Evidence Based Library and Information Practice

Evaluating Bibliographic Referencing Tools for a Polytechnic Environment

Gina Brander, Erin Langman, Tasha Maddison and Jennifer Shrubsole

Article abstract

Objective – This paper analyzes the design process for a toolkit for appraising emerging and established bibliographic reference generators and managers for a particular student population. Others looking to adapt or draw from the toolkit to meet the needs of users at their own institutions will benefit from this exploration of how one team developed and streamlined the process of assessment.

Methods – The authors implemented an extensive initial evaluation using a checklist and comprehensive rubric to review and select reference tools. This work was guided by a matrix of categories from Marino (2012), Bates (2015), and other literature. As the tools were assessed using the toolkit, the components of the toolkit were evaluated and revised. Toolkit revisions were based on evaluators' feedback and lessons learned during the testing process.

Results – Fifty-three tools were screened using a checklist that reviewed features, including cost and referencing styles. Eighteen tools were thoroughly evaluated using the comprehensive rubric by multiple researchers to minimize bias. From this secondary testing, tools were recommended for use within this environment. Ultimately the process of creating an assessment toolkit allowed the researchers to develop a streamlined process for further testing. The toolkit includes a checklist to reduce the list of potential tools, a rubric for features, a rubric to evaluate qualitative criteria, and an instrument for scoring.

Conclusion – User needs and the campus environment are critical considerations for the selection of reference tools. For this project, researchers developed a comprehensive rubric and testing procedure to ensure consistency and validity of data. The streamlined process in turn enabled library staff to provide evidence based recommendations for the most suitable manager or generator to meet the needs of individual programs.

Evidence Based Library and Information Practice

Research Article

Evaluating Bibliographic Referencing Tools for a Polytechnic Environment

Gina Brander
Reference & Information Services Librarian
Saskatchewan Polytechnic Library
Regina, Saskatchewan, Canada
Email: gina.brander@saskpolytech.ca

Erin Langman
Nursing Liaison Librarian
Saskatchewan Polytechnic Library
Regina, Saskatchewan, Canada
Email: erin.langman@saskpolytech.ca

Tasha Maddison
OER & Copyright Librarian
Saskatchewan Polytechnic Library
Saskatoon, Saskatchewan, Canada
Email: maddisont@saskpolytech.ca

Jennifer Shrubsole
Learning & Teaching Librarian
Saskatchewan Polytechnic Library
Moose Jaw, Saskatchewan, Canada
Email: jennifer.shrubsole@saskpolytech.ca

Abstract

Objective – This paper analyzes the design process for a toolkit for appraising emerging and established bibliographic reference generators and managers for a particular student population. Others looking to adapt or draw from the toolkit to meet the needs of users at their own institutions will benefit from this exploration of how one team developed and streamlined the process of assessment.

Methods – The authors implemented an extensive initial evaluation using a checklist and comprehensive rubric to review and select reference tools. This work was guided by a matrix of categories from Marino (2012), Bates (2015), and other literature. As the tools were assessed using the toolkit, the components of the toolkit were evaluated and revised. Toolkit revisions were based on evaluators' feedback and lessons learned during the testing process.

Results – Fifty-three tools were screened using a checklist that reviewed features, including cost and referencing styles. Eighteen tools were thoroughly evaluated using the comprehensive rubric by multiple researchers to minimize bias. From this secondary testing, tools were recommended for use within this environment. Ultimately the process of creating an assessment toolkit allowed the researchers to develop a streamlined process for further testing. The toolkit includes a checklist to reduce the list of potential tools, a rubric for features, a rubric to evaluate qualitative criteria, and an instrument for scoring.

Conclusion – User needs and the campus environment are critical considerations for the selection of reference tools. For this project, researchers developed a comprehensive rubric and testing procedure to ensure consistency and validity of data. The streamlined process in turn enabled library staff to provide evidence based recommendations for the most suitable manager or generator to meet the needs of individual programs.

Introduction

Saskatchewan Polytechnic does not provide access to a subscription-based reference management tool. A task force of four librarians and two library technicians reviewed products for students requiring the use of a generator and/or manager. A key consideration was that training on the product and usability should not be prohibitive, as most students spend a limited amount of time in library instructional sessions. The review also looked at the availability of free basic features, accuracy, technical support, and storage options.

Reference managers are robust tools that allow users to save bibliographic metadata and organize sources into folders and subfolders.

Users can create references for individual citations, or entire bibliographies. Many reference managers also include the ability for multiple users to collaborate on a shared library. Some managers include web browser plug-ins for adding citations and references quickly and easily, as well as word processing add-ons that allow users to search their citation library and automatically add citations and references. Reference generators, on the other hand, are simple tools for creating references or bibliographies. They do not have the ability to save references beyond an individual session, and do not have the additional features of many reference managers. Table 1 outlines the main differences between reference managers and generators in greater detail.

Table 1

Features of Reference Managers vs. Generators

Features	Reference Managers	Reference Generators
Create in-text citations	✓	Usually
Create bibliographies	✓	✓
Save and organize references	✓	✗
Collaborate with other users	✓	✗
Browser integration	✓	✗
Word processing add-on	Usually	✗
Save PDFs or other files	Usually	✗
Entirely web-based	Rarely	✓

The most likely users of reference management tools among Saskatchewan Polytechnic's student population include:

- students working on a group project who wish to share resources,
- students working on a capstone project with the expectation of compiling a large number of resources, and
- students in degree programs with a high expectation of research and writing skills, such as Bachelor of Science in Nursing.

The most likely users of reference generators include:

- students in shorter programs, and
- students with assignments requiring few sources.

Due to the range of student needs, both generators and management tools are equally valid in this environment. As a result, both were analyzed as part of this research project.

