

## Research Article

# Peer-to-Peer Evaluation: Kritik Pilot at Two Post-secondary Canadian Institutions

Kimberley D. Ryan<sup>1\*</sup> and Kristen A. Hardy<sup>2</sup>

<sup>1</sup>Brandon University, Faculty of Health Studies, Department of Psychiatric Nursing, Canada

<sup>2</sup>Contract Instructor, University of Winnipeg and Brandon University, Canada

\***Corresponding author:** Kimberley D. Ryan, Associate Professor, Brandon University, Faculty of Health Studies, Department of Psychiatric Nursing; 270- 18<sup>th</sup> Street , Brandon, Manitoba, R7A 6A9, Canada; **Tel:** (204) 727 9628; **Fax:** (204) 571 8568

**Received:** December 18, 2022; **Accepted:** December 20, 2022; **Published:** December 30, 2022

## Abstract

In 2021/2022, educators at two Canadian universities chose to use a novel pedagogical approach to undergraduate program learning and instruction. Kritik, an on-line platform was selected because of its reputation for encouraging peer-to-peer learning and evaluation through use of Bloom's Taxonomy of Learning to promote critical thinking skill development in learners. In this article, we outline our experiences while piloting use of Kritik, provide a critical assessment of its use and share our pedagogical lessons learned.

**Keywords:** Peer evaluation, Peer feedback, Undergraduate learning and instruction, Kritik

## Introduction

The inspiration for this article originated in 2021/2022 after our pilot of Kritik, an on-line, peer-to-peer learning and evaluation platform used for the purpose of promoting student engagement in learning. Kritik allowed us to step away from being the primary source of all knowledge, and to instead use a collaborative educational approach to facilitate the learning process. In this article, we share our experiences, provide a critical assessment of Kritik use, and outline our pedagogical lessons learned.

## Explication of Concepts

A *peer* has been defined to be one who is of equal standing with another, especially based on age, grade, or status [1]. In this article, a *peer* is understood to be an undergraduate student. Despite their similarities and often being used interchangeably, the terms *review*, *assessment* and *evaluation* have distinct meanings. A *review* is identified to be the act of giving one's opinion about the quality of a product, and *assessment* is a systematic basis for making inferences about student learning and development, while *evaluation* is the determination of the value, nature, character, or quality of something [1]. *Feedback* is the transmission of evaluative or corrective information about an action, event, or process to the original or controlling source [1].

The terms *peer feedback* and *peer evaluation* are used for the purposes of this article. *Peer feedback* is the "rich detailed comments" [2 p. 280] given by peer student evaluators and communicated to students with the aim of providing "guidance on the quality of the student's work and understanding of the subject matter" [3 p. 753]. *Peer feedback* is given during the process of *peer evaluation*; an organized, systematic process wherein students use a rubric,

a standardized measurement tool to evaluate peer assignment compositions completed on the Kritik platform.

## Background

We teach undergraduate courses at two Canadian universities located in the province of Manitoba. Kristen teaches DIS-2200/WGS- 2264, Disabilities, Sexualities and Rights, cross-listed between the Disability Studies Program and Women's and Gender Studies Department, as well as REL-2570, Sexuality in the Religious Context, in the Religion and Culture Department at the University of Winnipeg. I teach both 69:161, an introductory interpersonal communication course and 69:473, a fourth-year philosophical perspectives in psychiatric nursing course at Brandon University.

Over the past two decades, I have incorporated the use of peer evaluation into my courses. One of my previously taught fourth-year courses involved students completing peer evaluation of their seminar presentations. The assignment mark was derived from an average of peer evaluation scores and my own. In a third-year course, students conducted peer evaluation of a case study analysis, completed, and posted on an on-line platform by each student. In this case, too, the assignment mark was based on an average between peer review scores and mine. In yet another course, this time an introductory course, after viewing an assigned video, students posted their responses to questions about the video to an on-line platform for peer evaluation by fellow classmates. The assignment mark was again based on an average of student peer review scores and mine.

It is noteworthy to share that over the years, I have repeatedly encountered significant resistance to peer evaluation assignments from several students. Course end student evaluation included comments suggesting that the majority of opposition was aligned with

assumptions and beliefs that use of peer evaluation is a mechanism for offloading the course instructor's work onto students.

