TEACHING WEB EVALUATION

A cognitive development approach

Web evaluation has been a standard information literacy offering for years and has always been a challenging topic for instruction librarians. Over time, the authors had tried a myriad of strategies to teach freshmen how to assess the credibility of Web sites but felt the efforts were insufficient. By familiarizing themselves with the cognitive development research, they were able to effectively revamp Web evaluation instruction to improve student learning. This article discusses the problems of traditional methods, such as checklists; summarizes the cognitive development research, particularly in regards to its relationship to the ACRL Information Literacy Standards; and details the instructional lesson plan developed by the authors that incorporates cognitive development theories.
INTRODUCTION

As the amount of information available to students has exploded exponentially, it has become increasingly critical for instruction librarians to teach students not only how to find sources but also how to evaluate them. This demand falls in line with the ACRL Information Literacy Competency Standards for Higher Education; Standard 3, Performance Indicator 2 reads: “The information literate student articulates and applies initial criteria for evaluating both the information and its sources” (Association of College and Research Libraries, 2000, p. 11).

At Radford University, Web evaluation has been a standard offering for course-integrated library sessions for many years. It has been an especially popular topic for first-year composition classes. Over time, the librarians have tried numerous strategies, but they never felt their efforts were adequate; after library instruction sessions, professors reported that students were still showing minimal ability to analyze online sources. However, by familiarizing themselves with the cognitive development research, the librarians at Radford University were able effectively to revamp Web evaluation instruction, and as a result, improve student learning relative to that subject matter.

The literature shows that some form of Web source evaluation instruction is necessary for undergraduate students. There are many reasons for this, but the primary ones are that Web sources are so popular, and that the quality of those sources varies widely. Students tend to favor the use of Web materials over others because they are easier to find (Biddix, Chung, & Park, 2011). Research also demonstrates that users tend to rate visual presentation of materials more highly than any other, somewhat more reliable criteria (Fogg, 2003). In a study by Project Information Literacy, student respondents rated authority and currency as the top criteria for choosing which sources to incorporate into academic research (Head & Eisenberg, 2010). However, research that looks at actual student behavior shows that the reality of choosing sources might not follow this ideal (Hogan & Varnhagen, 2012; Flanagin & Metzger, 2007). Oftentimes, students will trust the first results that a search engine provides and those with brand recognition (Hargittai, Fullerton, Menchen-Trevino, & Thomas, 2010). Taken together, these studies lead to concerns that students may choose sites that lack credibility (Metzger, 2007).

HISTORY OF WEB EVALUATION INSTRUCTION

To address these gaps in student learning, instruction librarians have tried many approaches to teach Web evaluation skills. Checklists were a popular technique in the late 1990s-early 2000s as Internet sites became acceptable resources. Librarians took criteria used to evaluate print sources and adapted them for Websites, creating checklists used to evaluate Web sources. While these checklists of criteria have many different acronyms and mnemonic devices attached (such as CRAAP), most address Authority, Accuracy, Currency, Bias and Relevancy (Metzger, 2007).

Most of these exercises begin with a lecture on the criteria before instructors provide pre-selected good and bad Websites and direct students to use the checklist to assess these Websites (Kapoun, 1998). The sample Websites are very clearly “good” or “bad,” and sometimes include hoax and extremist sites (Doyle & Hammon, 2006; Mathson &
Lorenzen, 2008). Some early exercises involved different sets of checklists, and students would have to match the correct checklist with the appropriate type of Website before evaluating (Tate & Alexander, 1996).

**HISTORY OF WEB EVALUATION AT RADFORD UNIVERSITY**

Since 2001, the authors’ strategies for teaching Web evaluation have mirrored techniques discussed in the library literature. Responding to student feedback that evaluating sources was too amorphous, the librarians created a checklist with a built-in rating system for each category. For example, when looking at authorship, students would check whether the site’s author was A) an expert in the field (2 points) B) Journalist (1 point) C) Author has personal experience (1 point) D) Author is named but cannot tell much about him or her (0 points) E) No author (-1 point) F) Author is a student (-2 points). Students could run a Website through the checklist and add up the category points. Sites that scored within the highest bracket would be deemed “Excellent,” while those in the lowest bracket would be deemed “Inappropriate.”