Aims

This paper describes the process of developing a comprehensive rubric and selection process to identify the best reference tool(s) to meet the needs of students at Saskatchewan Polytechnic. The emphasis of this paper is on the process used to create and revise the assessment toolkit, rather than on the end results of the evaluation. Others looking to adapt or draw from the toolkit to meet the needs of students at their own institutions will benefit from this analysis of how the team developed and streamlined the process of assessment.

Context

Saskatchewan Polytechnic offers 12 advanced certificates, 37 certificates, 37 applied certificates, 52 diplomas, three degrees, two post-graduate certificates, and 24 apprenticeship programs (Saskatchewan Polytechnic, 2017). Programs range from two months to four years, and represent most sectors of the economy.

The varying expectations of programs regarding research requirements and the accuracy of citations pose specific challenges for the library. Many programs do not prepare students to engage in practice-based research or publishing, and instead focus on teaching students about the ethical use of information by correctly citing their sources. Often, assignments for these projects require a low number of sources. Since the focus for many programs is on the "why" rather than the "how" of citing sources, many students would benefit from a simple tool that generates citations and references, and which requires less than half an hour to set up and learn.

At the other end of the needs spectrum are the institution's diploma and degree program students. These programs assign research-intensive, sometimes semester-long projects, which are best managed using tools that offer options for storage and collaboration. The institution's focus on student participation in applied research also engenders the need for more robust tools.

Since the length of programs and student needs vary substantially, the task force entered into the project with the understanding that more than one tool would be selected for recommendation, and that both generators and managers would be investigated. They were also aware that the Polytechnic's librarians do not necessarily have the opportunity to instruct students in the use of reference tools, since many programs have minimal interactions with the library outside of their orientations, and various programs are offered by distance with these students receiving no library instruction. Qualitative criteria (i.e., criteria assessed based on individual experience) were thus considered important in determining whether a tool could be easily accessed, learned, and adopted with little support or previous experience.

Literature Review

It is well established that students have

difficulty creating bibliographic references (Blicblau, Bruwer, & Dini, 2016; Stevens, 2016). Both students and researchers express frustration with creating references and often find the process aggravating and tedious (Antonijević & Cahoy, 2014; Stevens, 2016). Evidence indicates that bibliographic referencing tools may alleviate these negative emotions (Stevens, 2016).

Recommending reference tools is a common task in libraries (Childress, 2011). Because many people will continue to use the same tool even if not fully satisfied with it, it is important that librarians give advice based on evidence and the users' needs rather than personal preference (Antonijević & Cahoy, 2014; Blicblau et al., 2016). When matching reference tools to users' needs, librarians need to consider both management tools that store sources long-term, and generating tools that create a copyable reference without the need for long-term storage (Childress, 2011).

Relatively few researchers have evaluated reference tools; instead, most authors discuss situations and ways to apply the tools (Childress, 2011; Lorenzetti & Ghali, 2013; Stevens, 2016). While there is no common methodology for evaluating reference tools (Tramullas et al., 2015), many analyses compare functions and features of the tools (Homol, 2014; Imperial College London Library, 2017; Universitätsbibliothek Technische Universität München, 2016). Most of the evaluations are based on the needs of university students at the graduate level, or on professional researchers within a particular field (Kratochvíl, 2016; Lorenzetti & Ghali, 2013). Even if a user group is defined, most studies do not detail user needs as an initial step (Tramullas, Sánchez-Casabón, & Garrido-Picazo, 2015).

Unlike user groups discussed in the literature, students at polytechnics have a wide range of backgrounds, prior credentials, and work experience (Berger, Motte, & Parkin, 2009; Canadian Information Centre for International

Credentials and Council of Ministers of Education Canada, 2016; Polytechnics Canada, 2015). Polytechnics offer trade and technological education, as well as adult education, health sciences, and business programs (Canadian Council on Learning, 2010; Saskatchewan Polytechnic, 2017). Within polytechnics, there is a focus on applied research within a specific industry (Canadian Council on Learning, 2010). Likewise, workplace information literacy that is contextualized to the program is important (Bird, Crumpton, Ozan, & Williams, 2012). Workplace information literacy, unlike academic information literacy, has collective approaches to information, and supports the use of "noncanonical sources" (Inskip, 2014; Lloyd, 2011).

Evaluations by researchers tend to focus on the literature within a particular field, often in the health sciences (Gilmour & Cobus-Kuo, 2011; Kratochvíl, 2016). When testing, researchers often use articles, with books and book chapters as other commonly tested formats (Gilmour & Cobus-Kuo, 2011; Homol, 2014; Kratochvíl, 2016). Testing referencing tools for the polytechnic environment should include articles and books, but also other formats important to, and commonly used within, polytechnics - grey literature and web pages (Imagine Easy Solutions & EasyBib.com, 2014; Kelly, 2015; Kratochvíl, 2016; Marsolek, Cooper, Farrell, & Kelly, 2018; Oermann et al., 2008).

Most evaluations in the literature use a checklist approach or a list of features. If qualitative comments are included, there is no definition of terms used, e.g. "good" ease of use without explaining what "good" means (Universitätsbibliothek Technische Universität München, 2016, p. 15). One framework was found in Marino's (2012) "Fore-Cite: Tactics for Evaluating Citation Management Tools." Marino asks the reader to consider the following:

- The environment (the tools available),
- The user in your unique situation,

- The purpose of the software, such as a generator or a manager,
- System and browser requirements,
- Accessibility,
- Features important to your users,
- Vendor support available, and
- True cost of ownership.

Complementing Marino is Bates (2015) SECTIONS model for selection of educational media:

- Students,
- Ease of use,
- Costs,
- Teaching functions, or if the design promotes learning,
- Interaction: of student with the technology, of student to instructor, and of student to student,
- Organisational issues, such as institutional support and barriers,
- Networking with others outside the course, and
- Security and privacy.

