What's Format Got to Do with It?

The Role Information Formats Play in Evaluating Search Results

Wednesday 17 March 2021
3:00-4:30 PM EDT
Presenters

Brittany Brannon
Research Support Specialist
OCLC Research
brannonb@oclc.org

Tara Cataldo
Biological Sciences Librarian
University of Florida
ttobin@ufl.edu
@TaraTCataldo

Robin Kear
Liaison Librarian, English
University of Pittsburgh
rlk25@pitt.edu
@rkear

Ixchel M. Faniel
Senior Research Scientist
OCLC Research
fanieli@oclc.org
@imfaniel
Researching students’ information choices: Determining the identity and judging credibility in digital spaces (RSIC)

https://guides.uflib.ufl.edu/RSIC
“Google Generation are **format agnostic** and have little interest in the containers (reports, book chapters, encyclopedia entries) that provide context and wrapping for information ‘nuggets’.”

(Williams & Rowlands, 2007)
Research questions

• How do STEM students, ranging from 4th grade to graduate school, determine the credibility of online information resources?
  • What cues from a web search results screen do students use to judge the credibility of online information resources?
  • How do students’ characteristics influence their credibility judgments of online information resources?

• Do STEM students from 4th grade to graduate school differentiate among different types of online information resources during point-of-selection?
  • What cues from web search results screens do the students use to identify online information resources?
  • How do students’ characteristics and experiences influence their identification behaviour?

(Cataldo et al., 2020)
Research methods

Data Collection
• Pre-screen survey
• Simulation
  • Pre-interview
  • Information choice tasks
  • Post interview

Participants
• 4th – 5th grade
• 6th – 7th grade
• 9th – 12th grade
• Community college
• University undergraduate
• University graduate

(Cataldo et al., 2020)
The simulation

http://ufdc.ufl.edu/IR00010570/00001
Four of the simulation tasks

HELPFUL?

CITABLE?

EVALUATE - Credibility

IDENTIFY - Container
Containers

“recognizable ways of collecting and presenting texts of particular genres”

(Brannon et al., 2021)
Genres are...

“typified rhetorical actions based in recurrent situations”

(Miller, 1984, p. 159)
<table>
<thead>
<tr>
<th></th>
<th><strong>Genres</strong></th>
<th><strong>Containers</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Unit</strong></td>
<td>Document or text</td>
<td>Publication</td>
</tr>
<tr>
<td>Rhetorical action</td>
<td>Singular</td>
<td>Compound</td>
</tr>
<tr>
<td>Organizing principle</td>
<td>Communicative act</td>
<td>Curation</td>
</tr>
<tr>
<td>Status</td>
<td>Complete</td>
<td>Complete?</td>
</tr>
</tbody>
</table>
RSIC formats

Containers
1. Blog
2. Book
3. Conference Proceeding
4. Journal
5. Magazine
6. News
7. Preprint
8. Website

Genres
- Article
- Chapter
- Post...

Modes
- Text
- Images
- Audio...

File types
- pdf
- mp4...
Finding #1

Do students care about the container of the information resources they use?
Importance of knowing the container

- Elementary School (N=29): 67% Yes, 33% No
- Middle School (N=30): 70% Yes, 30% No
- High School (N=26): 85% Yes, 15% No
- Community College (N=30): 97% Yes, 3% No
- Undergraduate (N=30): 97% Yes, 3% No
- Graduate (N=30): 100% Yes, 0% No

(Valenza et al., 2019)
Not *format agnostic*, but they are *format confused*

RSIC study: Percentage of containers identified correctly by student count (N=175)

(Cyr et al., 2021)
Observations of Container Collapse

Students creating the wrong citations because they identified the container wrong

Students losing points on an assignment where references had to be journal articles but they misidentified other resources as journal articles
Can exposure to nature alleviate pain, distress during chemotherapy?

