
- Van der Vleuten, C.P., Schuwirth, L.W., Driessen, E.W., Dijkstra, J., Tigelaar, D., Baartman, L.K. and van Tartwijk, J. (2012), "A model for programmatic assessment fit for purpose", *Medical Teacher*, Vol. 34 No. 3, pp. 205-214.
- Venkatesh, V., Morris, M.G., Davis, G.B. and Davis, F.D. (2003), "User acceptance of information technology: toward a unified view", *MIS Quarterly*, Vol. 27 No. 3, pp. 425-478.
- Vlăsceanu, L., Grünberg, L. and Pârlea, D. (2004), *Quality Assurance and Accreditation: A Glossary of Basic Terms and Definitions*, Unesco-Cepes Bucharest.
- Wang, F. and Hannafin, M.J. (2005), "Design-based research and technology-enhanced learning environments", *Educational Technology Research and Development*, Vol. 53 No. 4, pp. 5-23.

### Corresponding author

Lonneke H. Schellekens can be contacted at: [l.h.schellekens@uu.nl](mailto:l.h.schellekens@uu.nl)

---

For instructions on how to order reprints of this article, please visit our website:

[www.emeraldgroupublishing.com/licensing/reprints.htm](http://www.emeraldgroupublishing.com/licensing/reprints.htm)

Or contact us for further details: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)