scope
Contemporary Research Topics

learning & teaching 16

November 2026

Article

https://doi.org/10.34074/scop.4016004

DEVELOPING A TEACHING PRESENCE RUBRIC: THEORY-BASED EVALUATION OF ONLINE COURSE AND TEACHING QUALITY IN HIGHER EDUCATION

Patrick Gillett, Tina van Eyk, Patrick Bruck,
Carolyn Seton and Kylie Day

Published by Otago Polytechnic Press.

CC-BY the authors.
© illustrations: the artists or other copyright owners or as indicated.

Article https://doi.org/10.34074/scop.4016004

DEVELOPING A TEACHING PRESENCE RUBRIC: THEORY-BASED EVALUATION OF ONLINE COURSE AND TEACHING QUALITY IN HIGHER EDUCATION

Patrick Gillett, Tina van Eyk, Patrick Bruck, Carolyn Seton and Kylie Day

INTRODUCTION

The use of student feedback to evaluate course and teaching quality in higher education has a long and contentious history (Alderman et al., 2012; Cunningham et al., 2023; Richardson, 2005). In particular, while it is well-accepted that the student perspective is an essential input for continuous improvement, strong criticism has been levelled at the most common feedback method: student evaluation of teaching (SET) surveys.

The negative position on SET is extensively supported by rigorous studies, including those that have identified serious validity concerns. These issues include response bias (see Heffernan, 2022), low participation rates (see Lipsey & Shepperd, 2021), and limited generalisability (see Curby et al., 2020). Research by Esarey and Valdes (2020) also reveals that, even when these concerns are accounted for, SET data is associated with a high error rate. This result adds further weight to the argument put forward by Nargundkar and Shrikhande (2014) that student ratings of course and teaching quality are substantially influenced by non-instructional factors such as class size, discipline type, and course level.

In addition to the strong quantitative data, criticism of the SET method is also supported by qualitative analysis, with Cunningham et al. (2023) and Lakeman et al. (2022) both providing disturbing examples of offensive student comments in SET reports. Although the number of objectionable remarks is small, the impact on instructor well-being can be substantial. To avoid potential harm, Lakeman et al. (2022) posit that many instructors ignore feedback altogether, thereby rendering it worthless for continuous improvement.

While research undertaken to date offers good reason to question whether SET is an appropriate method for evaluating course and teaching quality, its ongoing use in higher education is unlikely to end. This is largely because institutions have come to rely on it for its role in accreditation compliance and as an endorsement for academic appointments (Cunningham et al., 2023). Recognising its multifunctionality, many have argued for SET to be used alongside other course and teaching evaluation methods (Alderman et al., 2012; Esarey & Valdes, 2020; Lakeman et al., 2022). As detailed by Esarey and Valdes (2020), a diversified strategy can lead to more reliable evaluation outcomes as the inherent limitations of one particular method are unlikely to dominate the collective result. Among the few methods that may be suitable to complement the SET is the course evaluation rubric (Gregori-Giralt & Menéndez-Varela, 2021; McInnes et al., 2024).

COURSE EVALUATION RUBRICS

Increased attention in the higher education literature to the assessment of online courses and teaching quality has led to the development of numerous course evaluation rubrics (also referred to as "course quality" rubrics). Notable examples include the Quality Matters Higher Education Rubric, Seventh Edition (Quality Matters, 2023), and the SUNY Online Course Quality Review Rubric (SUNY, 2023). While such instruments have considerable merit, they also possess several limitations that should not be ignored. Among them is the lack of statistical testing to substantiate the rubric's validity and reliability (Yuan & Recker, 2015) and the absence of a theoretical foundation to support the rubric design (Lee et al., 2020). In terms of the latter, the rubrics are based on a 'best-practice' approach, which involves using subjective interpretations of quality to establish the evaluation targets and their individual standards. What constitutes "quality" in higher education, however, is a widely contested topic (McInnes et al., 2024). Consequently, there is considerable potential for variability in the application and interpretation of the instruments due to a focus on irrelevant or indistinct criteria. Rather than strengthening the evaluation result, adopting a diversified strategy that includes an existing course evaluation rubric is likely to introduce further credibility concerns.

AIM

This research project aims to develop a robust 'theory-based' rubric to aid the evaluation of online course and teaching quality. The research context is higher education programmes delivered entirely online as well as blended types, with the latter being restricted to their online components. The number of fully online and blended courses was not recorded separately; both were analysed together, and delivery mode was not examined as a separate factor. The online context was chosen over traditional 'in-situ' programmes due to ongoing concerns regarding its effectiveness and lasting appeal (see Brown, 2023). In particular, lower retention and completion rates are often associated with online programmes (Stone & Springer, 2019), which suggests that the students who take part in them are less successful and less satisfied with their educational experience. The research aim explicitly follows the strategy specified by Garivaldis et al. (2022), which is to address such concerns by demonstrating how established theories can be applied to advance the management and practice of online education.

The Quality Improvement Collaborative (QIC) model was used to establish a multi-disciplinary team for the research and to create the Teaching Presence rubric. The theory considered most appropriate for the research aim was Garrison et al.'s (2000) Community of Inquiry (COI) model. Recognised as one of the most heavily cited theories in the online education literature (Caskurlu, 2018; Richardson et al., 2012), COI contends that an effective educational experience is characterised by meaningful inquiry and deep learning, and that these outcomes are best achieved through the active engagement of students in a community-type learning environment. To facilitate this setting, the model gives attention to three central components: teaching presence, social presence, and cognitive presence. While social presence fosters meaningful connections, cognitive presence enables critical enquiry and knowledge construction.

When examining the COI model, it became evident to the project team that incorporating all three components into a single rubric would result in an extensive and unwieldy instrument. Although the rubric might be theoretically sound, the substantial administrative burden would severely limit its practical value. A rubric based on the single "teaching presence" component was

chosen as an effective alternative for two key reasons. The first is the strong interrelationship that exists among the three COI components, while the second is the widely held view that teaching presence acts as a unifying force that binds all three components to create a community of inquiry (Garrison et al., 2000; Vaughan et al., 2013). Consequently, the focus on teaching presence maintains the research aim while also offering greater practical value. As detailed in the following sections of this article, the design and validation of such a rubric followed a systematic process, beginning with an extensive review of the teaching presence literature.