By Andrea Laimbeer  Updated 4:57 pm CDT, Tuesday, October 29, 2019

2. Harrison's principles of internal medicine.
Summary: "Harrison's is intended for learners throughout their careers. For students,

UF: Off-campus access limited to current UF students & employees
UF ONLINE See Link to Connect

3. Harrison's principles of internal medicine.
Published: New York: McGraw-Hill, 1950-
Publishing History: 1st ed. (1950)-

Jacksonville UFHSCL Borland
UF HEALTH SCIENCE CENTER LIBR General Collection
Observations of Container Collapse

• Creating the wrong citations because they misidentified the container (Similar to STEM)
• Misidentifying Book Review container as journal article
  • Not understanding the place of book review in scholarship
• Misidentifying Book Chapter as journal article
1. 'It' and 'Olé' in 1930: The Structural Economy of Faulkner's Complex Words


Subjects: American literature; 1900-1999; Faulkner, William (1897-1962); As I Lay Dying (1930); 'Red Leaves'; novel; Southern American novelists; short story; language; Marxist approach; Voloshinov, Valentin Nikolaevich (1895-1936)
Importance of knowing the container

Valenza et al., 2019
How have you observed container collapse in your work?

- I have never encountered this. 33%
- I encounter this when I work with our users. 33%
- I encounter this when I work with our systems. 33%
- I encounter this in both places. 33%
Finding #2

What makes students more and less successful at identifying containers?
Predicted change in percentage of containers correctly identified

- *Cohort*
- *Confidence*
- *Parent w/ Degree*
- *Genre*
- *Source*
- *Aboutness*
- Visual appearance
- URL
- Google Result Snippet
- *Total Simulation Duration*
- Different Containers
- *Resource Clicks*

**Demographics**

**Cues**

**Behaviors**

*Significant at 95% or above*

N=175

(foundings from Cyr et al., 2021)
Predicted change in percentage of containers correctly identified

- **Cohort**: 2.5
- **Confidence**: 2.0
- **Parent w/ Degree**: 1.5
- **Genre**: 5.0
- **Source**: 4.0
- **Aboutness**: 3.0
- **Visual appearance**: 2.0
- **URL**: 1.0
- **Google Result Snippet**: 0.5
- **Total Simulation Duration**: 0.0
- **Different Containers**: 0.0
- **Resource Clicks**: -1.5

*Significant at 95% or above

N=175

(findings from Cyr et al., 2021)
Genres are...

“typified rhetorical actions”
(Miller, 1984, p. 159)
Predicted change in percentage of containers correctly identified

- *Cohort* (Demographics)
- *Confidence* (Demographics)
- *Parent w/ Degree* (Demographics)
- *Genre* (Cues)
- *Source* (Cues)
- *Visual appearance* (Behaviors)
- *Aboutness* (Behaviors)
- *URL* (Behaviors)
- *Google Result Snippet* (Behaviors)
- *Total Simulation Duration* (Behaviors)
- *Different Containers* (Behaviors)
- *Resource Clicks* (Behaviors)

