Student Peer Assessment Tools Working Group

2020 Final Report & Recommendations

Student Peer Assessment Tools Working Group

29-July-2020

Executive Summary

In November 2019, UBC established a Student Peer Assessment (SPA) Tools Working Group for faculty and staff from both campuses. The group's purpose was to evaluate SPA tools broadly and recommend which to support at UBC, both to facilitate review of the peer collaboration process (e.g., students review other students' contribution to a group) and peer products (e.g., students review other students' work).

The group compiled market scans, UBC tool usage statistics, previous pilot and usability data, and evaluation processes used at other institutions; developed and ran a SPA tools instructor survey; and discussed or evaluated a variety of potentially suitable SPA tools. This combined information led to developing two sets of core criteria for evaluating tools against one another.

In order to expedite a recommendation ahead of online courses in fall 2020 and support the pivot to online teaching, the working group focused on a shortlist of tools, based on those that best fit the criteria extracted from the survey results and the group's own experiences. Members explored and then rated these tools against the core criteria and collaborated on lists of pros and cons. The outcomes and further discussion resulted in a recommendation of five tools:

- **iPeer** for reviewing the peer collaboration process, though this recommendation is made with the assumption further development by UBC will be supported to enhance the tool's ease of use and bring it into alignment with the core pedagogical needs it does not currently meet
- **peerScholar** for reviewing peer products using traditional peer review that can be customized to suit most contexts
- **CLAS**, **ComPAIR**, and **PeerWise** for reviewing peer products for special situations (video annotation, comparative decision making, or multiple-choice item writing), which may also help facilitate access to instructors who are new to student peer assessment

Additionally, the working group recommended that UBC support SPA tool use through:

- **An iPeer working group** of instructors to guide the needed further development of iPeer, similar to what has been done for ComPAIR
- **Improved tool explanations, documentation, and promotion** through the Learning Technology Hub and other communication channels

- **A public means of sharing ratings of and feedback on SPA tools**, so instructors have somewhere to exchange experiences they've had in this space as it continues to evolve

Background

Objective

In November 2019, UBC established a Student Peer Assessment (SPA) Tools Working Group with faculty and staff to evaluate tools in this space and recommend which to support within UBC's Learning Technology Environment (LTE). These tools include those that facilitate review of the **peer collaboration process** (e.g., another student's contribution to a group) and **peer product** (e.g., another student's work). These will be referred to by the shorthand "process" and "product" for the rest of this document.

To support the evaluation, the working group sought to do several things:

1. Elicit and analyze requirements from the UBC community to inform decision-making.
2. Based on this analysis and their own experience, determine the priority pedagogical practices that should be supported by UBC's LTE.
3. Using these priority practices, review incumbent and other available market tools, with a special focus on three areas:
    - Incumbent tools not yet meeting UBC's needs
    - Multiple incumbent tools occupying the same or similar space
    - Functional gaps where there is pedagogical demand
4. Establish a rubric for making a comparative determination about the tools.
5. Make a recommendation to the learning technology governance groups about which
    - Incumbent tools to continue supporting or developing from the nine UBC-supported tools:
        - Canvas Peer Review
        - CLAS (Collaborative Learning Annotation System)
        - ComPAIR
        - edX Open Response Assessments
        - edX Peer Instruction
        - iPeer

        - PeerMark (Turnitin)
        - peerScholar
        - PeerWise
    - Incumbent tools to sunset
    - New tools to shortlist for comparative pilot implementation or adoption

Limitations

Completing any technical evaluations, official privacy impact assessments, security investigations, or budget-level cost analysis was considered outside the scope of this group. Recommendations were made primarily on the basis of pedagogical alignment.