Both Marino and Bates take student-centered approaches, making them excellent starting points for developing a research methodology.

Methods

With no common testing methodology and a student population quite different from those studied in the literature, the team built a methodology from the ground up. Guided by Longsight (2013), the team decided to design a rubric that included outcomes and functions, as well as features. Including the outcomes and functions in the rubric provided qualitative data on the way people interacted with each reference tool. For this reason, user experience was partially incorporated into the evaluation of the reference tools. A rubric would also improve the decision-making process, since multiple people made the final choices (ASQ, n.d.).

Criteria were gathered from the Marino (2012) and Bates (2015) frameworks, and other research literature. Notable lists of criteria were Wikipedia ("Comparison," 2016), Universitätsbibliothek Technische Universität München (2016), and the University of Toronto Libraries (2016). Both librarians and faculty members teaching referencing styles provided input on essential tool features. While reviewers did not agree on all points, commonly indicated criteria by librarians and faculty became minimal requirements in the rubric. See Appendix A for the final list of criteria.

Without a pre-existing rubric, the team adapted a matrix for the evaluation of learning management systems. The Longsight (2013) matrix, which shared many similar criteria with the Bates and Marino models, used a four-point rating scale (with four being the highest rank for a criterion). The researchers wanted to present the final results as a letter grade, and a four-point scale allowed easy calculation of both a percentage and a four-point grade average for conversion into a letter grade. To indicate the minimum requirement of each criterion, a shading system was used (see Table 2). Shaded areas indicate acceptable levels; lightly shaded areas represent barely acceptable rankings; and no shading indicates unacceptable levels.

Evaluators highlighted the words in a cell to indicate their decision (see Table 2). If evaluators had additional comments on a criterion and ranking, they were invited to add these to the relevant cell (see Table 3). During decision-making, and especially if there were inconsistencies in rankings by reviewers, comments were considered in addition to the numerical values.

The rubric was not normed due to time constraints; however, at least two people tested each reference tool to increase reliability and reduce bias. Rankings assigned by testers were fairly consistent, unless browser ad and popup blockers were active. During the testing process, the rubric was revised twice based on feedback

from testers. The testing process itself was also refined and developed into an assessment toolkit:

1. A checklist of basic features that are easy to measure (Appendix C)
2. A rubric capturing more complex features (Appendix D)
3. A rubric to evaluate qualitative criteria (Appendix E)
4. An instrument to score features, an Excel file (Brander, Maddison, Langman and Shrubsole, 2019)

As the process was refined, the team moved from exclusively focusing on reference managers to also including generators.

Based on a list from Wikipedia ("Comparison," 2016) and other literature (Beel, 2014; "Bibliographic," n.d.; G2 Crowd, n.d.), an inventory of tools was compiled (see Appendix B). For local reasons, some tools were excluded, such as mobile apps and Microsoft Word.

The final version of the process started with a checklist to reduce the number of potential tools (Appendix C). This checklist included easily assessed "deal-breakers" such as cost, referencing style(s) available, local computing requirements, and type of installation (e.g., all software requiring a user to set up a server was eliminated). Since a tool either did or did not have the criteria identified, a single person was adequate to complete the checklist. If there was any uncertainty as to whether a tool met the criterion, it moved to the next stage.

In the second stage, a rubric was used to examine additional features (see Appendix D). Each evaluator noted the type of device, operating system, and browser used for testing, but this information was only used if there were discrepancies between evaluations. Except for the criterion about viewing on mobile devices, evaluators used a desktop or laptop during testing. Evaluators used an automatic tool for input, such as a browser extension, for all but

Table 2
Example of Reviewer's Highlighting, with Illustration of Shading Showing (1) Unacceptable, (2) Barely Acceptable, and (3;4) Acceptable Levels

CRITERION	1	2	3	4
Privacy settings	No policy nor statement on privacy	Privacy policy or statement exists. May involve third parties.	Privacy policy or statement exists. No third parties involved. Server located outside Canada	Privacy policy or statement exists. No third parties involved. Server located in Canada

Table 3
Evaluators' Ratings with Comments

Privacy settings Mendeley evaluator #2	2. Privacy policy or statement exists. May involve third parties *Not clear where the server is located, but headquarters for Elsevier is in the UK. Third party information is controlled by individual user. Can sign up through Facebook, as well as connect Scopus author information to social network.*
Privacy settings Zotero evaluator #1	2. Privacy policy or statement exists. May involve third parties *Server located outside of Canada, site contains links to third-party web sites, but they do not share information to third parties. An open source project, and apps are created by a third party.*

two sources. These two sources tested manual input options so as to explore all input methods. With ad and pop-up blockers disabled, evaluators' results were consistent, so in future only one evaluator could complete this stage.

During the second stage, a list of the formats supported by each tool was created and compared. The formats were then placed into categories, including books and articles (print), photographs and maps (images), video and music (multimedia), web pages and software (Internet and computers), statues (legal), speeches (verbal), and other. Online versions of

formats were placed into the same category as the hard copy; for instance, ebooks were considered print-based, and streaming videos were placed into multimedia. Within the tool, only the main categories of formats were considered, and not subdivisions between hard-copy and electronic versions.

The final stage (Appendix E) tested experiential factors, such as ease of set-up. Since a person's previous experience impacted the evaluation, more than one person was needed for this stage. As with the previous stage, browser ad and pop-up blockers were disabled. Two librarians and

two technicians tested each tool. While the librarians had expertise in bibliographic styles and with reference tools, each technician had experience in one area, but not the other. The team did not include anyone who was inexperienced with both computing and citing. Evaluators tested the full process during this stage.

Results of the accuracy evaluations were not reliable, since reference tools were not retested after revision of the accuracy evaluation methodology. The instrument used to score the features of individual tools also underwent several revisions. The version that was current at the time of the publication of this article is available (Brander, Maddison, Langman and Shrubsole, 2019).