**PROBLEMS ENCOUNTERED**

At Radford University, two major problems were encountered with the checklist method. First, though critical thinking was encouraged, students utilizing the checklist tended to slide down the slippery slope of dualistic thinking. The worksheet’s rating system was employed frequently in a simple “right” or “wrong” approach, and students placed more weight on those categories where the system could be employed most easily. This was particularly apparent with the “where” category, where a quick glance at the URL could determine a rating. Hence, a .com site weighed in as “bad” even in cases where the site was highly reputable and written by an expert in the field.

The second problem was intertwined with the first: the difficulty of analyzing the websites to determine their credibility in some categories. The inherent nebulous nature of websites did not allow the criteria to be applied as neatly as in published sources with more rigid guidelines in place. The material needed to be contemplated or carefully analyzed, and in some cases, outside sources needed to be consulted in order to determine credibility.

The “who” category proved to be particularly problematic for students. Based on the difficulties of locating author information, assessing what was there, and the occasional necessity of looking elsewhere for information about the authors, overwhelmed students would turn to simpler categories to help make a determination. The end result was that students were not learning how truly to evaluate Websites, but simply how to determine the quickest and simplest way to run a site through a list of criteria.

Problems with this model are also reflected in the professional literature. In the mid-2000s, numerous studies revealed a disconnect between the checklist models and how students actually evaluate Websites. Sometimes, the questions asked in criteria may not be a good match in certain circumstances. They can be too simplistic or overly complicated for a specific site. They may be unrealistic, involve too many steps to evaluate information, and impractical; students won’t incorporate the checklist in their own evaluation process due to confusion or

Meola (2004) recommends a more practical context method in which students are encouraged critically to compare several Websites on the same topic and evaluate the context the source appears in (edited, reviewed, via a fee-based database, etc.). Comparing free sources alongside each other allows students to analyze content and verify accuracy (Dahl, 2009). One study assessed Web evaluation skills through a one-minute paper assessment tool, confirmed that the checklist method didn’t work, and discussed plans to move to the context method (Choinski & Emanuel, 2006).

SEEKING A SOLUTION

Inspired by this context research, the authors transitioned to a system that asked students to dig deep into a source and describe what they found rather than simply checking boxes. In this incarnation, students looked at either a specific website or compared two sites on a similar topic. Worksheets included questions related to the standard criteria and directed students to complete steps such as: use a reference book to find out more about the author or sponsoring organization, analyze questionable content, and consider the absence or inclusion of references. Students would then offer opinions on the websites based on what they discovered.

This instructional method change seemed to advance students’ evaluative skills. “Light bulb moments” could be witnessed as students began to see the value of considering different facts when analyzing websites. However, students still wanted to apply what had been relevant and specific to the particular websites utilized during the class session to websites they found themselves, even in cases where the criteria were not applicable. If the librarian underscored the importance of looking for the Website’s references in class, a student may select a website outside of class without consideration for author expertise or relevancy of content, focusing instead on the presence of a reference list. This decision could also be biased by student familiarity with a website, such as About.com, where the fact that “everyone uses it” trumped locating information about the author(s). Even after sessions in which the students seemed to excel at the evaluation worksheet, the professors reported that their class would backslide into using simplistic criteria when choosing sources on their own. “It appeared on the first page of Google results” was commonly cited by freshmen as a good reason to use a Website.

Frustrated by this lack of knowledge retention, librarians decided completely to overhaul how Web evaluation was taught and conducted a literature review outside of the library literature for ideas. During this process, they were struck by research in the realm of cognitive development.

COGNITIVE DEVELOPMENT

In the 1960s, William Perry and his colleagues at the Bureau of Study Counsel at Harvard University conducted a qualitative longitudinal study of male Harvard undergraduates and female Radcliffe undergraduates in order to document their experiences across four years of college (Perry, 1970). The students in Perry’s study met with Bureau staff at different points in their college careers for open-ended talks during which they reflected about their past academic year. Based on this study, Perry (1970) described
nine positions that students move through during their college career. Positions 1 and 2, grouped as dualism, describe many students beginning their college careers with the belief that there are definite right and wrong answers. To a dualistic student, success depends upon listening to authority figures to receive the “right” answers (Perry, 1970).