*Significant at 95% or above

N=175

(findings from Cyr et al., 2021)
<table>
<thead>
<tr>
<th>Reference Type</th>
<th>EndNote Online</th>
<th>Mendeley</th>
<th>Zotero</th>
<th>Sotahoe</th>
<th>M3 Word Reference: EasyBib</th>
</tr>
</thead>
<tbody>
<tr>
<td>Article</td>
<td>Journal Article</td>
<td>Book</td>
<td>Other</td>
<td>Article in periodical</td>
<td>Article in periodical</td>
</tr>
<tr>
<td>Audio</td>
<td>Audio Recording</td>
<td>Audio</td>
<td>Other</td>
<td>Book</td>
<td>Book</td>
</tr>
<tr>
<td>Bill Resolution</td>
<td>Bill Resolution</td>
<td>Audio</td>
<td>Other</td>
<td>Book</td>
<td>Book</td>
</tr>
<tr>
<td>Book</td>
<td>Book</td>
<td>Book</td>
<td>Other</td>
<td>Conference paper</td>
<td>Conference paper</td>
</tr>
<tr>
<td>Book Section</td>
<td>Book Section</td>
<td>Conference Paper</td>
<td>Other</td>
<td>Book section</td>
<td>Book section</td>
</tr>
<tr>
<td>Book, Edited Collection</td>
<td>Book, Edited Collection</td>
<td>Conference Proceed</td>
<td>Other</td>
<td>Book section</td>
<td>Book section</td>
</tr>
<tr>
<td>Computer Program</td>
<td>Computer Program</td>
<td>Conference Paper</td>
<td>Other</td>
<td>Book section</td>
<td>Book section</td>
</tr>
<tr>
<td>Court Case</td>
<td>Case</td>
<td>Conference Paper</td>
<td>Other</td>
<td>Conference paper</td>
<td>Conference paper</td>
</tr>
<tr>
<td>Journal Article</td>
<td>Journal Article</td>
<td>Conference Paper</td>
<td>Other</td>
<td>Journal article</td>
<td>Journal article</td>
</tr>
<tr>
<td>Law Statute</td>
<td>Article</td>
<td>Conference Paper</td>
<td>Other</td>
<td>Journal article</td>
<td>Journal article</td>
</tr>
<tr>
<td>Magazine Article</td>
<td>Magazine Article</td>
<td>Conference Paper</td>
<td>Other</td>
<td>Journal article</td>
<td>Journal article</td>
</tr>
<tr>
<td>Map</td>
<td>Map</td>
<td>Map</td>
<td>Other</td>
<td>Map</td>
<td>Map</td>
</tr>
<tr>
<td>Monograph</td>
<td>Monograph</td>
<td>Monograph</td>
<td>Other</td>
<td>Monograph</td>
<td>Monograph</td>
</tr>
<tr>
<td>Newspaper Article</td>
<td>Newspaper Article</td>
<td>Monograph</td>
<td>Other</td>
<td>Monograph</td>
<td>Monograph</td>
</tr>
<tr>
<td>Personal Comm</td>
<td>Personal Comm</td>
<td>Patent</td>
<td>Other</td>
<td>Patent</td>
<td>Patent</td>
</tr>
<tr>
<td>Video</td>
<td>Video</td>
<td>Video</td>
<td>Other</td>
<td>Video</td>
<td>Video</td>
</tr>
<tr>
<td>Web Page</td>
<td>Web Page</td>
<td>Web Page</td>
<td>Other</td>
<td>Web Page</td>
<td>Web Page</td>
</tr>
<tr>
<td>Webinar</td>
<td>Webinar</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Article</td>
<td>Conference Article</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Paper</td>
<td>Conference Paper</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
<tr>
<td>Conference Proceed</td>
<td>Conference Proceed</td>
<td>Webinar</td>
<td>Other</td>
<td>Webinar</td>
<td>Webinar</td>
</tr>
</tbody>
</table>

Citation management reference types
Predicted change in percentage of containers correctly identified

- Significant at 95% or above

*Predicted change in percentage of containers correctly identified

N=175

(findings from Cyr et al., 2021)
Aboutness and IL thresholds

• Aboutness is a negative predictor of correctly identifying a container
  • Discerning aboutness is the one of the hardest information literacy (IL) thresholds to pass.
• Related to these ACRL Framework dispositions:
  • ‘resist the tendency to equate format with the underlying creation process’
  • ‘realize that information sources vary greatly in content and format and have varying relevance and value…’
Journal article fixation

- Fixation on finding only the journal article container
  - Commonly see students asking to find ‘only journal articles’
  - Container is sometimes substituted for credibility analysis
Which of these findings is most surprising to you?