Results

Results of testing revealed that none of the reference managers available at the time of testing were outstanding choices for Saskatchewan Polytechnic. Despite this temporary setback, the project enabled the team to identify the best tools currently available based on defined institutional and student needs. Six tools (three reference managers and three generators) out of the original fifty-three were recommended as a result of testing: RefME, Zotero, Mendeley, CiteFast, Citation Machine, and EasyBib.

Given that none of the reference managers were identified as ideal choices for Saskatchewan Polytechnic, future assessments are likely as the reference tool terrain evolves and new tools become established. The third and final toolkit version (see appendices) was determined by the team to have achieved a better balance of simplification and detail than the first two versions. Toolkit components include:

1. A checklist of basic features that are easy to measure (Appendix C)
2. A rubric capturing more complex features

3. A rubric to evaluate qualitative criteria (e.g.., one's experience of the tool)
4. An instrument to score features

The final toolkit, developed and revised throughout the process of testing, offers an efficient, adaptable, and evidence based method for future testing.

Discussion

Throughout the process of developing and revising the toolkit, the authors experienced a number of challenges. While some of the challenges were anticipated and unavoidable, others informed toolkit changes and established new evaluator expectations. The lessons learned throughout this process will improve future iterations of testing.

Be Adaptable

One of the key takeaways for the authors was the need to be flexible. Fluid criteria such as institutional software updates, library platform changes, style guide edition and reference tool version updates, and business instability can complicate or prolong the testing process. The sale of the seemingly well-established product RefME in the middle of testing illustrates the state of flux testers work within. Testing does not occur in a vacuum, and unavoidable challenges may necessitate revisions or a return to the drawing board.

Simplify Where You Can

The several toolkit versions tested ultimately led to a final set of documents that match efficiency with accuracy. Key changes included the addition of a checklist, and modification to the number of rubric testers based on the nature of the information being evaluated.

The team determined that the addition of a checklist would speed up the overall process, since some criteria in the rubric did not require a scale to evaluate whether or not tools met basic

requirements. The checklist allows a single person to quickly assess descriptive criteria and eliminate tools that do not meet the most basic requirements. Using this checklist, the team was able to quickly reduce the number of tools that required in-depth assessment from 53 to 18. The rubric was retained to assess the remaining criteria, which required finer distinctions in evaluation.

The team also decided to split the original rubric into two separate rubrics to improve reliability. Individual assessments of qualitative factors were found to vary substantially, since these are influenced by testers' previous experiences with other citation management tools. For example, the reviewer ratings were found to range widely for the "self-efficacy" criterion, which considered the level of instruction or consultation required for initial setup and use of a tool. To address this inconsistency without requiring more than one tester to evaluate quantitative criteria for 18 tools, the final version of the toolkit placed qualitative criteria in a separate, shorter rubric. Separating these criteria allowed two testers to evaluate and compare criteria that were open for interpretation, while requiring only one tester to evaluate yes/no criteria.

Limitations

This iteration of testing did not compare the capabilities of standalone browser extensions and applications for mobile devices. Investigations into the functionality of tools in mobile environments will gain more relevance as students' dependence on handheld devices for computing continues to increase.

Only a cursory evaluation of user experience and accessibility was performed. User testing would provide rich details about experience and is an additional testing method to be considered in the future. Both usability and accessibility testing would require different methodologies than the ones applied in this project.

Since reference tools often have issues with languages other than English (Libson, 2018), English-language materials alone were tested. Datasets were also not tested, despite their importance, due to numerous issues around citing this format (Kelly, 2016).

Lastly, while the review investigated available privacy policies, input from legal experts is needed for further development of these criteria.

Conclusion

Much was learned throughout this project as the rubric and testing procedures were refined. Without an assessment of local user needs, there is no basis for informed decision making. It is critical, therefore, to understand the users' needs within a particular institutional environment, and to adapt the toolkit accordingly.

Not all criteria required testing by multiple reviewers and reducing the number of people involved expedited the process. Descriptive criteria (e.g., cost or support available) yielded generally consistent results across reviewers and did not require multiple evaluators. Using a checklist of essential criteria instead of a rubric also accelerated the process. On the other hand, multiple viewpoints were vital when considering experiential criteria (e.g., level of complexity of processes).

Saskatchewan Polytechnic Library has ultimately benefited from the development and application of these assessment tools to help identify the most appropriate bibliographic referencing tools for the student population. The assessment tools have allowed the Library to provide evidence based advice that can be tailored to the needs of individual users. As a result, library staff are better equipped to aid students with reference tools within a polytechnic setting.

As reference tools continue to evolve, there is a strong possibility that the toolkit will require revision. In future iterations, an emphasis

should be also placed on user experience testing and evaluation of accessibility. For now, however, the toolkit provides a foundation for ongoing testing of reference tools at Saskatchewan Polytechnic. The toolkit also provides a starting place for others looking to draw from or adapt it to meet students' needs at their institutions.