Perry found that by the time students in his study completed their freshman year, they had reached one of the multiplicity positions (Positions 3 and 4). These students accepted that there is not always a “right” answer to every question and that every person has an opinion that is as good as anyone else’s (Perry, 1970). For a student to move into a relativistic position (Positions 5 and 6), they must become aware that there are very few “right” answers, but that most knowledge is contextual (Perry, 1970). Most students in Perry’s study did not move into relativistic positions until the end of their college careers, if they attained this level at all. Perry found that very few college students are able to move into the positions of commitment (Positions 7, 8, and 9) because they are not ready to come to great conclusions about values and occupations to create a “way of life” before graduation (Perry, 1970). Due to the nature of the university and the time period, Perry’s findings may not translate perfectly to the current higher education population; however, a cautious comparison may be made to modern undergraduates.

In both this landmark study and later research, incoming undergraduate students saw the world in terms of right/wrong, black/white, good/bad and progressed gradually to a stage where they could appreciate differing points of view by the time they were graduated (Perry, 1970; King & Kitchener, 1994).

**Cognitive Development and Information Literacy**

What does cognitive development research mean for information literacy instruction? According to Rebecca Jackson (2007), “information literacy standards may include many competencies that are beyond the cognition level of the students librarians encounter” (p. 30). Librarians may become frustrated at students who expect answers to be provided to them, but dualistic students believe that there is one “right” answer to most problems and that authority figures possess those answers. Constance Mellon (1982) explains that dualistic students “have little patience with alternative search strategies… and with the complexities of information retrieval” (p. 80). After all, if there is only one “right” answer, why should the student consult multiple sources to find it?

Students at early stages of cognitive development may have a particularly hard time evaluating their information sources using skills identified in Standard 3 of the Information Literacy Competency Standards for Higher Education (ACRL, 2000). Jackson (2007) notes that the performance indicators and outcomes listed under Standard 3 “call for skills that are far beyond what the average freshman student can accomplish” (p. 30). As a result, students may look for an easy way out or a resource that will evaluate sources for them. According to Michael Lorenzen (2001), “the nature of the Web and the difficulty it presents in verifying information, means that students in the early stages of Perry’s Scheme are going to have difficulty in using the Web appropriately” (p. 153). Many of the students Lorenzen (2001) interviewed “felt that if a Website was indexed by Yahoo! the information found on the Web site was reliable” (p. 161). Therefore, they
didn’t feel that they needed to verify information found on the Web or evaluate Web sources at all. The dualistic viewpoint of most college freshmen can cause problems for librarians attempting to teach Web evaluation classes since students are not ready to master the skills necessary to critically assess Web sources.

While mastery may be out of reach, freshmen can—and must—begin to learn the basics of evaluation. Most colleges and universities require students to conduct research from the first year. At Radford University, research papers are required in two general education courses (Core 102 and Core 201) that are taken by freshmen and sophomores. Professors encourage students to use articles and books from the library; the books and articles have been through some review process and therefore tend to be more critical, but in truth, the lure of Google is too great. Students will use items from the open Web and need at least some rudimentary training in evaluation to select credible sources.

NEW APPROACH

Based on the cognitive development literature, the authors knew that first-year students would likely still be in the dualism stage. The librarians decided to use a constructivist approach to Web evaluation. In a constructivist environment, students learn by doing. They pull from their own personal experience in order to give context to the information they encounter (Booth, 2011; Cooperstein & Kocevar-Weidinger, 2004). The constructivist Web evaluation exercise emphasizes self-learning and is adaptable for either 50 or 75-minute library sessions. Students are divided into groups of two or three and are given a worksheet that divides the exercise into three activities. In the first activity, students develop their own criteria to evaluate Websites. In the second activity, students decide what would pass as a gold standard Website. The third and final activity is structured like a competition. The students are given a topic and must find a Website that fits the gold standard criteria that they developed in the previous activity. (Appendix A contains a copy of the student worksheet.)

In the first activity, students are introduced to a Website that is not credible. Working in groups of two or three, students are instructed to determine five reasons why the given Website is not credible. Students are given five to seven minutes to complete this activity, and are then asked to share their findings with the class. The librarian listens to the class discussion and writes group responses on the white board. This conversation about the Website’s shortcomings organically leads to the development of general criteria for evaluation. For example, students often supply responses that fit nicely into the 5Ws, or the who, what, when, where, and why categories. They typically discover who is the author of the Website and recognize that he is not someone that can be considered an expert in the subject. Next, they typically point out that the text of the site is poorly written and full of typographical errors, so the what category is lacking. The creation date, or the when, of the Website is deemed outdated. Often times, students mention the where or the domain name of the Website. The librarian would then take that opportunity to discuss domain names and how they are not always the best benchmark to use when deciding whether or not a source is credible. Lastly, students notice that the language used throughout the Website is very biased. This looks at why the Website was created in the first place. As the discussion unfolds, the librarian groups her whiteboard notes with the who,
what, when, where, and why labels. It is then explained that these criteria, which the students developed themselves, can be employed in any source evaluation. Rather than framing the criteria as a checklist, these general categories are viewed as context-sensitive. Students can now use their own experience with a “bad” Website to predict what features may be ideal for “good” Websites on any particular topic.

By identifying the sample site as a “bad” Website from the beginning, the librarian creates a safe environment for the class. Students know the Website is not credible, so they can concentrate on finding supporting evidence rather than worrying they might give the wrong answer about credibility or suitability of the source. It also affirms the students’ initial dualistic feelings: there are good sites and there are bad sites. By not challenging students’ assumptions at the beginning, the librarian can concentrate on the importance of contextualizing criteria rather than teaching oversimplified guidelines.

The second activity of the class allows students to set what their “gold standard” Website would look like. Students are given a specific topic to research and are asked to specify the features of a gold standard Website for it. Using the who, what, when, where, and why categories, students (continuing to work in their assigned groups) set benchmarks for each criterion. For the who, students decide what kind of profession would be the most credible. For the what, students consider what specific topic they want the site to discuss. For the when, students think about how current the site should be to provide the most accurate information. For the where, students consider what kind of domain they would like to host the information. Lastly, for the why, students decide what the intent of the site should be. Once each group has decided on their standards, the class must come to consensus about their “gold standard” for each category through open discussion.

This exercise builds on the previous one after students have achieved success and feel comfortable talking about evaluating sources. In an effort to push them out of the dualistic mindset, the librarian-led discussion focuses on more multiplistic and relativistic views. A Website might be perfect for one use, but dreadful for another. For example, a student would not want to use the infamous Martin Luther King, Jr. site hosted by a White Power group (http://www.martinlutherking.org) for a biography on the civil rights leader, but she might cite it as an example in a paper on how hate groups distort history. As students work with their teams to create their own criteria, the librarian circulates and encourages students to provide reasons for their suggestions.

Once the “gold standard” has been set, it is time to move on to the last activity. Each group is given five to seven minutes to use Google to find a Website that best approaches the “gold standard” they have established. They are directed to record the Website’s name, URL, and their reasons for choosing this source on their worksheet. After the allotted time has passed, a competition begins. Each group shares the Website they chose and why they feel it is a “gold standard” source. Points are awarded based on how closely each site meets the “gold standard” that was established in each category. The group with the most points wins a small prize.

The final competition provides the opportunity for students to apply what they have learned and discussed in the first two activities. Application is often difficult to fit
into a one-shot library instruction session, but such exercises give the librarian much better insight as to whether the students actually learned the material. According to Fink’s taxonomy (2003), the application level promotes higher-order thinking by adding critical thinking to foundational knowledge. This is where the rubber meets the road; the students may have succeeded in developing context-sensitive criteria for Web evaluation, but are they able to follow through and use these skills to find a credible site?

An additional element to this exercise is the competition factor. Much has been written about the gamification of library instruction (Danforth 2011; Kim 2012) and the role of competition in learning (Attle & Baker, 2007). At Radford University, the authors witnessed these theories in action. Once a prize (like candy or library pens) was offered, students became much more engaged. As each group presented their “gold standard” site, the librarian asked other groups to comment for judging purposes. Since they had a vested interest in being judged ‘best,’ students were much more likely to offer sound critiques of other groups’ chosen Websites. This interaction also gave the students who were not presenting an active role in the process, reducing ‘fade out’ when not in the spotlight.

ASSESSMENT AND FEEDBACK

The instruction team employed an observational assessment, comprised of both qualitative and quantitative components, to evaluate this exercise. The assessment required the completion of a standardized form by the instruction librarian, an immediate reflection on the session’s qualitative success, a quantitative analysis of the student worksheets to see if objectives were met, and a post-review qualitative reflection (See Appendix B).

The quantitative indicators used in the assessment analyzed whether students completed their worksheets and located and recorded relevant high-quality websites. Success was indicated with 75% of attendees achieving the benchmark, partial success was 50-75% achievement, and little success was less than 50% of attendees meeting the benchmark.