Membership

UBC Vancouver	UBC Okanagan
Peter Graf (Arts) - Chair Manuel Dias (CTLT) - Co-Chair Elisa Baniassad (Science) Silvia Bartolic (Arts) Rik Blok (Science) Agnes d'Entremont (Applied Science) Letitia Englund (CTLT) Shawna Faber (Education) Tim Kato (LT Hub) Jennifer Walsh Marr (Arts / Vantage) Peter Ostafichuk (Applied Science)	Yang Cao (Engineering) Vania Chan (CTL) Zoe Soon (Health and Exercise Science)

Evaluation Process

Information Gathering

To better understand the existing landscape around SPA tools at UBC and beyond, the working group undertook a number of exercises:

- Compiled existing SPA tool materials to review[1], including market scans, UBC tool usage statistics, previous pilot and usability data, and evaluation processes used at other institutions
- Developed and ran a SPA tool instructor survey[2] (in February / March 2020)
- Discussed or evaluated a variety of potentially suitable SPA tools

Survey Results

Overall, 102 instructors completed the survey across sixteen Faculties / Schools, with highest response rates coming from Applied Science (23%), Business (16%), Science (10%), and Arts (9%) for UBC Vancouver and Creative & Critical Studies (9%) and Arts & Sciences (6%) for UBC Okanagan.

SPA tool use

Respondents reported using SPA tools for having students assess five main areas: other individuals' teamwork (72%), other individuals' assignments (59%), other groups' assignments (individually) (32%), the feedback received on an individual or group assignment (24%), and other groups' assignments as a group (18%).

UBC SPA tool awareness

When asked about their knowledge of the current list of UBC-supported SPA tools, respondents reported no awareness of most tools. A majority were aware of three tools: iPeer (68% were aware), Canvas Peer Review (61%), and PeerMark (58%). Other tools were largely unrecognized: ComPAIR was known to 43%, while 20% or fewer respondents were aware of the rest (CLAS, edX Open Response Assessments, edX Peer Instruction, peerScholar, and PeerWise).

SPA tool experience

Of those rating an experience with SPA tools, most respondents reported tools meeting some or most but not all of their needs. No tools stood out as particularly exceptional, though CATME had generally more favourable and Moodle Peer Review less favourable overall reactions. But these outcomes occurred with very small samples of users (5 and 6, respectively).

[1] https://docs.google.com/document/d/1vV0bk23lFzFCWImpEb5_vU92F73uYoIEKKvy5YNw5zs/edit
[2] https://docs.google.com/document/d/1OsZpO2ECBoWKqZqACbkIj-dyflVEh-7yMDKlrHKWo68/edit

Drawbacks experienced

In discussing unmet needs when using SPA tools, respondents noted three pedagogical pain points: it's hard to train students to review well (10%), it's unclear if there are learning benefits from the process for students (9%), and it's challenging to know and apply best practices (7%).

From a technology point of view, respondents talked about tools having poor usability (29%), missing features that would help in their specific use case (16%), having poor or no support for using groups (13%), lacking rubric flexibility (12%), missing an automatic and meaningful peer review mark calculation (12%), presenting technical issues (12%), having poor or no Canvas integration (9%), requiring a significant time commitment to set up (9%), and creating difficulties with assessing feedback students give, customizing grades, and overcomplicating the process (all 7%).

Benefits experienced

In talking about needs met when using SPA tools, respondents mentioned getting insight into group dynamics (22%) and clear benefits to student learning (19%).

From a technology point of view, respondents pointed to how tools supported students exchanging feedback (24%), simply facilitated other complex review process logistics (22%), had good usability (16%), provided an automatic and meaningful peer review mark calculation (14%), had good Canvas integration (10%), offered flexible rubrics (9%), used a simple review process (9%), supported use in large courses (7%), enabled self-assessment (7%), and let students numerically mark each other (5%).

Needs for SPA tools when reviewing peer collaboration processes

For reviewing the peer collaboration process, instructors in open-ended responses mainly discussed wanting to

- Better understand group dynamics
- Distribute group marks fairly
- Help students improve their teamwork skills

The primary SPA tool technical needs mentioned included allowing students to numerically mark peer teamwork (15%), calculating a suggested mark adjustment based on peer reviews for

instructors to see and use (12%), enabling students to exchange feedback (10%), and supporting students in self-assessing their own teamwork (6%).