References

Antonijević, S., & Cahoy, E. S. (2014). Personal library curation: An ethnographic study of scholars' information practices. *Portal: Libraries and the Academy*, *14*(2), 287–306. https://doi.org/10.1353/pla.2014.0010

ASQ. (n.d.). Decision matrix. Retrieved from http://asq.org/learn-about-quality/decision-making-tools/overview/decision-matrix.html

Bates, A. W. (2015). *Teaching in a digital age: Guidelines for designing teaching and learning for a digital age*. Retrieved from https://opentextbc.ca/teachinginadigitalage

Beel, J. (2014, January 15). Comprehensive comparison of reference managers: Mendeley vs. Zotero vs. Docear [Web log post]. Retrieved from http://www.docear.org/2014/01/15/comprehensive-comparison-of-reference-managers-mendeley-vs-zotero-vs-docear

Berger, J., Motte, A., & Parkin, A. (Eds.). (2009). *The price of knowledge: Access and student finance in Canada* (4th ed.). Montreal, QC: The Canada Millennium Scholarship Foundation. Retrieved from http://www.yorku.ca/pathways/literature/Access/The%20Price%20of%20Knowledge%202009.pdf

Bibliographic listings (n.d.). Retrieved from http://dirtdirectory.org/

Bird, N. J., Crumpton, M., Ozan, M., & Williams, T. (2012). Workplace information literacy: A neglected priority for community college libraries. *Journal of Business & Finance Librarianship*, *17*(1), 18–33. https://doi.org/10.1080/08963568.2012.630593

Blicblau, A. S., Bruwer, M., & Dini, K. (2016). Do engineering students perceive that different learning and teaching modes improve their referencing and citation skills? *International Journal of Mechanical Engineering Education*, *44*(1), 3–15. https://doi.org/10.1177/0306419015624186

Brander, G., Maddison, T., Langman, E., & Shrubsole, J. (2019). Scoring instrument for reference tools. UAL Dataverse, V1. https://doi.org/10.7939/DVN/7PMKTO/EX8VFM

Canadian Council on Learning. (2010). *Challenges in Canadian post-secondary education: Navigating post-secondary education in Canada: The challenge of a changing landscape*. Retrieved from ERIC database (ED525096).Canadian Information Centre for International Credentials, and Council of Ministers of Education, Canada. (2016). *Canada's education systems*. Retrieved from https://www.cicic.ca/docs/PTeducation/Canada-s-Education-Systems-PDF.pdf

Childress, D. (2011). Citation tools in academic libraries: Best practices for reference and instruction. *Reference & User Services Quarterly*, *51*(2), 143–152. https://doi.org/10.5860/rusq.51n2.143

Comparison of reference management software. (2016, December 2). In *Wikipedia*. Retrieved from https://en.wikipedia.org/wiki/Comparis

on_of_reference_management_software

G2 Crowd. (n.d.). Best reference management software. Retrieved from https://www.g2crowd.com/categories/reference-management

Gilmour, R., & Cobus-Kuo, L. (2011). Reference management software: A comparative analysis of four products. *Issues in Science and Technology Librarianship, 66.* https://doi.org/10.5062/F4Z60KZF

Homol, L. (2014). Web-based citation management tools: Comparing the accuracy of their electronic journal citations. *The Journal of Academic Librarianship*, *40*(6), 552–557. https://doi.org/10.1016/j.acalib.2014.09.011

Imagine Easy Solutions, & EasyBib.com. (2014). *Perspectives on student research skills in K-12 and academic communities*. Retrieved from http://info.easybib.com/hs-fs/hub/222136/file-1990871677-pdf/InfoLitReport-Part2.pdf?t=1414690132007

Imperial College London Library. (2017). *Reference management software comparison*. Retrieved from http://www.imperial.ac.uk/media/imperial-college/administration-and-support-services/library/public/Reference-management-software-comparison.pdf

Inskip, C. (2014, June). *Information literacy is for life, not just for a good degree: A literature review*. Retrieved from http://discovery.ucl.ac.uk/1448073/

Kelly, M. (2015). Citation patterns of engineering, statistics, and computer science researchers: An internal and external citation analysis across multiple engineering subfields. *College & Research Libraries*, *76*(7), 859–882.

https://doi.org/10.5860/crl.76.7.859

Kratochvíl, J. (2017). Comparison of the accuracy of bibliographical references generated for medical citation styles by EndNote, Mendeley, RefWorks and Zotero. *The Journal of Academic Librarianship*, *43*(1), 57–66. https://doi.org/10.1016/j.acalib.2016.09.001

Lisbon, A. H. (2018). Multilingual scholarship: Non-English sources and reference management software. *The Journal of Academic Librarianship, 44*(1), 60–65. https://doi.org/10.1016/j.acalib.2017.12.001

Lloyd, A. (2011). Trapped between a rock and a hard place: What counts as information literacy in the workplace and how is it conceptualized? *Library Trends*, *60*(2), 277–296. https://doi.org/10.1353/lib.2011.0046

Longsight. (2013). *Criteria for the evaluation of learning management systems*. Retrieved from https://www.ncat.edu/provost/docs/lms/LMS%20Evaluation%20Tool.pdf

Lorenzetti, D. L., & Ghali, W. A. (2013). Reference management software for systematic reviews and meta-analyses: An exploration of usage and usability. *BMC Medical Research Methodology*, *13*. https://doi.org/10.1186/1471-2288-13-141

Marino, W. (2012). Fore-cite: Tactics for evaluating citation management tools. *Reference Services Review*, *40*(2), 295–310. https://doi.org/10.1108/00907321211228336

Marsolek, W., Cooper, K., Farrell, S., & Kelly, J. (2018). The types, frequencies, and findability of disciplinary grey literature within prominent subject databases and

academic institutional repositories. *Journal of Librarianship and Scholarly Communication*, *6*(1). https://doi.org/10.7710/2162-3309.2200

Oermann, M. H., Nordstrom, C. K., Wilmes, N. A., Denison, D., Webb, S. A., Featherston, D. E., … Striz, P. (2008). Information sources for developing the nursing literature. *International Journal of Nursing Studies*, *45*(4), 580–587. https://doi.org/10.1016/j.ijnurstu.2006.10.005

Polytechnics Canada. (2015). What is a polytechnic. Retrieved from http://www.polytechnicscanada.ca/polytechnic-advantage/what-polytechnic

Saskatchewan Polytechnic. (2017). *Impact that matters: 2016-2017 annual report*. Retrieved from https://saskpolytech.ca/about/about-us/reports-and-statistics/annualreport/documents/2016-17%20Sask%20Polytech%20Annual%20Report.pdf