During the pilot of this assessment method in Spring 2012, three classes that focused on Web evaluation (out of a total seven classes on the subject during the semester) were assessed. The librarians collected 24 worksheets (which represented 47 students, as some worked together in groups). Two of the three classes assessed had ‘Level Three Success’ on both indicators, showing a grasp of the nature of web evaluation; the other class had success with students completing the worksheet, but only ‘Level 2: Partial Success’ on the criteria of students recording relevant websites.

In the Spring 2013 semester, the librarians evaluated nine of the 14 total Web evaluation classes taught. A total of 152 student worksheets were collected, representing 152 students assessed; eight out of nine classes had ‘Level Three Success’ on both indicators; one class had success on students completing the worksheet, with ‘Level 2: Partial Success’ on students recording relevant websites.

After comparing the observational assessment data to earlier anecdotal evidence collected by the librarians on the checklist method, librarians are confident that the constructivist method effectively addressed their earlier concerns and helped students meet the goals of the lesson. While
the data cannot directly be compared to the anecdotal evidence, it does provide some basis for validating the change in approach. On a more informal note, faculty feedback has been very positive. Teaching faculty who had previously expressed frustration with their students’ inability to evaluate sources after a library session reported a great improvement following the new workshop structure. The freshmen and sophomores selected more appropriate Websites for their research papers and provided solid reasons that mimicked the contextual criteria discussed in the library sessions.

Professors in attendance have also responded positively to the simplification of the criteria to the 5 Ws. These are terms the students have previously learned, so there is no jargon to fight through. One professor shared, “I’m so glad you don’t use the word ‘authority’- what does that even mean? I always think of the police coming to get me.”

CONCLUSION

In a perfect world, information literacy would be scaffolded throughout the curriculum and students would not be expected to achieve higher-order skills, such as Web evaluation, until they are juniors or seniors at the relativist stage of their cognitive development. At most universities with traditional-aged students, however, freshmen and sophomores are assumed to be beyond dualistic thinking and ready to dive into evaluation. Such assumptions can lead to frustration among teaching faculty, librarians, and the students themselves. By exploring the literature on cognitive development and applying the lessons learned with a constructivist framework, the librarians were able to improve greatly the student learning outcomes from Web evaluation exercises. The authors discovered that by starting students in an activity that accepts their natural dualistic thinking and then easing them towards more multiplistic and relativist viewpoints, the students’ abilities to critique websites and choose appropriate ones for their projects greatly improved.

REFERENCES


Dahl, C. (2009). Undergraduate research in the public domain: The evaluation of non-


**APPENDIX A—WEB EVALUATION WORKSHEET**

**Exercise 1:** Go to [http://joedawson.org/Interests/SmokersRights/Essays/issues1.html#smoke](http://joedawson.org/Interests/SmokersRights/Essays/issues1.html#smoke)

This is a bad website. With your teammates, list at least five reasons why your professor would not want you to use this website for your paper.

1.
2.
3.
4.
5.

**Exercise 2: Creating a Gold Standard**

Your cousin has heard that you should not drink bottled water that’s been sitting in a hot car because the plastic bottles leak a toxic substance that increases the drinker’s chance of developing cancer. As breast cancer runs in your family, this is an issue dear to your heart. What characteristics would you want to see (who, what, when, where, why) in a website you would be willing to use to advise your cousin about whether it’s safe to drink the water.

**Exercise 3: Find a website that most closely meets the gold standard criteria developed by the class.**

Name of website: 
URL: 
Reasons for choosing:

**APPENDIX B — OBSERVATIONAL ASSESSMENT OF LIBRARY INSTRUCTION**

Reflection: How did the session go? (Should be completed before looking at students’ worksheets.)

Assessment Rubric

<table>
<thead>
<tr>
<th>Number of students in class:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of worksheets collected:</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Level 3: Success</th>
<th>Level 2: Partial Success</th>
<th>Level 1: Little Success</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students completed the worksheet</td>
<td>More than 75% of attendees</td>
<td>50%-75% of attendees</td>
<td>Less than 50% of attendees</td>
</tr>
<tr>
<td>Students recorded relevant websites on the worksheet</td>
<td>More than 75% of worksheets</td>
<td>50%-75% of worksheets</td>
<td>Less than 50% of worksheets</td>
</tr>
</tbody>
</table>

Post reflection: Having reviewed the worksheets, comment on how successful you think the session was and what, if any, things you would change for next time.