Predicted change in percentage of containers correctly identified

- **Cohort**
- **Confidence**
- **Parent w/ Degree**
- **Genre**
- **Source**
- **Aboutness**
- **Visual appearance**
- **Google Result Snippet**
- **Total Simulation Duration**
- **Different Containers**
- **Resource Clicks**

*Significant at 95% or above

N=175

(findings from Cyr et al., 2021)
Finding #3

What container label did students use for an information resource in JStor?
Container labels applied to an article in the magazine *Science News* in JStor

<table>
<thead>
<tr>
<th>Container labels</th>
<th>Blog</th>
<th>Book</th>
<th>Conference Proceeding</th>
<th>Journal</th>
<th>Magazine</th>
<th>News</th>
<th>Preprint</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School (n=26)</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>Community College (n=30)</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>10</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Undergraduate (n=30)</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>16</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>Graduate (n=30)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>3</td>
<td>10</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Total (n=116)</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>47</td>
<td>15</td>
<td>31</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

(Valenza et al., 2019)
Container labels applied to an article in the magazine *Science News* in JStor

<table>
<thead>
<tr>
<th></th>
<th>Blog</th>
<th>Book</th>
<th>Conference Proceeding</th>
<th>Journal</th>
<th>Magazine</th>
<th>News</th>
<th>Preprint</th>
<th>Website</th>
</tr>
</thead>
<tbody>
<tr>
<td>High School</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>7</td>
<td>4</td>
<td>9</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td>(n=26)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community College</td>
<td>1</td>
<td>4</td>
<td>0</td>
<td>10</td>
<td>5</td>
<td>8</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>(n=30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undergraduate</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>16</td>
<td>3</td>
<td>4</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>(n=30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Graduate</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>14</td>
<td>3</td>
<td>10</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>(n=30)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total (n=116)</td>
<td>2</td>
<td>8</td>
<td>0</td>
<td>47</td>
<td>15</td>
<td>31</td>
<td>1</td>
<td>12</td>
</tr>
</tbody>
</table>

(Valenza et al., 2019)
JSTOR in the humanities

• JSTOR is commonly used in high schools; undergraduates are familiar with this resource
• Undergraduates consider it gold-standard credible in the humanities, no matter what the container is
• Undergraduates seem to assume content in JSTOR is in a journal container, even graduate students
Branding and source recognition

• JSTOR = journal
• Springer = journal
• Wiley = textbook
• Gale = K12
Which cues do you find most misleading when trying to identify containers?

appearance, url, titles, publishers, article, know
Finding #4

How does container affect helpfulness, citability, and credibility judgments?
High school to graduate school helpfulness and citability judgments

Students who selected Helpful/Citable

(findings from Buhler et al., 2019)
High school to graduate school credibility judgments

(data from Buhler et al., 2019)

(findings from Buhler et al., 2019)
High school to graduate school opinions of the same information in different containers and genres

(findings from Buhler et al., 2019)
High School to Graduate School Helpfulness Judgments

![Bar graph showing students who selected Helpful/Citable by type of content: Articles and Press Release.](findings from Buhler et al., 2019)

N=116
High school to graduate school helpfulness and citability judgments

Students who selected Helpful/Citable

N=116

(findings from Buhler et al., 2019)
High school to graduate school credibility judgments

(findings from Buhler et al., 2019)
High school to graduate school opinions of the same information in different containers and genres

(findings from Buhler et al., 2019)

N=116

@RSICstudy  #containercollapse
Container = credibility?

• Faculty desire for credibility in sources
  • Expressed as ‘help students find credible sources’
  • Sometimes is focused on container as credibility
  • Incorporated into instruction in library resources
  • Quickly breaks down on the ‘open’ web
Public dissemination of research

- Confusion around popular news synopsized from research
  - Complicated container confusion
  - Which to use and cite?

Created by Ragal Kartidev
From the Noun Project
Other than scholarly journals, what is the most common container you send your users to for credible information?

- Magazine: 0%
- News: 0%
- Website: 20%
- Blog: 60%
- Book: 20%
- Other: 0%
References