TEACHING PRESENCE

Anderson et al.'s (2001) model was selected as it provides the foundational operationalisation of teaching presence within the Community of Inquiry framework and directly defines the three functions assessed by the rubric. The research team determined early that Anderson's model was best for our practical analysis of learning environments as the model is universally applied by teaching scholars. It recognises the critical role of three particular teaching functions in achieving "personally meaningful and educationally worthwhile learning outcomes" (p. 2). The three functions—design and organisation, facilitating discourse, and direct instruction—encapsulate the work undertaken to plan and administer the course, advance the cognitive engagement of students, and disperse subject knowledge. The higher education literature supports the positive influence of teaching presence on student satisfaction and perceptions of learning (Arbaugh, 2008; Caskurlu et al., 2020).

The three teaching functions can involve contributions from multiple course stakeholders. For example, the "design and organisation" function will benefit from a coordinated effort involving the programme director, course instructor, and instructional designers. Vaughan et al. (2013) also point out that it is not just the course instructor who can lead conversations on the course content and provide direct instruction. Guest presenters, for example, can also make meaningful contributions. It is this potential for collaboration and purpose among key stakeholders that underscores the strength of the Teaching Presence model.

In this sense, the project is concerned not only with measurement but also with developing a shared, theory-informed language for discussing online teaching quality in higher education. This is particularly relevant in contemporary tertiary learning and teaching contexts, where institutions are increasingly seeking evaluation approaches that move beyond student satisfaction alone and instead support reflective practice, peer dialogue, and course enhancement. By developing a rubric grounded in Teaching Presence, this paper contributes to scholarship on learning and teaching as both a practical and collegial field of inquiry.

Design and organisation

Anderson et al. (2001) describe the design and organisation function as a collection of teaching activities undertaken to establish "the process, structure, evaluation, and interaction components" of an online course (p. 5). The function effectively provides a "grand design of the course," one that is explicit and transparent in terms of its learning goals and the actions students should undertake to achieve them (p. 6). Explicitness and transparency are particularly important for online students as the awareness of learning expectations and related norms that often accrue through classroom attendance may not be available to them (Anderson et al., 2001).

As detailed in Table 1, a targeted review of the higher education literature was conducted to identify and synthesise teaching activities associated with the teaching presence functions. Specifically, the review examined how scholars have conceptualised and operationalised the design and organisation function in contemporary learning environments. This process revealed seven distinct teaching activities that directly support design and organisation.

While many of the activities identified are consistent with those initially proposed by Anderson et al. (2001), the collection has been expanded with the inclusion of two topics of more recent research: "equity and inclusion" and "constructive alignment."

Activity	Contributors	Description
Course Aim	Programme Director, Course Instructor	A concise description of the critical course features, including: 1) the subject matter, 2) the method(s) used to explore the subject matter, 3) the breadth and depth of study, and 4) why the course should be studied (Macquarie University, 2022).
Course Learning Outcomes (CLOs)	Programme Director, Course Instructor, Instructional Designer	Individual statements identifying "what a student will be able to do or demonstrate at the completion of a certain sequence of learning (course, program)" (Popenici & Millar, 2015, p. 4). Each statement should be framed by three distinct components: 1) action verb, 2) content, and 3) context (University of New South Wales, 2022).
Module Design	Course Instructor, Instructional Designer	The chosen learning material for each course module is presented in a manner that effectively enables a particular CLO to be achieved. This objective can be assisted by 'signposting,' which clearly establishes the relationship between particular sections of module content and other course material, such as learning activities and assessments (Mayer, 2019).
Learning Activities	Course Instructor, Instructional Designer	Clearly explained activities with timely feedback to encourage student engagement with the learning material and other course participants (Chi & Wylie, 2014).
Equity and Inclusion	Course Instructor, Instructional Designer	Initiatives to support greater access, engagement, and learning outcomes for disadvantaged and under-represented students. Key examples include the provision of closed captioning and audio transcripts with lecture recordings (Martin et al., 2021), and the use of Open Education Resources (Willems & Bossu, 2012).
Assessment Design	Course Instructor, Instructional Designer	Establishing a clear purpose for undertaking the assessment task by explicitly targeting one or more CLOs (Bearman et al., 2016). A strong purpose is further enhanced by assessments that are relevant to, and apply, the particular learning, as these allow students to develop specific work-relevant skills (Bretag et al., 2019).
Constructive Alignment	Course Instructor, Instructional Designer	Teaching and learning quality are achieved through a focused curriculum design process (Biggs, 1996). In particular, student engagement is stimulated when clear and mutually reinforcing links are established between key course design elements, such as the content, learning outcomes, activities, and assessments (Biggs, 2014).

Table 1. 'Design and organisation' teaching activities.

In the examples in Table 1, we distinguish between course instructors and instructional designers. Roles performed by the former include drawing on their subject matter expertise to ensure seamless live facilitation of the courses with a focus on student learning and engagement while the latter are typically brought on at the design phase of course development to focus on interactive activities, assessments, and multimedia. In the design of courses, they generally work together to finalise the course offerings and delivery.

Facilitating discourse

Akyol et al. (2011) describe the "facilitating discourse" function as "enabling and encouraging the construction of personal meaning as well as shaping and confirming mutual understanding" (p. 232). Through this process, students are empowered to make the content their own and are thus more able to understand and apply it effectively. The term "discourse" is adopted over "conversation" to distinguish communication with a learning focus from communication with a social orientation (Anderson et al., 2001).

The distinct teaching activities associated with facilitating discourse that emerged from the literature are presented in Table 2. As the descriptions show, achieving a positive learning outcome is aided by instructors' efforts to encourage and guide discourse among students, and between themselves and students (Garrison & Akyol, 2013; Oliver & Shaw, 2003; Shea et al., 2003).

Activity	Contributors	Description
Instructor Immediacy	Course Instructor	Instructors create an active, open, and focused learning environment by making themselves psychologically available to students (Mehrabian, 1969). For example, by posting their contact details and available meeting times on the LMS, instructors show that they are welcoming and approachable (Caskurlu et al., 2021). Stronger familiarity and immediacy can be achieved by providing a biography and recording a welcome video message (Kelly & Zakrajsek, 2021; Martin et al., 2018).
Encouraging Discourse	Course Instructor, Guest Presenter	For discourse to be encouraged, the opportunity to participate in dialogue with others must first be made available. Liu et al. (2007) identify two distinct methods for doing so. The first involves instructors' informal efforts to reach out to students and ask whether they need help with the course material. The second is a planned learning activity in which the instructor sets a specific question aligned with a course or module learning outcome.
Guiding Discourse	Course Instructor, Guest Presenter	While participant interaction is a fundamental tenet of the Community of Inquiry framework, it does not by itself ensure that an effective learning experience is realised (Garrison et al., 2000). Online discussions, in particular, require the active involvement of instructors to guide and shape the discourse toward the desired learning outcome (Anderson et al., 2001).

Table 2. 'Facilitating discourse' teaching activities.

Direct instruction

When describing the direct instruction function, Anderson et al. (2001) emphasise the responsibility teachers have for providing "intellectual and scholarly leadership" (p. 8). While facilitating discourse plays a largely supportive role in learning, in "direct instruction" the teacher is a skilled expert, adept at directing and scaffolding the online learning experience to achieve a positive educational outcome (Garrison & Akyol, 2013; Vaughan et al., 2013). As reported in Table 3, the higher education literature reveals five distinct teaching activities that directly embody these skills.

Activity	Contributors	Description
Orientation Module	Course Instructor, Instructional Designer	An orientation module targets specific course-level needs, such as raising students' awareness of the course purpose and structure (Martin et al., 2019). Fundamental details include the class timetable and assignment schedule, while deeper support identifies the key challenges students are likely to face and how to manage them.
Teaching and Learning Methods	Course Instructor, Instructional Designer	To establish an effective online learning opportunity, Richardson et al. (2012) emphasise the importance of orienting students to the role of the online learner. Advice should be given on the methods to be used to deliver the course content and maximise learning outcomes. This information is particularly important when unfamiliar learning technologies are involved (Oliver & Shaw, 2003).
Instructor Insights	Course Instructor	Learning materials drawn from a range of sources reflect the instructor's strong understanding of the discipline and are particularly effective when insights and analogies are grounded in personal experience (Caskiurlu et al., 2021).
Instructional Quality	Course Instructor	Effective subject matter is of little value without the instructor's ability to discuss it coherently and in a structured manner. An added challenge for online instructors is to attain the required standard of communication using an electronic video conferencing platform, such as Zoom or Microsoft Teams. Among the advice given is for instructors to keep their webcam on (Kelly & Zakrajsek, 2021) and to position the camera at eye level, as this increases students' perception of credibility (Ramlatchan & Watson, 2020). The quality of lecturing is also improved when the content is well organised and the speaker makes eye contact, whether that is in person or online. Ensuring eye contact online can be achieved by consciously looking directly at the camera lens rather than at the faces on the screen and by ensuring that the video window is as close to the lens as possible.
Feedback	Course Instructor	Providing high-quality feedback is a fundamental requirement in higher education due to its strong influence on student learning (Evans, 2013; Hattie & Timperley, 2007). Factors that contribute to poor feedback include insufficient detail and an overly authoritative tone (Jonsson, 2013), being late, ambiguous, or irrelevant (Henderson et al., 2019), being unbalanced (for example, lacking positive recognition), and being unfair (Lizzio & Wilson, 2008).

Table 3. 'Direct instruction' teaching activities.

METHOD

Several principles of a Quality Improvement Collaborative (QIC) were used to establish the project team and to create the Teaching Presence rubric. Common in the healthcare sector, QICs emphasise a shared learning approach where stakeholders from multiple areas (usually teams of experts) come together to exchange knowledge and insights on a particular topic (Wells et al., 2018). In keeping with this tenet, the project team drew one member with recent online teaching experience from five separate faculties within the same university. Establishing the team in this manner ensured that a range of discipline-specific perspectives would be considered and suited the 'learning by doing' approach adopted by the team. Described by Schouten et al. (2008) as a key QIC principle, this technique involves repeated cycles of testing, sharing details of the experience within the team, and analysing results to further refine the quality improvement effort.

Procedure

After obtaining approval from the university's Human Research Ethics Committee (2022/053), a multi-stage course evaluation procedure commenced. During the first stage, the project team developed an initial 15-item Teaching Presence rubric based on findings from the literature review (reported above in Tables 1, 2, and 3). Four courses from different programmes of study that fell within the specialist disciplines represented by the team members were then selected by all five team members to pilot test this initial rubric. Before testing the rubric, permission to evaluate each course was obtained from the respective course instructors.

The team then met to comprehensively review the results of the pilot-stage rubric and evaluation procedure. Much of the discussion focused on a small number of rubric items that received different ratings from multiple team members or were flagged as difficult to evaluate. To address the concerns raised, amendments were made to clarify the purpose, meaning, and criteria of these particular items. Extra rubric items were also added to target distinct components of several teaching activities. For example, the "course learning outcomes" item initially developed was broken down into four separate items. At the end of this first stage, the updated rubric with improvements to the criteria wording now totalled 21 items.

In Stage 2, the 21-item rubric was tested independently by all five project members on a new set of eight courses drawn from across the university's five faculties. These courses were distinct from those examined in the previous stage. As with the first stage, permission to do so was obtained from each respective course instructor. During this second stage, a panel of experts convened from within the university reviewed the design to ensure the content validity of the rubric and its items. The panel consisted of three Associate Directors of Education (ADE) for a particular faculty and three experienced learning designers. One-on-one interviews conducted by the project lead and lasting approximately 60 minutes were undertaken with each panel member.

The expert panel review, together with analysis of the stage two rubric results, highlighted the need to refine the wording of several rubric criteria. The review also indicated that the existing "module design" and "assessment design" items each encompassed multiple dimensions. To improve clarity and measurement precision, these two items were separated into distinct criteria that captured their different components. After doing so, the rubric now contained 23 items. Upon completing the project's second stage and expert panel review, the finalised 23-item version of the Teaching Presence rubric was ready to be comprehensively tested. Twelve undergraduate courses from across the five faculties and including first-, second-, and third-year levels, were selected for this purpose.

In addition to testing a larger number of courses, the third and final project stage included a procedural change to reduce the cognitive load on project members. In Stages 1 and 2, all five project members applied the rubric independently to each course. In Stage 3, differently constituted teams of three project members independently applied the rubric to each of 12 courses. Team membership was rotated so that 10 of the 12 courses were evaluated by a different combination of raters. This variation enabled the project team to examine the consistency of ratings across differently constituted teams; substantial disagreement among raters would have indicated that further refinement of the rubric was required. Stage 3 produced 36 completed rubric administrations for analysis. The procedural variations across the three stages are summarised in Table 4.

Project stage	Number of courses tested	Number of raters per course	Number of rubric administrations	Number of rubric items
1	4	5	20	15
2	8	5	40	21
3	12	3	36	23

Table 4. Project stages and procedural details.
Note: Each stage involved a distinct set of courses.
Across the three stages, the rubric was tested on 24 unique courses, resulting in 96 rubric administrations.
Findings from each stage informed the subsequent refinement of the rubric.

Statistical analysis

The rubric criteria described five distinct quality levels (Exemplary to Ineffective). Poorly designed rubrics with ambiguous, lengthy descriptors are often open to misinterpretation, leading to wide score ranges among raters and inaccurate evaluation outcomes. Measures of inter-rater agreement address this concern by statistically testing how closely two or more raters agree when evaluating a particular phenomenon (Quarfoot & Levine, 2016). A high level of agreement provides confidence that the rubric descriptors have been interpreted consistently and accurately by the raters.

Inter-rater agreement for the ordinal rubric ratings on the five-point scale was assessed using Gwet's Agreement Coefficient 2 (Gwet's AC2) and the Real Statistics Resource Pack for Microsoft Excel (Zaiontz, 2023). Gwet's AC2 values range from 0 to 1, where 0 indicates absolute disagreement and 1 indicates perfect agreement. Values below 0.20 are categorised as a poor level of agreement. Gwet's AC2 values between 0.21 and 0.40, 0.41 and 0.60, 0.61 and 0.80, and 0.81 and 1.00 are categorised as fair, moderate, good, and very good to perfect levels of agreement, respectively.

RESULTS

As reported in Table 5, Gwet's AC2 coefficients for the ratings assigned to individual rubric items ranged from 0.218 to 0.956. An item-level appraisal based on the inter-rater agreement analysis produced the following profile:

Very good (VG): Items 1–6, 8, 10–14, 17, 19–20, 23
Good (G): Items 7, 15, 16, 18, 22
Moderate (M): Item 9
Fair (F): Item 21.

Gwet's AC2 coefficients indicated good to very good inter-rater agreement for 21 of the 23 rubric items. Agreement was moderate for Item 9 and fair for Item 21. These two items also had comparatively large standard errors and wide confidence intervals, indicating greater uncertainty in their agreement estimates. The findings suggest that the wording or application of Items 9 and 21 may require further examination in future iterations of the rubric. All remaining items demonstrated good or very good inter-rater agreement.

Rubric item (rated on a five-point scale)	Gwet's AC2	Interpretation of agreement	Standard error for subjects	95% confidence interval for subjects	Standard error overall	95% confidence interval overall
Item 1	0.885	Very good	0.036	0.803; 0.968	0.069	0.732; 1.000
Item 2	0.849	Very good	0.053	0.733; 0.966	0.085	0.663; 1.000
Item 3	0.916	Very good	0.021	0.869; 0.962	0.022	0.868; 0.963
Item 4	0.915	Very good	0.039	0.829; 1.000	0.046	0.812; 1.000
Item 5	0.880	Very good	0.048	0.775; 0.986	0.090	0.683; 1.000
Item 6	0.897	Very good	0.046	0.797; 0.993	0.071	0.740; 1.000
Item 7	0.717	Good	0.085	0.531; 0.903	0.090	0.519; 0.915
Item 8	0.926	Very good	0.036	0.847; 1.000	0.045	0.827; 1.000
Item 9	0.584	Moderate	0.156	0.239; 0.928	0.301	-0.080; 1.000
Item 10	0.916	Very good	0.048	0.812; 1.000	0.053	0.799; 1.000
Item 11	0.828	Very good	0.075	0.664; 0.993	0.132	0.538; 1.000
Item 12	0.843	Very good	0.048	0.737; 0.949	0.122	0.574; 1.000
Item 13	0.840	Very good	0.054	0.722; 0.959	0.094	0.634; 1.000
Item 14	0.943	Very good	0.023	0.892; 0.994	0.029	0.878; 1.000
Item 15	0.647	Good	0.135	0.349; 0.945	0.178	0.256; 1.000
Item 16	0.613	Good	0.091	0.413; 0.813	0.237	0.091; 1.000
Item 17	0.808	Very good	0.049	0.700; 0.902	0.102	0.584; 1.000
Item 18	0.741	Good	0.165	0.377; 1.000	0.260	0.168; 1.000
Item 19	0.851	Very good	0.051	0.739; 0.962	0.056	0.727; 0.975
Item 20	0.839	Very good	0.084	0.656; 1.000	0.145	0.521; 1.000
Item 21	0.218	Fair	0.231	-0.291; 0.728	0.439	-0.748; 1.000
Item 22	0.712	Good	0.081	0.534; 0.900	0.135	0.414; 1.000
Item 23	0.956	Very good	0.016	0.922; 0.990	0.023	0.904; 1.000

Table 5. Gwet's AC2 inter-rater agreement for ratings of rubric items.
Note: Agreement coefficients indicate the consistency of the ratings assigned by raters to each rubric item using the five-point scale. Interpretation thresholds were poor, ≤ 0.20; fair, 0.21–0.40; moderate, 0.41–0.60; good, 0.61–0.80; and very good to perfect, 0.81–1.00.

DISCUSSION

The aim of this research project was to develop a robust theory-based resource to aid the evaluation of course and teaching quality, while also contributing to contemporary scholarship and practice in tertiary learning and teaching. The need for such a resource follows extensive criticism of SET surveys and their use as a standalone evaluation method in Higher Education. Complementing SET with a dissimilar evaluation method can yield a more reliable outcome, as the combined result is not affected by the inherent limitations of a single technique. Such a complementary resource will have widespread appeal given the substantial growth of online and blended programmes in recent years and their importance to the educational strategy of many higher education institutions worldwide (Brown, 2023). A limitation is that the number of fully online and blended courses was not recorded separately. The two delivery contexts were analysed together, so rubric performance could not be compared by delivery mode.

The resource developed from this project, a Teaching Presence rubric (available on request), achieves the stated aim in several ways. Most notably, the instrument is based on a well-regarded theoretical framework explicitly designed for online teaching and learning. The rubric items and criteria were derived from teaching activities identified in the higher education literature rather than from unsupported assumptions about good practice. Its strong theoretical basis distinguishes the rubric from other course and teaching evaluation instruments that are predominantly based on perceptions of good practice.

Contributing to the rubric's efficacy is the meticulous procedure employed in its development. Starting with a careful review of the literature, 15 activities across the three teaching presence functions were identified (see Tables 1, 2, and 3), expanding the scope of activities established by Anderson et al. (2001) in their Teaching Presence framework. The activity descriptions guided the development of the rubric criteria with five distinct quality levels (Exemplary to Ineffective). Administering the rubric over two initial project stages and analysing the experience further improved the criteria wording. The concerted effort to establish clear and concise evaluative criteria is critical to the utility and effectiveness of the rubric method in general (Davidson, 2005) and to higher education course evaluation rubrics specifically (Gregori-Giralt & Menedez-Varela, 2021).

The rubric criteria enable clear decisions about quality, as evidenced by generally strong inter-rater agreement. This feature further underscores the robust approach taken to develop the instrument as few rubrics in the higher education literature had undergone statistical analysis (Yuan & Recker, 2015). Rating scores for all but two of the 23 rubric items achieved a "good" to "very good" level of agreement. The strength of the findings is supported by the number and diversity of courses included in the final testing stage, the use of three independent raters per course, and the rubric's five-point scoring scale. The generally good to very good levels of inter-rater agreement indicate that most rubric items were interpreted and applied consistently across raters. This provides preliminary evidence that the wording and intended meaning of most items were sufficiently clear to support comparable ratings. However, the lower agreement and greater uncertainty associated with Items 9 and 21 indicate that these items may require further examination and refinement.

A further contribution of the project lies in the way the rubric supports collegial and interdisciplinary engagement with teaching practice. Because the instrument makes explicit the design features and pedagogical activities associated with Teaching Presence, it offers more than a technical scoring

framework. It provides a structured basis for peer review, shared interpretation, and reflective discussion about online teaching quality. In this respect, the rubric may help foster the kind of professional community that learning and teaching scholarship depends upon.

Further improvement to the rubric can be achieved by clarifying items 9 and 21, which recorded "moderate" and "fair" levels of agreement, respectively. Both items focus on features of the learning activities presented in the individual course modules. Item 9 targets their design, while Item 21 targets the feedback provided to students. In terms of the criteria for item 21, post-hoc discussions among the project team revealed some uncertainty about the "Limited" grading category and its reference to the quantity and value of feedback provided. It is suggested that future users of the rubric amend the grading category to emphasise and target at least one module learning outcome.

In terms of its practical application, the Teaching Presence rubric offers considerable value. In particular, faculty teams can employ the rubric to evaluate courses identified as underperforming by SET data. This peer approach to course and teaching quality leverages a valuable asset not captured by the SET method: the knowledge, skills, and experience of teaching and learning design professionals. Importantly, the rubric provides faculty teams with considerable purpose and structure by targeting specific elements of course and teaching quality, as informed by the Teaching Presence framework. Used in this way, the instrument can also support staff development by informing reflective conversations, peer-supported course reviews and Performance Review and Planning discussions. In this respect, the rubric may help foster the kind of professional collaborative learning community that learning and teaching scholarship depends upon. This reflects the project team's experience of administering the rubric, particularly the insights gained from reviewing courses outside their disciplines and encountering unfamiliar approaches to teaching presence.

One practical aspect of implementing the rubric is the time required. Administration times for the 36 rubrics collected ranged from 105 to 132 minutes. While this is not considered excessive, given the valuable outcomes achieved, it followed an extensive period of familiarisation. Familiarisation occurred through repeated application of the rubric, team discussion of rating differences and iterative review of the item descriptors during the first two project stages. These activities provided practical training in the rubric's purpose, design and implementation. After familiarising themselves with its items and level descriptors, future first-time users should set aside at least three hours to apply the rubric effectively to a course.

CONCLUSION

In line with its aim, the project has produced a practical, theory-based instrument for evaluating course quality and teaching in online and blended higher education programs. Developed and tested through a rigorous process to ensure validity, reliability, and minimisation of bias, the Teaching Presence rubric offers a valuable non-student perspective to complement SET data. Its strong theoretical grounding in the Community of Inquiry framework, together with the consistency demonstrated through inter-rater reliability testing, distinguishes it from many existing course evaluation rubrics and strengthens its value for higher education practice.

Beyond its evaluative function, the rubric contributes to tertiary learning and teaching by supporting structured peer review, professional reflection, and course renewal. It provides educators, instructional designers, and faculty teams with a shared framework for discussing the

quality of online and blended learning environments in explicit, theory-informed, and practically applicable ways. This is particularly important in a sector where student evaluation data often remains dominant despite its limitations, and where there is a growing need for complementary approaches that can better support scholarly teaching practice.

Conceptually, the rubric translates observable teaching practices and relational dimensions of teaching presence into a structured framework for evaluation. Its value lies in supporting consistent assessment while retaining the professional judgement, collegial interpretation and contextual understanding required to evaluate teaching practice. Ultimately, the rubric offers institutions, instructors, and instructional designers a valuable resource for evaluating and improving the quality of online and blended courses, while also contributing to broader conversations in learning and teaching scholarship about how educational quality is understood, discussed, and enhanced in higher education.

Patrick Gillett (https://orcid.org/0000-0002-7697-2426) is a Lecturer (Teaching Scholar) with the Faculty of Business, Law, and Arts at Southern Cross University. Senior positions held within the faculty include MBA Director and Academic Integrity Officer. His research interests focus on the use of educational technologies.

Tina van Eyk (https://orcid.org/0000-0001-9770-8133) is an Adjunct Fellow with the Centre for Teaching and Learning at Southern Cross University and has worked as a learning designer, associate lecturer and project coordinator. Tina holds a Master of Online Education and her research interests include designing online and blended learning experiences and flexible professional learning.

Patrick Bruck (https://orcid.org/0000-0002-2848-8840) is a Lecturer at Southern Cross University, Faculty of Education, Australia. He is enrolled in a Doctor of Education (EdD) at Central Queensland University with a focus on the teaching of proportional reasoning in Australian schools. Patrick has held numerous roles in Australia's education sector, including as a teacher, Head of Department, Deputy Principal and Principal in high schools, predominantly in Queensland. He has also taught internationally in Asia and Europe.

Carolyn Seton's (https://orcid.org/0000-0002-2566-5389) research interests span computing education, GenAI in education, English as an Additional Language learners, cognitive load, and sense of belonging in STEM. Her teaching experience encompasses introductory programming, systems analysis, database, and application development, with a focus on inclusive, evidence-based pedagogical practice.

Kylie Day (https://orcid.org/0000-0003-0202-7808) is a Dharawal Aboriginal scholar and Senior Lecturer based on Bundjalung Country in northern New South Wales. She coordinates the Bachelor of Indigenous Knowledge at Southern Cross University. Her research engages with Indigenous Knowledge systems, decolonising pedagogies, and the Scholarship of Learning and Teaching.

REFERENCES

Akyol, Z., Vaughan, N., & Garrison, D. R. (2011). The impact of course duration on the development of a community of inquiry. *Interactive Learning Environments, 19*(3), 231–246. https://doi.org/10.1080/10494820902809147

Alderman, L., Towers, S., & Bannah, S. (2012). Student feedback systems in higher education: A focused literature review and environmental scan. *Quality in Higher Education, 18*(3), 261–280. https://doi.org/10.1080/13538322.2012.730714

Anderson, T., Rourke, L., Garrison, D. R., & Archer, W. (2001). Assessing teaching presence in a computer conferencing context. *Journal of Asynchronous Learning Networks, 5*(2), 1–17. https://doi.org/10.24059/olj.v5i2.1875

Arbaugh, J. (2008). Does the Community of Inquiry Framework predict outcomes in online MBA courses? *International Review of Research in Open and Distributed Learning, 9*(2), 1–21. https://doi.org/10.19173/irrodl.v9i2.490

Bearman, M., Dawson, P., Boud, D., Bennett, S., Hall, M., & Molloy, E. (2016). Support for assessment practice: Developing the Assessment Design Decisions Framework. *Teaching in Higher Education, 21*(5), 545–556. https://doi.org/10.1080/13562517.2016.1160217

Biggs, J. (1996). Enhancing teaching through constructive alignment. *Higher Education, 32*(3), 347–364. https://doi.org/10.1007/BF00138871

Biggs, J. (2014). Constructive alignment in university teaching. *HERDSA Review of Higher Education, 1*, 5–22. https://www.tru.ca/__shared/assets/Constructive_Alignment36087.pdf

Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., van Haeringen, K., Saddiqui, S., & Rozenberg, P. (2019). Contract cheating and assessment design: Exploring the relationship. *Assessment and Evaluation in Higher Education, 44*(5), 676–691. https://doi.org/10.1080/02602938.2018.1527892

Brown, S. (2023). Online, distance, blended: It's all just education. In S. Brown, L. Amrane-Cooper, D. Baume, S. Hatzipanagos, P. Powell, S. Sherman & A. Tait (Eds.), *Online and distance education for a connected world* (pp. 1–12). UCL Press. https://doi.org/10.2307/j.ctv2wk7261.8

Caskurlu, S. (2018). Confirming the subdimensions of teaching, social, and cognitive presences: A construct validity study. *The Internet and Higher Education, 39*, 1–12. https://doi.org/10.1016/j.iheduc.2018.05.002

Caskurlu, S., Maeda, Y., Richardson, J. C., & Lv, J. (2020). A meta-analysis addressing the relationship between teaching presence and students' satisfaction and learning. *Computers and Education, 157*, Article 103966. https://doi.org/10.1016/j.compedu.2020.103966

Caskurlu, S., Richardson, J. C., Maeda, Y., & Kozan, K. (2021). The qualitative evidence behind the factors impacting online learning experiences as informed by the community of inquiry framework: A thematic synthesis. *Computers and Education, 165*, Article 104111. https://doi.org/10.1016/j.compedu.2020.104111

Chi, M. T. H., & Wylie, R. (2014). The ICAP Framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist, 49*(4), 219–243. https://doi.org/10.1080/00461520.2014.965823

Cunningham, S., Laundon, M., Cathcart, A., Bashar, M. A., & Nayak, R. (2023). First, do no harm: Automated detection of abusive comments in student evaluation of teaching surveys. *Assessment and Evaluation in Higher Education, 48*(3), 377–389. https://doi.org/10.1080/02602938.2022.2081668

Curby, T., McKnight, P., Alexander, L., & Erchov, S. (2020). Sources of variance in end-of-course student evaluations. *Assessment and Evaluation in Higher Education, 45*(1), 44–53. https://doi.org/10.1080/02602938.2019.1607249

Davidson, E. J. (2005). *Evaluation methodology basics: The nuts and bolts of sound evaluation*. Sage Publications.

Esarey, J., & Valdes, N. (2020). Unbiased, reliable, and valid student evaluations can still be unfair. *Assessment and Evaluation in Higher Education, 45*(8), 1106–1120. https://doi.org/10.1080/02602938.2020.1724875

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research, 83*(1), 70–120. https://doi.org/10.3102/0034654312474350

Garivaldis, F., McKenzie, S., Henriksen, D., & Studente, S. (2022). Achieving lasting education in the new digital learning world. *Australasian Journal of Educational Technology, 38*(4), 1–11. https://doi.org/10.14742/ajet.8331

Garrison, D., & Akyol, Z. (2013). The Community of Inquiry theoretical framework: In the context of online and blended learning. In M. G. More (Ed.), *Handbook of distance education* (2nd ed., pp. 104–119), Routledge. https://doi.org/10.4324/9780203803738

Garrison, D. R., Anderson, T., & Archer, W. (2000). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education, 2*(2), 87–105. https://doi.org/10.1016/S1096-7516(00)00016-6

Gregori-Giralt, E., & Menėndez-Varela, J-L. (2021). The content aspect of validity in a rubric-based assessment system for course syllabuses. *Studies in Educational Evaluation, 68*, Article 100971. https://doi.org/10.1016/j.stueduc.2020.100971

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81–112. https://doi.org/10.3102/003465430298487

Heffernan, T. (2022). Sexism, racism, prejudice, and bias: A literature review and synthesis of research surrounding student evaluations of courses and teaching. *Assessment and Evaluation in Higher Education, 47*(1), 144–154. https://doi.org/10.1080/02602938.2021.1888075

Henderson, M., Ryan, T., & Phillips, M. (2019). The challenges of feedback in higher education. *Assessment and Evaluation in Higher Education, 44*(8), 1237–1252. https://doi.org/10.1080/02602938.2019.1599815

Jonsson, A. (2013). Facilitating productive use of feedback in higher education. *Active Learning in Higher Education, 14*(1), 63–76. https://doi.org/10.1177/1469787412467125

Kelly, K., & Zakrajsek, T. D. (2021). *Advancing online teaching: Creating equity-based digital learning environments*. Routledge. https://doi.org/10.4324/9781003442936

Lakeman, R., Coutts, R., Hutchinson, M., Massey, D., Nasrawi, D., Fielden, J., & Lee, M. (2022). Stress, distress, disorder and coping: The impact of anonymous student evaluation of teaching on the health of higher education teachers. *Assessment and Evaluation in Higher Education, 47*(8), 1489–1500. https://doi.org/10.1080/02602938.2022.2060936

Lee, J. E., Recker, M., & Yuan, M. (2020). The validity and instructional value of a rubric for evaluating online course quality: An empirical study. *Online Learning, 24*(1). https://doi.org/10.24059/olj.v24i1.1949

Lipsey, N., & Shepperd, J. (2021). Examining strategies to increase student evaluation of teaching completion rates. *Assessment and Evaluation in Higher Education, 46*(3), 424–437. https://doi.org/10.1080/02602938.2020.1782343

Liu, X., Magjuka, R. J., Bonk, C. J., & Lee, S. (2007). Does sense of community matter? An examination of participants' perceptions of building learning communities in online courses. *Quarterly Review of Distance Education, 8*(1), 9–24. https://www.proquest.com/docview/231180604/8AB92EC26004C59PQ/2?accountid=16926

Lizzio, A., & Wilson, K. (2008). Feedback on assessment: Students' perceptions of quality and effectiveness. *Assessment and Evaluation in Higher Education, 33*(3), 263–275. https://doi.org/10.1080/02602930701292548

Macquarie University. (2022, April 7). *Unit descriptions*. https://lt.arts.mq.edu.au/curriculum- design/constructive-alignment/writing-unit-descriptions/

Martin, F., Bolliger, D. U., & Flowers, C. (2021). Design matters: Development and validation of the Online Course Design Elements (OCDE) Instrument. *International Review of Research in Open and Distance Learning, 22*(2), 46–71. https://doi.org/10.19173/irrodl.v22i2.5187

Martin, F., Wang, C., & Sadaf, A. (2018). Student perception of helpfulness of facilitation strategies that enhance instructor presence, connectedness, engagement and learning in online courses. *The Internet and Higher Education, 37*, 52–65. https://doi.org/10.1016/j.iheduc.2018.01.003

Mayer, R. E. (2019). Thirty years of research on online learning. *Applied Cognitive Psychology, 33*, 152–159. https://doi.org/10.1002/acp.3482

Mehrabian, A., (1969). Some referents and measures of nonverbal behavior. *Behavior Research Methods and Instrumentation, 1*(6), 203–207. https://doi.org/10.3758/BF03208096

McInnes, R., Hobson, J. E., Johnson, K. L., Cramp, J., Aitchison, C., & Baldock, K. L. (2024). Online course quality evaluation instruments: A scoping review. *Australasian Journal of Educational Technology, 40*(2), 55–75. https://doi.org/10.14742/ajet.8978

Nargundkar, S., & Shrikhande, M. (2014). Norming of student evaluations of instruction: Impact of non-instructional factors. *Decision Sciences Journal of Innovative Education, 12*(1), 55–72. https://doi.org/10.1111/dsji.12023

Oliver, M., & Shaw, G. P. (2003). Asynchronous discussion in support of medical education. *Journal of Asynchronous Learning Networks, 7*(1). https://doi.org/10.24059/olj.v7i1.1863

Popenici, S., & Millar, V. (2015). *Writing learning outcomes: A practical guide for academics*. Melbourne Centre for the Study of Higher Education. The University of Melbourne. www.melbourne- cshe.unimelb.edu.au/__data/assets/pdf_file/0007/2296861/MCSHE-Learning-Outcomes-Guide-web- Nov2015-rev2021.pdf

Quality Matters. (2023). *Higher Education rubric, seventh edition*. https://www.qualitymatters.org/qa-resources/rubric-standards/higher-ed-rubric

Quarfoot, D., & Levine, R. A. (2016). How robust are multirater interrater reliability indices to changes in frequency distribution? *The American Statistician, 70*(4), 373–384. https://doi.org/10.1080/00031305.2016.1141708

Ramlatchan, M., & Watson, G. S. (2020). Enhancing instructor credibility and immediacy in online multimedia designs. *Educational Technology Research and Development, 68*(1), 511–528. https://doi.org/10.1007/s11423-019-09714-y

Richardson, J. T. E. (2005). Instruments for obtaining student feedback: A review of the literature. *Assessment and Evaluation in Higher Education, 30*(4), 387–415. https://doi.org/10.1080/02602930500099193

Richardson, J. C., Arbaugh, J. B., Cleveland-Innes, M., Ice, P., Swan, K. P., & Garrison, D. R. (2012). Using the Community of Inquiry framework to inform effective instructional design. In L. Moller & J. B. Huett (Eds.), *The next generation of distance education: Unconstrained learning* (pp. 97–125). Springer. https://doi.org/10.1007/978-1-4614-1785-9_7

Schouten, L. M., Hulscher, M. E., van Everdingen, J. J., Huijsman, R., & Grol, R. P. (2008). Evidence for the impact of quality improvement collaboratives: Systematic review. *BMJ, 336*(7659), 1491–1494. https://doi.org/10.1136/bmj.39570.749884.BE

Shea, P., Fredericksen, E., Pickett, A., & Pelz, W. (2003). A preliminary investigation of 'teaching presence' in the SUNY Learning Network. *Elements of Quality Online Education: Practice and Direction, 4*(2), 279–290. https://doi.org/10.24059/olj.v7i2.1856

Stone, C., & Springer, M. (2019). Interactivity, connectedness and 'teacher presence': Engaging and retaining students online. *Australian Journal of Adult Learning, 59*(2), 146–169.

SUNY. (2023, December 10). *The SUNY Online Course Quality Review Rubric OSCQR*. https://oscqr.suny.edu/

University of New South Wales. (2022, August 8). *Learning outcomes: What is a learning outcome?* https://www.teaching.unsw.edu.au/learning-outcomes

Vaughan, N., Cleveland-Innes, M., & Garrison, D. (2013). *Teaching in blended learning environments: Creating and sustaining communities of inquiry*. AU Press. https://doi.org/10.15215/aupress/9781927356470.01

Wells, S., Tamir, O., Gray, J., Naidoo, D., Bekhit, M., & Goldmann, D. (2018). Are quality improvement collaboratives effective? A systematic review. *BMJ Quality & Safety, 27*(3), 226–240. https://doi.org/10.1136/bmjqs-2017-006926

Willems, J., & Bossu, C. (2012). Equity considerations for open educational resources in the glocalization of education. *Distance Education, 33*(2), 185–199. https://doi.org/10.1080/01587919.2012.692051

Yuan, M., & Recker, M. (2015). Not all rubrics are equal: A review of rubrics for evaluating the quality of open educational resources. *International Review of Research in Open and Distributed Learning, 16*(5), 16–38. https://doi.org/10.19173/irrodl.v16i5.2389

Zaiontz, C. (2023, February 18). *Real statistics resource pack software* (Release 8.5). https://real- statistics.com/free-download/real-statistics-resource-pack/

GUIDELINES FOR CONTRIBUTORS

Submissions for *Scope (Learning & Teaching)* should be sent by 30 April for review and potential inclusion in the annual issue to: megan.kitching@op.ac.nz or trish.chaplin-cheyne@op.ac.nz.

Please consult the information for contributors below and hardcopy or online versions for examples.

All submissions will be peer reviewed. Peer review comments will be sent to all submitters in due course, with details concerning the possible reworking of documents where relevant. All submissions must include disclosure of whether and how AI was used in writing the work. All final decisions concerning publication of submissions will reside with the Editors. Opinions published are those of the authors and not necessarily subscribed to by the Editors or Otago Polytechnic.

Contributors retain copyright in their submissions and must obtain permission for the use of any material that is not their own included in their submission. Contributors grant the publishers permission to deposit the published work in our institutional repository.

Except for images, mātauranga Māori, and where otherwise indicated, content will be made available as open access on a Creative Commons Attribution license 4.0. 

All images are copyright to the respective artists and photographers.

Mātauranga Māori is a significant component of Aotearoa/New Zealand's heritage, and sharing mātauranga Māori facilitates inter-cultural dialogue and understanding that is in the national interest. However, we recognise that the originating Māori community and/or individual has the primary interest as kaitiaki over the mātauranga and we are therefore committed to ensuring that the sharing, promotion and innovation based on mātauranga Māori respects and enhances its cultural and spiritual integrity, as well as that of the originating community and/or individual.

Submission formats include: editorials, articles, perspectives, essays, research notes and work in progress. Other formats will also be considered.

High standards of writing, proofreading and adherence to consistency through the APA referencing style are expected. For more information, please refer to the APA Publication Manual, 7th edition; and consult prior issues for examples. A short biography of no more than 50 words; as well as title; details concerning institutional position and affiliation (where relevant); contact information (postal, email and telephone number) and ORCID number should be provided on a cover sheet, with all such information withheld from the body of the submission. Low resolution images with full captions should be inserted into texts to indicate where they would be preferred, while high resolution images should be sent separately. Enquiries about future submissions can be directed to megan.kitching@op.ac.nz.