Stevens, C. R. (2016). Citation generators, OWL, and the persistence of error-ridden

references: An assessment for learning approach to citation errors. *The Journal of Academic Librarianship*, *42*(6), 712–718. https://doi.org/10.1016/j.acalib.2016.07.003

Tramullas, J., Sánchez-Casabón, A. I., & Garrido-Picazo, P. (2015). Studies and analysis of reference management software: A literature review. *El Profesional de La Información*, *24*(5), 680–688. https://doi.org/10.3145/epi.2015.sep.17

Universitätsbibliothek Technische Universität München. (2016). *Reference management software comparison - 6th update (June 2016)*. Retrieved from https://mediatum.ub.tum.de/doc/1320978/1320978.pdf

University of Toronto Libraries, Gerstein Science Information Centre. (2016). *What's the best citation management software for me?* Retrieved from http://guides.library.utoronto.ca/c.php?g=250610&p=1671260

Appendix A
List of Evaluative Criteria

Students (including access)	Ease of Use	Costs	Teaching (Referencing) Functions
● Integration with word processors ● Mobile devices (e.g. app) ● Portability ● Syncing	● Ease of set-up ● Design and layout ● Ease of adding information ● Auto-fill options ● Ease of editing info ● Signaling (e.g. date format, incomplete info) ● Intuitiveness of making references ● Intuitiveness of making citations ● Self-efficacy	● Free or freemium ● If freemium, extent of use available at no cost ● If freemium, cost of storage for free	● Accuracy* ● Citation style(s) ● Content stored (e.g. metadata only) ● Formats (e.g. journal articles) ● Generic form for any format ● Importing and exporting files (e.g. RIS) ● PDF compatibility ● Storage available
Interaction (including with learning materials)	Organisational (Institutional) Issues	Networking (externally, e.g. social media)	Security and Privacy
● Collaborating and sharing ● Manual entry available ● Organization and discovery ● Presence of ads	● Authentication ● Browser(s) ● Installation requirements ● Operating system(s) ● Support available ● Training materials available	[nothing in this category]	● Business stability ● Data security ● Privacy settings

* Accuracy includes correct information, correct presentation (e.g. punctuation, capitalization), correct format of source, no empty fields, no fields missing, and variance in accuracy based on method of addition (e.g. browser tool vs. database search within tool).

Based on Bates (2015)

Appendix B
Tools Examined, by Category

Reference Generators

- APA Style Wizard*
- BibMe*ǂ
- Citation Machine *ǂ
- Citation Producer
- CiteFast*ǂ
- CiteMaker*ǂ
- Citethisforme*ǂ
- ClassTools
- EasyBib*ǂ
- Google Scholar*ǂ
- KnightCite*ǂ
- Make Citation
- NCSU Citation Builder
- Noodle Tools Express*ǂ
- OttoBib

Reference Managers

- Aigaion
- Bebop
- BibBase
- BibDesk
- Biblioscape
- BibServer
- BibSonomy
- Bibus
- Bookends
- Citavi
- Citelighter
- CiteULike*
- Colwiz*
- Docear
- EndNote
- EndNote Basic*
- EWWOWW (WizFolio)*
- F1000Workspace

- JabRef
- KBibTeX
- Mendeley*
- Noodle Tools
- Paperpile
- Papers
- Pybliographer
- Qiqqa
- ReadCube
- refbase
- RefDB
- RefME*
- Reference Manager
- Referencer
- RefWorks
- SciRef
- Sente
- Wikindx

Zotero*

* Tool examined using rubrics.
ǂ Tested as generator; manager option available.

Appendix C
Step 1: Checklist of Initial Criteria

Highlight the features that each tool has. You may highlight more than one option in a category. Repeat rows as necessary.

Name of Tool	Cost (Exclude Trial Period)	Type of Tool (free version)	Operating Systems	Citation Styles	Installation
	Free Freemium (free version, pay to upgrade) Fee to use	Generator (cannot save) Manager (save references)	Windows Mac	APA (6th ed.) CSE Chicago MLA (8th ed.) IEEE	Desktop (application) Desktop (server set up) Web-based (online account) Browser tool

Appendix D
Step 2: Expanded Features of Reference Tools

Tool Tested: _____

Name of Reviewer: _____

Thank you for helping us determine which generator(s) are best for our students. Your input will be very important to making a decision.

Instructions
- Use the list of sources attached in the email:
  - When manually adding sources, please ignore special information for some items. This is for automatic entry, which may pull this information even if not used.
  - Please add information by an automatic method (e.g. bookmarklet), except where indicated.
    - If more than one way of adding information automatically, please try all methods.
  - After making notes about the automatic method, please 'fix' information before making the reference.

- Please highlight text that best describes your experience under each criterion.
  - If the criterion does not apply to your tool, please skip that line.
  - If you need to highlight two squares, please explain why both squares apply to the tool.

- There are areas to note observations.
  - Please feel free to add extra observations, as you feel are needed.

- If you test two different citation styles, please fill out two forms (the tool may act differently for each style).

- Please copy and paste the bibliography into a Word document.
  - Upload bibliography to [] with your initials, tool name, and citation format in filename.
  - We will be noting how accurate the tool is.

- Please upload a copy of this form, with your initials and name of tool in the filename, to […].

Please check options, and highlight text or make notes as needed.