When rating the importance of specific SPA tool features for reviewing the peer collaboration process, about 60% or more of respondents selected the following as must-have features:

- Easy to use for instructors and students (78%)
- Students can leave qualitative comments (70%)
- Instructors can control release of peer feedback (65%)
- Instructors can download qualitative comments written by students (53%)
- Integration with Canvas classlist (58%)
- Instructors can download quantitative marks given by students (58%)

Features which were rated must-have by less than 60% but given importance by about 60% or more of respondents included the following features:

- Integration with Canvas groups (55%)
- Students can self-assess (51%)
- Integration with Canvas grades (51%)
- Students can give quantitative marks (49%)
- Ability to create groups in tool (44%)
- Instructors can customize, calculate final mark in tool (40%)
- Option for student training/practice before real reviews (23%)
- Instructors receive suggested grade adjustment / multiplier (21%)
- Provides vetted questions instructors can use (19%)
- Includes visualization or reporting on how each group is doing (18%)
- Includes learning analytics for deeper dive into student behaviour (18%)

Finally, features which were must-have by less than 60% and given importance by less than 60% of respondents included the following:

- Can be used in large courses (21%)
- UBC controls development (9%)

Needs for SPA tools when reviewing peer products

For reviewing the peer product, instructors in open-ended responses primarily discussed wanting students to

- Learn to recognize work quality
- Share quality formative feedback with each other

The main SPA tool technical needs that respondents mentioned included enabling students to exchange feedback (30%), letting students submit written work (25%), allowing students to numerically mark peer work (8%), letting students submit presentations (8%), letting students submit group work (7%), supporting students contextually commenting on peer work (6%), and calculating a mark generated by the peer reviews for instructors to see and use (5%).

When rating the importance of specific SPA tool features for reviewing peer product, about 60% or more of respondents selected the following as must-haves:

- Easy to use for instructors and students (91%)
- Students can leave qualitative comments (76%)
- Instructors can download qualitative comments written by students (70%)
- Instructors can download quantitative marks given by students (68%)
- Integration with Canvas classlist (67%)
- Instructors can leave qualitative comments on student work (62%)
- Instructors can control release of peer feedback (61%)

Features which were rated must-have by less than 60% but given importance by about 60% or more of respondents included the following:

- Reviewers can be automatically assigned by tool (56%)
- Instructors can customize, calculate final mark in tool (54%)
- Integration with Canvas grades (52%)
- Integration with Canvas groups (52%)
- Instructors can assign quantitative marks to student work (52%)
- Reviewers can be manually assigned by instructors (48%)
- Support for group submission of assignments (46%)
- Students can give quantitative marks (40%)

- Students can self-assess (39%)
- Allows individual review of group assignments (38%)
- Allows students to directly annotate assignments they review (34%)
- Supports use in large courses (33%)
- Allows group review of group assignments (28%)
- Option for students to submit assignment revision (25%)
- Allows students to numerically mark feedback they receive (22%)
- Option for student training/practice before real reviews (22%)
- Includes learning analytics for deeper dive into student behaviour (14%)

Finally, features which were rated must-have by less than 60% and given importance by less than 60% of respondents included the following:

- Students can know whose assignment they are reviewing (17%)
- Students can know who reviewed their assignment (17%)
- UBC controls development (8%)

(Visualizations for all the survey results[3] are available in Qualtrics.)

Final Determination

The combined information led to developing two sets of core criteria (see Appendix A for the full lists) for comparing tools in each group to one another. In order to expedite a recommendation ahead of online courses in fall 2020, the working group decided to focus on a shortlist of tools, which included the following.

- Tools to Review Peer Collaboration Process (e.g., another student's contribution to a group):
    - CATME
    - iPeer
    - ITP Metrics
    - Teammates
    - WebPA
- Tools to Review Peer Product (e.g., another student's work):
    - Aropä

[3] https://ql.tc/fAwLjV

- Canvas Peer Review
- CLAS
- ComPAIR
- CrowdGrader
- edX Open Response Assessments
- edX Peer Instruction
- Moodle Peer Review
- PeerMark (Turnitin)
- peerScholar
- PeerWise

Members split up the tools and individually evaluated them against the core criteria, using a 1-3 rating system; members also rated the importance of the criteria, considering both their own contexts and priorities raised in survey responses. Because this evaluation occurred during the rapid pivot to fully online teaching and learning at UBC in March 2020, the working group also more highly prioritized tools with lower barriers to entry (e.g., higher ease-of-use).

The tool values were weighted based on the criteria rating averages, then combined to give an overall sense of each tool's versatility. Additionally, members collaborated on lists noting what each tool did well and didn't do well (or didn't do at all).

This process helped surface subcategories for the tools to review product: seven that used traditional peer review, i.e., students submit work for select peers to review (Aropä, Canvas Peer Review, CrowdGrader, edX Open Response Assessments, Moodle Peer Review, PeerMark, and peerScholar) and four that offered a specialty take with a useful variation on how students review work (CLAS, ComPAIR, edX Peer Instruction, and PeerWise). The outcome also flagged tools with overlap.

Members researched, discussed, or demonstrated tools with overlap and identified pedagogical use cases for those that stood alone. The group ultimately recommended five incumbent tools[4] for continued central support.

[4] Though it should be noted that some tools not recommended will still be supported, since they are built in to larger platforms instructors may continue using (e.g., Canvas Peer Review).

Recommendations

Technology

These recommendations were made with the assumption that UBC will seek to fund any costs on both the instructor and student sides, so that use of centrally-supported systems is free, and instructors will not be constrained to the limitations of free versions of the systems, incur personal costs, or pass on costs to students to access these tools.

It is also important to note that, unlike this static report, the SPA tools space is ever-evolving, especially during these rapidly changing times. The recommendations should therefore be seen as reflecting members' best effort at choosing the most suitable tools, given the available products, features, and integrations at the end of July 2020.

Tools to review peer collaboration process

In the space of tools that focus on review of process, the working group recommended **iPeer**.

Teammates was removed from this list because it required a Google login to access, which creates privacy issues for students. WebPA was also removed because it didn't offer anything unique and hadn't been updated in over ten years. This left iPeer, CATME, and ITP Metrics. The latter two are similar in that, unlike iPeer, they offer specific, fixed, vetted rubrics with which instructors run their evaluations.

Members felt that a prescribed approach to reviewing teams had pros and cons. On one hand, having preset rubrics increases instructor trust in the outcomes and decreases time they have to spend setting up in the tool. On the other hand, it significantly increases the time students spend to complete their reviews (as the required question lists can be quite extensive, especially for short group work assignments) and eliminates rubric flexibility.

In the end, the group agreed that supporting iPeer and adding vetted rubrics (with scores that can be tabulated automatically) as an option was a better solution than supporting two tools in this space. iPeer was seen as a highly flexible tool that could be shaped into an all-in-one to meet any use case, especially since UBC oversees development internally.

Tools to review peer product

Central support for traditional peer review

Traditional peer review means students submit work for select peers to review based on rubrics set by an instructor. For this purpose, the working group recommended **peerScholar**.

Members agreed that one tool would be sufficient in this space, so long as it could cover the majority of needs. Five tools were eliminated from this list because of difficulties with access / building Canvas integration (CrowdGrader, edX Open Response Assessments, Moodle Peer Review), a lack of flexibility (Canvas Peer Review), and/or potential issues with data privacy and handling (CrowdGrader, PeerMark).

This left two tools: Aropä and peerScholar. After demonstrations and further consideration, members selected peerScholar, primarily because of its apparent higher ease of use, one of the most critical criteria identified in the survey and by the working group in the current context.

Other reasons for preferring peerScholar included the past responsiveness of the vendor to adding new features or fixing bugs required by UBC faculty, the sleek look-and-feel for students, the existing integration with Canvas, and the substantial overall flexibility that offers multiple types of rubrics, feedback options, and student submission extensions.

Aropä seemed more confusing for students and instructors to pick up, had a less polished look and feel, and presented a rubric creation process that appeared more challenging. Additionally, most of the functionality available in Aropä was also available in peerScholar; the biggest benefit of Aropä was the lesser upfront cost associated with using an open-source tool.

Central support for specialty peer review

Specialty peer review means students submit work but there is a twist on how they will review each other's work (i.e., they will not only compare to an instructor rubric). For this purpose, the working group recommended the following:

- **CLAS**
- **ComPAIR**
- **PeerWise**

edX Peer Instruction was eliminated from this list largely because it did not fill a clear use case in the context of peer review[5]. Members felt the other tools offered clear pedagogical value that was distinct from what could be accomplished with peerScholar. They also believed offering a variety of tools that do more specialized things than peerScholar could help with uptake of the pedagogically-rich student peer assessment process generally.

CLAS enables highly interactive peer annotations of peers' multimedia; ComPAIR eases novices into peer review by inviting comparative judgement on pairs of peers' work; and PeerWise helps students study through peer review by answering and evaluating peers' study questions.

Training & Tool Discoverability

Beyond technology recommendations, the working group also discussed desired UBC support for SPA tool use.

iPeer working group formation

As part of recommending iPeer, the working group strongly advocated forming an instructor working group to guide its further development, similar to what has been run for ComPAIR. This group could work with the development team to make long-term plans and incorporate meaningful, pedagogically-effective features like premade rubrics, a student training round, and better visualization/scoring of review outcomes. It could also oversee other needed improvements like better usability for instructors and students.

Improved tool explanations, documentation, and promotion

Two notable issues that surfaced in the survey results were that many instructors: a) struggle with how to do peer review well in their courses and b) do not realize what tools are available to support this process.

[5] As the name suggests, edX Peer Instruction is intended to mirror the peer instruction process typically done during in-person classes, where students answer a question individually, discuss in small groups, then answer the question individually again.

The working group would like to see central development of new training materials and/or promotion of existing ones[6], for both instructors and students. Many members were also willing to help with these materials (perhaps basing some comparisons around the criteria spreadsheets developed for this evaluation), and survey respondents who gave contact points for follow-up could also potentially be called on to give input.

Ideally, documentation should cover both the general pedagogical best practices and specific technical orientations to the centrally-supported tools. It should also guide instructors at any level of experience to easily understand differences between the recommended tools in one central place.

Public sharing of ratings and feedback for SPA tools

Because the student peer assessment technology landscape is always changing, the working group also recommended giving instructors a way to share input (e.g., ratings, comments) on tools they've used and read what colleagues say about their use. This could also provide a way to let central support (e.g., Learning Technology Hub) know about promising new or emergent tools or highlight features or limitations for potential development requests.

[6] For example, Peter Graf and Catherine Rawn's general student peer assessment training already in the Canvas Learning Commons.

Appendices

Appendix A: Evaluation Criteria

Tools to Review Peer Collaboration Process

1. Summary of tool suitability for UBC
2. How does this tool help instructors understand group dynamics?
3. How does this tool support instructors marking group work more fairly?
4. How does this tool encourage students to improve their teamwork skills?
5. How does this tool help train students to review well?
6. Individuals assess individual teamwork
7. Easy to use
8. Quick assignment setup
9. Flexible rubrics for reviewing
10. Provides vetted questions to use
11. Instructors customize participation marks
12. Instructors manage group lists
13. Students write feedback on teamwork
14. Students write feedback to instructors specifically
15. Students numerically mark peers
16. Students log teamwork activities
17. Students self-assess own work
18. Students see teamwork marks received
19. Instructors control release of peer feedback
20. Instructors see suggested mark adjustment
21. Instructors can download feedback given
22. Instructors can download marks given
23. Instructors customize final mark
24. Integration with Canvas
25. Compliant with FIPPA
26. Scalability for large classroom use

Tools to Review Peer Product

1. How does this tool support students better recognizing work quality?
2. How does this tool support students exchanging quality formative feedback?
3. How does this tool help train students to review well?
4. Individuals assess individual peer work
5. Individuals assess group peer work
6. Groups assess group peer work
7. Individuals assess feedback they receive
8. Easy to use
9. Quick assignment setup
10. Flexible rubrics for reviewing
11. Tool automatically assigns reviews
12. Instructors manually assign reviews
13. Instructors customize participation marks
14. Support for group work
15. Students know whose work they review
16. Students submit written work
17. Students submit presentations
18. Students submit images
19. Students submit code
20. Students submit equations
21. Students submit videos
22. Students exchange feedback on work
23. Students contextually comment on work
24. Students numerically mark work
25. Students self-assess own work
26. Students write feedback on feedback received
27. Students numerically mark feedback received
28. Students resubmit work
29. Instructors control release of peer feedback
30. Instructors see peer-generated marks
31. Instructors can numerically mark work
32. Instructors can leave feedback on work
33. Instructors can numerically mark student feedback
34. Instructors can leave feedback on student feedback

35. Instructors can download feedback given
36. Instructors can download marks given
37. Instructors customize final mark
38. Integration with Canvas
39. Compliant with FIPPA
40. Scalability for large classroom use
41. Students can contribute anonymously