Kristen had previously experimented with incorporating peer evaluation into an online REL-2507 course, using the 'Discussions' feature of the D2L/Brightspace platform. Having sorted students into groups, they were asked to post their near-final drafts of a visual-art analysis assignment on-line. Each group member was instructed to carefully read the other members' submissions, and to offer actionable feedback. Completing all (four or five) of one's peer assessments with evidence of a good degree of care and effort earned each student 5% of their final mark, though students' drafts themselves were not graded. Students then made any changes they wished to their drafts and submitted the final papers to the instructor for assessment. The finished paper was worth 15% of students' final grade, and assessed in a conventional, holistic way by the instructor and teaching assistant (TA). One of the limitations of this peer-evaluation activity was that the process could not be made anonymous, limiting the critical feedback some students were comfortable providing. Overall, this activity resulted in respectful and supportive engagement with peers' work, and a reasonably high caliber of finished work, and encouraged Kristen's interest in using peer evaluation in future courses.

### **Kritik**

Kritik, a novel on-line platform developed with the aim of encouraging peer-to-peer learning and evaluation, integrates the use of Bloom's Taxonomy of Learning to promote critical thinking skill development in learners [4]. A rubric developed by the course instructor, in consultation with a Kritik design specialist, accompanies each peer-to-peer learning & evaluation activity. This rubric, comprised of evaluative criteria and specific criterion measures, serves as a frame of reference for students while conducting peer review and evaluation.

Each peer-to-peer learning and evaluation activity consists of *create*, *evaluate*, and *feedback* stages [4]. In the *create* stage, students follow the activity instructions, adhere to the criteria outlined in the corresponding rubric, and *create* an activity composition that is uploaded to the Kritik platform. During the *evaluate* phase, students provide both formative and summative feedback by evaluating assigned peers' activity compositions using the corresponding rubric and substantiate their assigned peer evaluation scores with written comments intended to be helpful when students complete their next Kritik composition. The *feedback* stage involves students providing each peer evaluator with feedback about how helpful their comments were [4]. Weighting is assigned to each of the three stages through use of an algorithm calibrated by a Kritik design specialist in consultation with the course instructor. A score is assigned for each of the three stages using the established algorithm with the overall Kritik activity mark being the combined score of all three stages

### **Motivations for Using Kritik**

We do not believe that the course instructor should be the central figure in any given course; achievement of sustained, quality learning is a cooperative task. Moreover, we opine that learning is less likely to occur if students are passive recipients and not actively engaged within

learning and instruction. As such, we strive to create student-centered learning environments with students actively engaged with one another while we provide guidance in the accomplishment of course learning objectives. One of the most significant advantages of Kritik, from this perspective, is that it keeps the learning process student centered. Although instructors are available to 'step in' to provide guidance with peer evaluations, or to override marks assigned by peer evaluators in any/all three stages, students interact primarily with each other. Further, if Kritik is used for a sequence of assignments in a course, students are teaching and learning from one another and so doing are engaged in an iterative process. Kritik levels the playing field wherein students are on somewhat more-even footing with the course instructor. Rather than being fundamental to the learning process, we consider ourselves to be facilitators of learning, readily available to assist and/or intervene if needed.

One of the more 'pragmatic' benefits of Kritik is the flexibility it offers to instructors in managing their grading work. While the platform does not completely take assessment duties away from the instructor, it does allow the instructor to choose how intensively they wish to be involved in the process, ranging from engaging only when students disclose a concern with a peer evaluation by submitting a 'grade dispute', to evaluating all Kritik compositions as well as commenting on and/or adjusting the score for students' work themselves. For those educators dealing with large grading loads and minimal teaching assistance, this is a decidedly attractive feature.

### **Kritik Implementation**

During the 2021/2022 academic year, Kristen piloted Kritik for DIS-2200/WGS-2264 and REL-2507 during Fall 2021 term, for DIS-2200/WGS-2264 again in the Winter 2022 term, and once more for the latter course in the condensed Spring 2022 term; and I piloted Kritik in 69:161 and 69:473 during Fall 2021 term as well as 69:161 again in the Winter 2022 term.

Both Kristen and I had some lead time to familiarize ourselves with the platform. We each worked closely with Kritik's instructional designers and technical-support people to navigate Kritik's range of features and to become familiar with the processes involved in using Kritik from both the faculty and student perspectives. Despite this support, and although we had prior experience with peer evaluation as well as a significant degree of confidence with design/development of grading matrices, we both encountered steep learning curves. In saying this, it is important to acknowledge that Kritik's support team were indispensable supports in our onboarding including platform set-up and initial use.

Kristen opted to use Kritik in REL-2507 for 'two and a half' of the written assignments. Specially, the initial draft (but not final paper) of a visual-art analysis activity, the final version of the course's media analysis and film analysis assignments. Each assignment component required students to apply their course-derived knowledge to analyze, respectively, a visual image, a media story, and a documentary film. In her first term of using Kritik in DIS-2200/WGS-2264, Kristen chose to use Kritik for most of the written assignments (weekly response papers, plus two film analyses). After receiving mixed student reviews,

adjustments (described below) were made to Kritik use in subsequent terms.

In both 69:161 sections of my course, students completed five Kritik based learning activities designed to increase writing proficiency. Two weeks completion time was allocated to each of the five activities, with six days to *create* the composition, five days for the peer *evaluation* and three days for the *feedback* stages. A different rubric accompanied each of the five Kritik based activities. 69:474 involved short essay compositions to address a different philosophical question for each of five Kritik based activities. The same time parameters used for 69:161, outlined above, were used for this course but all Kritik based activities had the same corresponding rubric.

Kristen and I made a point of soliciting frequent feedback from students about their experiences with Kritik. We both expressed an interest in receiving feedback in several asynchronous lecture videos and distributed reminders in our weekly email updates to students. Ultimately, it was the time spent during synchronous online class inviting students to discuss and share concerns about Kritik that we found to be most productive.

### **Pedagogical Issues/Concerns**

One of my reasons for deciding to use Kritik was its assurance of student anonymity throughout the peer evaluation process. Following completion of the first Kritik activity in 69:474, I received notification from several students that viewing the document 'properties' of uploaded Microsoft Word files permitted identification of the students for whom they were completing peer evaluations. After notifying Kritik technical support of this issue, I was advised that use of third party, Microsoft Word was the issue; that anonymity could only be assured by a copy and paste of student compositions directly into the Kritik application text box. I was also advised that disabling the file type submission option would prevent Word file uploads and ensure that students could only submit a composition via the copy/paste feature.

However, this copy/paste option was not without issue. None of the original formatting was retained after composition copy/paste. Those compositions that were not edited following copy/paste were difficult for peers to evaluate. Specifically, students shared difficulties with determining where paragraph breaks should be. Ultimately students disclosed their frustration with the additional time and focus required to complete an accurate peer review, when formatting was asked. Therefore, prior to third 69:474 assignment commencement, I advised all students of the necessity to pay particular attention to their composition formatting, and the need to edit submissions, after copy/paste completion within the Kritik application textbox.

In both my 69:161 Fall and Winter term offerings, following completion of the initial assignment, several students shared their lack of understanding that assignment completion involved participation in the three Kritik stages, others indicated they did not realize that they needed to evaluate all peer compositions assigned to them and many disclosed their duress with the requirement of learning a new platform in a very short time while being concurrently required to complete an assignment for marks.

In both 69:161 sections, 10 of 35 and five of 26 students respectively did not complete the first two Kritik activity compositions by the due date/time. Acceptance of late submissions affects the assignment of completed compositions for peer review, and I believed it unfair for those students who completed their peer evaluation in a timely manner to be assigned late composition(s) to evaluate. Therefore, I made the decision not to permit late composition submissions. Students were notified of this decision from the outset of course commencement. However, in circumstances where students notified me early of their failure to complete the assignment by due date/time, I did allow these students to proceed with completing the peer review and evaluation stages. In these instances, I assigned partial marks that reflected the effort each student had invested during completion of these last two Kritik stages.

All students in both 69:161 sections and 69:474 were assigned five peer compositions to evaluate. While reviewing the first activity compositions and corresponding peer evaluations, I noted significant variation in peer evaluation scores assigned by the 69:474 students. Specifically, 10 of 36 peers had assigned full marks for many if not all rubric criteria, while other peers who had been assigned the same student compositions to evaluate, had allocated less than half mark for the same rubric criteria. Upon further examination, I noted the peer evaluations with less than full marks assigned were closely aligned with my evaluations. To ensure that use of Kritik was accurate and therefore credible, I adjusted the composition scores for students who had received inaccurate peer evaluation scores and reduced the evaluation scores of the students who completed inaccurate evaluations. All students involved were apprised of my reasons for over-riding the previously assigned scores and in particular the students who had completed inaccurate evaluations were informed of the action required by them to ensure that I did not have to over-ride scores again.

One of Kristen's concerns, as a Disability Studies instructor and self-identified disabled person, was whether the inflexible deadlines within Kritik would create an accessibility issue for some students, including those with formal accommodations around time allowances for completing assignments, and those who encountered unexpected illnesses or personal/family circumstances. Outside of Kritik, Kristen has maintained a flexible policy of granting extensions on most assignments, especially if requested in advance. Somewhat surprisingly, despite having more than 100 students in the three courses combined, accessibility rarely became an issue. While not obvious to students, work could still be submitted late (with the instructor's approval) via Kritik's on-line help/support chat feature. If the work was uploaded before the deadline for the 'evaluation' stage, there was a good chance of its distribution for peer-assessment to those students who have left their 'evaluate' tasks until closer to that phase's due date/time. Even papers uploaded too late to be peer evaluated can be assessed within Kritik by the instructor.

For the most recent offering of DIS-2200/WGS-2264, Kristen experimented with allowing late submissions automatically. In this case, Kritik requires students to provide an explanation for the delay in their submissions, and then allows the instructor to accept or

reject the paper. In a class of approximately 40 students, this typically resulted in only one or two late submissions for a given activity, with improved odds of 'slightly' late papers being distributed to peers for evaluation because of the lack of a need to obtain pre-approval and Kritik's assistance to upload late work.

Partway through each section of DIS-2200/WGS-2264, during discussions with students about Kritik, some of Kristen's strongest students expressed misgivings about having their work assessed by peers, and their inability to score 'perfect' marks, given the way Kritik's algorithm calculates overall activity marks. A few students showed interest in receiving more ungraded, early term training in assessing peers' work accurately in Kritik. Some students also expressed concerns that their peers were marking too harshly. During one conversation, some students disclosed the belief that because one of the metrics during the evaluation stage asked them to rate how 'critical' aspects of the peer evaluation were, some believed they needed to "find faults" in the peer evaluations. Kritik's use of the term "critical" refers to how helpful and actionable the peer evaluation was for students [4]. This misinterpretation of the term "critical" was addressed during asynchronous online class and by email.

Kritik's 'grade dispute' function permits students to notify the course instructor of any peer evaluation concerns. Such 'disputing' promotes more focused checking by the course instructor of student work with dispute resolution accomplished by the instructor prior to finalizing activity grades. Repeatedly encouraging all students to use Kritik's 'grade dispute' function to flag any concerns about peer evaluations or make comments for the instructor on a specific paper also reduced in-class complaints about inaccurate grading. Both Kristen and I think that the presence of this function builds students' trust. For example, if students think their peers' evaluations are too low, upon notification of a grade dispute, the instructor or TA can focus in on the issue to reassess, and perhaps regrade, the work. In response to concerns over lower-than-expected marks by midterm during her first term using Kritik, Kristen emailed her students and requested that they "err on the side of generosity" in their evaluations. This tipped the subsequent marks noticeably upwards, even somewhat beyond what Kristen would have assigned had she evaluated students' work herself. Revision of rubrics in subsequent terms was helpful in preventing inappropriately low averages.

Also worrisome were the views, expressed by several students in the Fall DIS-2200/WGS-2264-section course evaluations that use of Kritik reflected "laziness" on the part of the instructor; that because the instructor's spot-checking papers was not visible to students unless marks were changed, or additional comments left, students perceived they were doing this work with no oversight. These concerns were addressed through changes made in implementing Kritik in the following winter and spring-term sections of the same course, with both a reduction in the number of course assignments for which Kritik was used, and instructor comments left on all spot-checked papers; a process that involved approximately a third of submitted papers each week, while recording which papers had been checked in any given week.

## **Pedagogical Lessons Learned**

My lessons learned include the necessity to include both a Kritik orientation for students and a strategy to promote student anonymity, as well as provision of clear expectations about the scoring and commentary required of peers during the evaluation stage. I did not anticipate the need for providing an orientation to Kritik at the beginning of my courses. Based on student concerns and my acquired insight into the difficulties associated with lack of familiarity with the Kritik platform, I believe that during the first week of coursework all students should participate in a mandatory orientation to Kritik that includes completion of a mock Kritik peer review assignment with no marks assigned.

Anonymity cannot be assured if students upload a composition, completed in a format such as MS Word, to Kritik. To ensure that anonymity is maintained throughout the course, all compositions must be uploaded to the Kritik application text box and the file type submission option disabled by course instructor. Students need to be advised of the requirement to edit composition(s) after upload to ensure assignments are formatted appropriately in advance of peer evaluation.

Some peers failed to exercise due diligence when evaluating composition(s) and instead assigned inaccurate score(s). To ensure accurate assignment of scores, it is crucial that a timely and thorough review be conducted of all peer evaluation commentary and score(s) assigned, particularly when significant variation exists among the assigned peer evaluation scores. In circumstances of inaccurate score(s) allocation, fairness in grading is assured by adjustment of Kritik assigned scores for both student who completed the composition and the peer who conducted the inaccurate evaluation.

Kristen's experiences also led to several conclusions about how to better implement Kritik. Having every written assignment for a course submitted and assessed via Kritik was not a popular option with students; maintaining at least one major activity as instructor-graded seems to have satisfied students who were concerned over the degree of instructor involvement in the process. Regular spot-checking with at least brief comments acknowledging an instructor's or TA's reading of the work has also made a significant difference with students' expressions of dissatisfaction. Instructor visibility and responsiveness to disputes, combined with reassurance and proactive training and support of students in the use of the platform, are perhaps the leading factors noted by Kristen in improving students' receptiveness to the use of Kritik.

Kristen is of the opinion that there are also some learning activities for which Kritik simply does not work very well. Specifically, for activities such as the REL-2507 film analysis that involves a set of common questions for each student on the same media, peer evaluation resulted in additional frustration. Students experienced less variety and more repetition in the materials they were required to assess, while offering less 'reward' in terms of exposure to different perspectives and ideas. It is also very challenging to design an effective rubric for assessing multiple questions effectively within the context of an individual activity, since only one rubric per activity may be used, and any marks assigned apply to the entire submission.

Although instructors have the option to adapt one of Kritik's rubric templates, for those instructors who decide to design their own, a significant degree of effort is required to design an effective rubric. Students, most of whom will be new to doing rubric-based assessment themselves, require a rubric that is clear, minimally complex, explains the specifics that students should use to determine assignment of scores for each criterion and an explanation of how peer evaluation scores align with the course grading scale. Since the rubric also needs to guide students on how to evaluate solid-but-not-exceptional papers, determining whether a composition that addresses all required criteria quite well should be assigned full marks, or not, is also important. Investing additional effort into this stage translates into more-accurate marks, less frustration, and complaints among peer evaluators, as well as less revisions to peer scoring by the instructor.

One area that remains a concern when using Kritik is identifying plagiarism. Despite being encouraged to contact us in situations where peer evaluators suspected plagiarism or other forms of academic dishonesty, no students did so. In contrast, spot-checking by both of us did reveal some instances of plagiarism.

Kritik representatives have shared plans for integration with the plagiarism-detection software such as Turnitin which may be of use for some courses. However, as is the case for traditional assessment methods, the easy availability of new automatic 'paraphrasing' applications to circumvent automated detection systems, is still a concern. Ultimately, diligent spot-checking by the instructor seems to be the most-effective way to monitor, detect, and correct academic integrity issues.

### **Use of Kritik in the Future**

Based on our experiences thus far and with consideration of our lessons learned, both Kristen and I have decided to incorporate the use of Kritik into our upcoming courses. It is noteworthy to share that our decision to use the platform will cost students a fee ranging from \$18.00 to \$24.00 Canadian per term. The lower fee is associated with adoption of Kritik by multiple faculty members at the same learning institution. Given we both concur that peer evaluation is a beneficial pedagogical approach to learning and instruction and that our experiences with Kritik have for the most part been positive, we believe adoption of Kritik for use in our courses will be money well spent.

### **References**

1. Merriam G, Merriam C, Webster N (2005) Merriam-Webster's Collegiate Dictionary (11<sup>th</sup> ed.) Springfield, MA: Merriam-Webster.
2. Liu NE, Carless D (2006) Peer feedback: the learning element of peer assessment. *Teaching in Higher education* 11: 279-290.
3. Parboteeah S, Anwar M (2009) Thematic analysis of written assignment feedback: Implications for nurse education. *Nurse Education Today* 29: 753-757. [[crossref](#)]
4. Murette C (2022) V1\_Kritik Product Breakdown.

### **Citation:**

Ryan KD, Hardy KA (2022) Peer-to-Peer Evaluation: Kritik Pilot at Two Post-secondary Canadian Institutions. *Integr J Nurs Med* Volume 3(3): 1-5.