Style Tested

	APA		MLA: 8th edition		IEEE (note version, if applicable)

How I Am Testing

Device			
Type	**Make & Model (e.g. Samsung S7)**	**Operating System**	**Version**
Desktop/laptop		Chromebook Linux Mac Windows	
Tablet		Android Blackberry iOS Windows	
Phone		Android Blackberry iOS Windows	

Browsers Supported

Browser	Full or Partial Support?	Version Tested
Chrome		
Edge		
Firefox		
Internet Explorer		
Safari		
Other:		

General Overview

Legend: Clear = unacceptable; light grey = barely acceptable; dark grey = acceptable

CRITERION	1	2	3	4
Setting up (Generator)	Before using, you must create an account and/or install software	No account to use, but some functions require software installation 5 steps.	No account to use and perform basic functions. Additional functions require software installation.	No account and no software installation required to use tool.
Setting up (Manager)	To use, you must install software and a browser tool. Syncs with online personal account.	To use, you must install software. A browser tool is optional. Syncs with online personal account.	To use, you must install a browser tool. Syncs with online personal account.	To use, you set up online personal account. No software installation necessary (may have optional browser tool).
Storage (Manager)	Must pay to store references (e.g. free generates only); or small free storage (<100 references, <100MB total) without option to buy additional storage.	Free storage limited by # of references (<100), or free storage limited by file size (<100MB). Option to buy additional storage.	Free storage limited by # of references (>100), or free storage limited by file size (>100 MB). Option to buy additional storage.	No storage limit in free version.
Collaborating and sharing (Manager)	Cannot co-create bibliography and cannot share references.	Can co-create bibliography, but can only share with co-creator.	Can co-create bibliography, and can share references. Can only share references publicly (except with co-creator).	Can co-create and share bibliography, and can control privacy of bibliography (e.g. can share publicly or privately).

CRITERION	1	2	3	4
Content stored (Manager)	Stores only metadata (info about sources).	Stores metadata, and can add files or notes, but not both.	Stores metadata, and can add files and notes. Cannot annotate PDF files.	Stores metadata, and can add files and notes. Can annotate PDF files.
Organization and discovery (Manager)	One method is available: searching, filtering, folders, or tagging.	You can search or filter sources, but folders or tagging requires payment.	You can search or filter sources, and use folders or tagging. Search does not include full-text (e.g. PDFs).	You can search or filter sources, and use folders or tagging. Search includes full-text (e.g. PDFs).
Ads (Disable ad blockers)	Contains ads. Ads are obtrusive (video that plays automatically, pop-up, animated gif).	Contains ads. Ads are static and/or videos that do not play automatically. Ads are a majority of the interface.	Contains ads. Ads are static and/or videos that do not play automatically, and are a minority of the interface.	No ads.
Viewing on mobile devices	Does not scale to mobile screen. Function impaired by lack of mobile version.	Does not scale to mobile screen. Functionality not impaired by lack of mobile version.	Scales to mobile screen. Basic functionality available, but not all features, in mobile version.	Scales to mobile screen. Full functionality available in mobile version.
App	No app available.	App over $5 or not available for Android and iOS.	App developed by third party (i.e. not by developer of original tool. Cost is under $5 and available for Android and iOS.	App developed by same entity as the original tool. Cost is under $5 and available for Android and iOS.

Support and Business Operations

CRITERION	1	2	3	4
Support	No support available, or support only available for a cost.	Email support only.	Support available both live and by email. Live support has limited	Support available both live and by email. Live support available 24/7.
Training materials	No training materials.	Print training materials available, no videos/tutorials available.	Print training materials available, with some videos/tutorials available.	Print training materials available, with extensive collection of videos/tutorials available.
Online help resources	No online help resources.	A users' manual or space (e.g. blog, user group) is accessible online.	Help files are accessible at each step of a process.	Contextually-appropriate help files provide assistance as appropriate. Pop-ups or rollovers provide "just-in-time" information for specific actions
Privacy settings	No policy nor statement on privacy.	Privacy policy or statement exists. May involve third parties.	Privacy policy or statement exists. No third parties involved. Server located outside	Privacy policy or statement exists. No third parties involved. Server located in Canada.
Data security (Manager)	Unknown information about backing up. No policy on data retrieval in event of suspension of	Backs up data. No policy on data retrieval in event of suspension of business operations.	Backs up data. Allows data retrieval if business suspends operations. No notification if operations suspended.	Backs up data. Allows data retrieval if business suspends operations. Will send notification if operations suspended.
Business stability	Business is a high-risk organization (e.g. past financial trouble, recent start-up without 'buzz').	Business is a recent start-up that has received attention.	Business has some maturity, but it is unclear if there is continuing investment in product development.	Business has some maturity, and has evidence of continuing investment in product development.

Adding Information

CRITERION	1	2	3	4
Method of adding info	Only automatic entry available (i.e. no manual entry).	Only manual entry available.	Manual entry available, plus one automatic method (database, bookmarklet or other browser tool).	Manual entry available. Includes multiple automatic methods (database, bookmarklet or other browser tool).
Missing/ required info	No warning about missing info. Required info not indicated.	Warns about missing info. Required info not indicated.	No warning about missing info. Required info indicated.	Warns about missing info. Required info indicated.
Autofill from searching in tool	No autofill from searching tool's internal database.	Autofill from searching internal database. Search/database has complicated interface.	Autofill from searching internal database. Search has simple interface, but usually does not find sources.	Autofill from searching internal database. Search has simple interface. Can usually find information on sources.
Browser tool (e.g. bookmarklet)	No browser tool.	Browser tool is not intuitive to use.	Browser tool is intuitive, but makes frequent errors.	Browser tool is intuitive, and retrieves metadata fairly accurately.
Importing sources	Cannot import sources.	Can add import files, but not a common file type.	Can add import files, but limited to one common file type (e.g. RIS).	Can add import multiple common files types (e.g. RIS, BibTeX).

CRITERION	1	2	3	4
Format of author name	No clue for order of author's name.	Clues for order of author's name are not obvious.	Order of author's name is obvious. Dealing with multiple authors not obvious.	Order of author's name is obvious. It is obvious how to include multiple authors.
Corporate authors	No option for corporate author.	Corporate author option is hard to find.	Corporate author option is obvious, but difficult to use.	Corporate author option is obvious, and easy to use.
Date format	No clue for date format. Format matters.	Clues for date format are not obvious.	Date format is obvious.	Date format does not matter (i.e. tool changes automatically).
Generic form for any format	No option available.	Includes generic format, but has very few fields available.	Includes generic format, but does not consider usability of presentation of options.	Includes generic format, and considers usability of presentation of options (e.g. 'chunks' options).
PDFs	No support available – manual entry.	Limited importing of metadata. Importing requires software or extra steps.	Limited importing of metadata. Importing does not require software or extra steps.	Full importing of metadata.

Making References and Citations

CRITERION	1	2	3	4
Bibliographic references	No references output.	Can only make individual references.	Can make full bibliographic list.	Can make full bibliographic list. Bibliography included within full document (i.e. write paper in tool).
Citations	No citation output.	Provides basic citation information (i.e. no page number).	Provides full citation information. Page number may be prompt only.	Full citation information provided in-context within full document.
Word processors	No special word processor support (e.g. paste text version).	Can export in MS Word format, or have MS Word plug-in. Support only for Windows.	Can export in MS Word format, or have MS Word plug-in. Support for Windows and Mac.	Can export in various formats, or have plug-ins for various word processors, including MS Word.
Order of references	Not applicable (e.g. does not create list of references).	Does not follow style's requirements for reference order.	Follows style order, but makes errors (e.g. alphabetize The).	Follows style order, and is accurate in order.
Indentation (APA, MLA)	No hanging indentation.	Hanging indentation in tool, but formatting disappears in MS Word.	Hanging indentation kept in MS Word, but other formatting needed.	Hanging indentation kept in MS Word, and no extra formatting required.
Capitalization	Keeps capitalization style of input.	Changes capitalization to citation style, but makes 1 error (e.g. does not capitalize subtitle).	Changes capitalization to citation style, but makes multiple errors.	Changes capitalization to citation style. There are no capitalization errors.

Exporting Sources

Options for exporting sources (e.g. RIS, MS Word, RTF).		Export individual sources only? Yes / No

Types of sources supported (e.g. reports, journal articles):

Any additional comments that may affect decision:

Estimated time to complete evaluation:

Appendix E
Step 3: Experiential Criteria of Reference Tools

Tool Tested: _____

Name of Reviewer: _____

Thank you for helping us determine which generator(s) are best for our students. Your input will be very
important to making a decision.

Instructions
- Use the list of sources attached in the email:
  - When manually adding sources, please ignore special information for some items. This is
    for automatic entry, which may pull this information even if not used.
  - Please add information by an automatic method (e.g. bookmarklet), except where
    indicated.
    - If more than one way of adding information automatically, please try all
      methods.
  - After making notes about the automatic method, please 'fix' information before making
    the reference.

- Please highlight text that best describes your experience under each criterion.
  - If the criterion does not apply to your tool, please skip that line.
  - If you need to highlight two squares, make sure to explain why.

- There are areas to note observations.
  - Please feel free to add extra observations, as you feel are needed.

- If you test two different citation styles, please fill out two forms (the tool may act differently for
  each style).

- Please copy and paste the bibliography into a Word document.
  - Upload Word bibliography to [] with your initials, tool name, and citation format in
    filename.
  - We will be noting how accurate the tool is.

- Please upload a copy of this form, with your initials and name of tool in the filename, to [].

Please check options, and highlight text or make notes as needed.

Style Tested

	APA		MLA: 8th edition		IEEE (note version, if applicable)

How I Am Testing

Device			
Type	**Make & Model (e.g. Samsung S7)**	**Operating System**	**Version**
Desktop/laptop		Chromebook Linux Mac Windows	
Tablet		Android Blackberry iOS Windows	
Phone		Android Blackberry iOS Windows	

Browser	
Type	**Version**
Chrome	
Edge	
Firefox	
Internet Explorer	
Safari	
Other:	**Name** / **Version**

Using the Tool

Legend: Clear = unacceptable; light grey = barely acceptable; dark grey = acceptable

CRITERION	1	2	3	4
Setting up/ Starting	Process is complicated and requires over 5 steps.	Process may be confusing to new users, but has under 5 steps.	Process is clear, but has over 5 steps.	Process is clear and has under 5 steps.
Design and layout	Complex and potentially confusing.	Interface is functional, but some features may be complex and/or confusing.	Functional interface that can be navigated with minimal training. Look is fairly simple and straightforward.	Intuitive interface that can be navigated with little or no training. Look is simple and straightforward.
Adding information (automatically)	Process is complicated, and contains extra screens/windows.	Process may be confusing to new users, but does not contain extra screens/windows.	Process is clear, but contains extra screens/windows.	Process is clear, and does not contain extra screens/windows.
Adding information (manually)	Difficult to find option for format. Cannot add all required information.	Easy to find option for format. Can add all required information.	Difficult to find option for format. Cannot add all required information.	Process is clear, and does not contain extra screens/windows.
Editing sources	Difficult to change information. Many changes needed.	Difficult to change information. Few changes needed.	Easy to change information. Many changes needed.	Easy to change information. Few changes needed.

CRITERION	1	2	3	4
Making references	Requires extra clicks during process. Process could be simplified.	Process simple, but may be confusing to new users.	Process simple and clear, but output may not be clear to new users.	Process simple and clear. Output is clearly identified.
Self-efficacy	For most users, requires instruction or consultation for set-up. Significant time required to maintain skills.	For most users, requires instruction or consultation for set-up. Requires regular use to maintain skills.	Most users should be able to set up with help materials alone. Requires regular use to maintain skills.	Most users should be able to set up with help materials alone. Minimal time required to maintain skills.

What did you like best about this tool?

What did you find most frustrating about this tool?

Any additional comments that may affect decision

Estimated time to complete evaluation: