User-defined valued metrics for electronic journals

Katherine Chew
Jim Stemper
Caroline Lilyard
Mary Schoenborn

Library Assessment Conference: Building Effective, Sustainable, Practical Assessment
Charlottesville, Virginia
October 29–31, 2012
Outline

• Background
• Data
• Research Questions
• Methodology

• Findings
• Case Studies
• Conclusions
• Next Steps
Inspirations

Deborah Helman, UW-Madison: Cancellation Criteria for Eng. Faculty

<table>
<thead>
<tr>
<th>Criteria</th>
<th>Ranking Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Journals that faculty cite most in the article they publish</td>
<td>2.3</td>
</tr>
<tr>
<td>Journals in which faculty publish articles</td>
<td>3.4</td>
</tr>
<tr>
<td>Journals with the highest campus usage statistics</td>
<td>3.8</td>
</tr>
<tr>
<td>Journals with the highest ISI impact factor ratings</td>
<td>4.2</td>
</tr>
<tr>
<td>Journals published by professional associations faculty belong to</td>
<td>4.4</td>
</tr>
<tr>
<td>Journals that have good business practices (ex: author copyright retention, non-restrictive access)</td>
<td>5.7</td>
</tr>
<tr>
<td>Journals for which faculty serve on editorial boards</td>
<td>6.3</td>
</tr>
<tr>
<td>Journals that publish articles by authors that cite COE faculty</td>
<td>6.3</td>
</tr>
<tr>
<td>Journals with the lowest cost-per-use statistics</td>
<td>6.7</td>
</tr>
</tbody>
</table>

Anderson, Wilson, Li, CA Digital Library: Journal Value Metrics Assessment

<table>
<thead>
<tr>
<th>Measurement Categories</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utility</td>
<td>Usage</td>
</tr>
<tr>
<td></td>
<td>Citations</td>
</tr>
<tr>
<td>Quality</td>
<td>Impact Factor</td>
</tr>
<tr>
<td></td>
<td>SNIP</td>
</tr>
<tr>
<td>Cost Effectiveness</td>
<td>Cost Per Use</td>
</tr>
<tr>
<td></td>
<td>Cost Per SNIP</td>
</tr>
</tbody>
</table>
Problems

1. Cost Per Use based on questionable assumptions:

   A user who clicks on a link a) downloads an article b) reads it & c) cites it.

2. Each metric has unique limitations:

   SFX loses the user at the vendor site.

   COUNTER is not available from every vendor.

   Web of Science doesn’t index all journals; impact factor factor is not localized or uniform across the disciplines.
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIENCE</td>
<td>AMER ASSN ADVANCEMENTS SCIENCE</td>
<td>Print + Online</td>
<td>0038-8075</td>
<td>4700-S; 7130-S; 1011-5; 2003-S</td>
<td>$14,050.00</td>
<td>23528</td>
<td>25131</td>
<td>24929.5</td>
<td>381</td>
<td>9</td>
<td>170</td>
<td>$0.02</td>
<td>$0.58</td>
<td>$0.00</td>
<td>$1,628.83</td>
</tr>
<tr>
<td>NATURE</td>
<td>NATURE PUBLISHING GROUP</td>
<td>Print + Online</td>
<td>0028-0836</td>
<td>2072-S; 4700-S; 1012-S; 1205-S</td>
<td>$64,212.00</td>
<td>18218</td>
<td>21435</td>
<td>19826.5</td>
<td>19</td>
<td>9</td>
<td>14</td>
<td>$3.53</td>
<td>$3.00</td>
<td>$3.24</td>
<td>$7,126.99</td>
</tr>
<tr>
<td>PEDIATRICS</td>
<td>AMER ACADEMY OF PEDIATRICS</td>
<td>Print + Online</td>
<td>0031-4005</td>
<td>1001-S</td>
<td>$568.00</td>
<td>9807</td>
<td>8697</td>
<td>9207</td>
<td>12288</td>
<td></td>
<td>12288</td>
<td>$0.06</td>
<td>$0.07</td>
<td>$0.06</td>
<td>$0.05</td>
</tr>
<tr>
<td>NEW ENGLAND JOURNAL OF MEDICINE - US ED</td>
<td>MASSACHUSETTS MEDICAL SOCIETY</td>
<td>Print + Online</td>
<td>0028-4793</td>
<td>1013-S</td>
<td>$6,900.00</td>
<td>9109</td>
<td>16169</td>
<td>12639</td>
<td>53775</td>
<td></td>
<td>53775</td>
<td>$0.75</td>
<td>$0.43</td>
<td>$0.55</td>
<td>$0.13</td>
</tr>
<tr>
<td>JAMA; JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION</td>
<td>AMERICAN MEDICAL ASSOCIATION</td>
<td>Print + Online</td>
<td>0038-7484</td>
<td>1013-S</td>
<td>$957.00</td>
<td>8961</td>
<td>12333</td>
<td>10847</td>
<td>31895</td>
<td></td>
<td>31895</td>
<td>$0.11</td>
<td>$0.08</td>
<td>$0.09</td>
<td>$0.03</td>
</tr>
<tr>
<td>PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA</td>
<td>NATIONAL ACADEMY OF SCIENCES</td>
<td>Print + Online</td>
<td>0027-8424</td>
<td>1012-S</td>
<td>$4,745.00</td>
<td>8714</td>
<td>10122</td>
<td>9418</td>
<td>45070</td>
<td></td>
<td>45070</td>
<td>$0.54</td>
<td>$0.47</td>
<td>$0.50</td>
<td>$0.11</td>
</tr>
<tr>
<td>HARVARD BUSINESS REVIEW</td>
<td>HARVARD BUSINESS REVIEW PROJECT MORE</td>
<td>Print</td>
<td>0017-8012</td>
<td>2031-S; 3318-S</td>
<td>$99.00</td>
<td>6705</td>
<td>8387</td>
<td>7546</td>
<td>0</td>
<td></td>
<td>0</td>
<td>$0.01</td>
<td>$0.01</td>
<td>$0.01</td>
<td>$0.01</td>
</tr>
<tr>
<td>HEALTH AFFAIRS</td>
<td>HEALTH AFFAIRS</td>
<td>Online</td>
<td>1088-S</td>
<td>1008-S</td>
<td>$458.00</td>
<td>5832</td>
<td>5472</td>
<td>5552</td>
<td>7159</td>
<td></td>
<td>7159</td>
<td>$0.08</td>
<td>$0.08</td>
<td>$0.08</td>
<td>$0.06</td>
</tr>
<tr>
<td>ANNALS OF INTERNAL MEDICINE</td>
<td>ACP ATTN CUSTOMER SERVICE</td>
<td>Print + Online</td>
<td>0003-4819</td>
<td>1001-S</td>
<td>$520.00</td>
<td>5041</td>
<td>4957</td>
<td>4999</td>
<td>8508</td>
<td></td>
<td>8508</td>
<td>$0.10</td>
<td>$0.10</td>
<td>$0.10</td>
<td>$0.06</td>
</tr>
<tr>
<td>CIRCULATION</td>
<td>LIPPINCOTT WILLIAMS &amp; WILKINS</td>
<td>Print + Online</td>
<td>0005-7322</td>
<td>1001-S</td>
<td>$915.00</td>
<td>4822</td>
<td>5245</td>
<td>5033.5</td>
<td>13055</td>
<td></td>
<td>13055</td>
<td>$0.19</td>
<td>$0.17</td>
<td>$0.18</td>
<td>$0.67</td>
</tr>
</tbody>
</table>
Affinity Strings

tc.grad.gs.chem_engr.phd

- Campus
- School
- Degree
- Degree level
- Area of Study
The Data

Utility or reading value
- Link Resolver requests *(SFX)*
- COUNTER downloads *(publishers)*
- Affinity Strings *(EZProxy)*

Quality or citing value
- Impact Factor *(ISI)*
- Eigenfactor *(Bergstrom Lab)*
- Citations by faculty *(LJUR from ISI)*

Cost effectiveness value
Journal Subscription Costs *(EBSCO)*
Research Questions

Is SFX a “good enough” substitute for vendor data?
Do departments vary greatly in their journal use?

Is Eigenfactor a substitute for impact factor?
Do measures predict which journals our faculty cite?

Can reading and citing values /cost create a meaningful cost-per-activity metric?
Methodology: Correlation Analysis

Pearson’s correlation coefficient, or R and R²

<table>
<thead>
<tr>
<th>Correlation</th>
<th>Negative</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>None</td>
<td>-0.09 to 0.00</td>
<td>0.0 to 0.09</td>
</tr>
<tr>
<td>Low</td>
<td>-0.3 to -0.1</td>
<td>0.1 to 0.3</td>
</tr>
<tr>
<td>Moderate</td>
<td>-0.5 to -0.3</td>
<td>0.3 to 0.5</td>
</tr>
<tr>
<td>Strong</td>
<td>-1.0 to -0.5</td>
<td>0.5 to 1.0</td>
</tr>
</tbody>
</table>

Correlation coefficient ranges
## Correlating Rankings of Journal Hit Lists

### SFX Top 10
1. SCIENCE
2. NATURE
3. NEW ENGLAND JOURNAL OF MEDICINE
4. JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION
5. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES
6. PEDIATRICS
7. HARVARD BUSINESS REVIEW
8. LANCET
9. HEALTH AFFAIRS
10. JOURNAL OF BIOLOGICAL CHEMISTRY

### COUNTER Top 10
1. ECONOMIST
2. NEW ENGLAND JOURNAL OF MEDICINE
3. AMERICAN JOURNAL OF PUBLIC HEALTH
4. PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES
5. JOURNAL OF BIOLOGICAL CHEMISTRY
6. TIME
7. NEWSWEEK
8. LIBRARY JOURNAL
9. HARVARD BUSINESS REVIEW
10. JOURNAL OF THE AMERICAN MEDICAL ASSOCIATION

---

[University of Minnesota Libraries]
```r
> plot(I(SFXAVE+1) ~ I(SFX09+1), data, log="xy")
> summary(m3 <- lm(log(SFXAVE+1) ~ log(SFX09+1), data))

Call:
lm(formula = log(SFXAVE + 1) ~ log(SFX09 + 1), data = data)

Residuals:
     Min      1Q  Median      3Q     Max
-0.7743 -0.1140 -0.0108  0.0855  6.4833

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    0.318580   0.014308  22.62  <2e-16 ***
log(SFX09 + 1) 0.936654   0.002851  328.29  <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2583 on 4780 degrees of freedom
Multiple R-squared:  0.9576, Adjusted R-squared:  0.9575
F-statistic: 1.079e+05 on 1 and 4780 DF,  p-value: < 2.2e-16

> plot(I(SISImpactAveMod+1) ~ I(EigenAveMod+.01), data, log="xy")
> summary(m7 <- lm(log(SISImpactAveMod+1) ~ log(EigenAveMod+.01), data))

Call:
lm(formula = log(SISImpactAveMod + 1) ~ log(EigenAveMod + 0.01),
    data = data)

Residuals:
     Min      1Q  Median      3Q     Max
-1.65892 -0.23634 -0.02975  0.18139  1.97139

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)     3.38867    0.03814   88.84  <2e-16 ***
log(EigenAveMod + 0.01) 0.58816    0.00952   61.78  <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.3729 on 3514 degrees of freedom
(1266 observations deleted due to missingness)
Multiple R-squared:  0.5207, Adjusted R-squared:  0.5205
F-statistic: 3817 on 1 and 3514 DF,  p-value: < 2.2e-16

> plot(I(CounterAveMod+1) ~ I(SFXAVE+1), data, log="xy", xlab="Average SFX", ylab="Average Counter")
> summary(m5 <- lm(log(CounterAveMod+1) ~ log(SFXAVE+1), data))

Call:
lm(formula = log(CounterAveMod + 1) ~ log(SFXAVE + 1), data = data)

Residuals:
     Min      1Q  Median      3Q     Max
-7.1822 -1.0238  0.1876  1.3188  3.1096
```
Results: Reading Value (SFX vs COUNTER)

n=4782; r=0.72, r²=0.52
Results: Citing Value (Impact Factor vs Eigenfactor)

- Impact Factor vs Eigenfactor:
  - n=3579; r=0.65, r²=0.42
  - n=3579; r=0.81, r²=0.65
  - n=3579; r=0.64, r² =0.8
Results: Citing Value (Cited By And SFX/COUNTER)

n=4782; r=0.66, r²=0.43

n=4782; r=0.54, r²=0.29
Affinity Strings

n=4782; r=0.90, r²=0.81

n=4782; r=0.72, r²=0.52
Case Study: Humphrey School of Public Affairs

\[ R^2 = 0.6426 \]

**Humphrey**

SFX & Affinity Strings

\[ n=29; \ r=0.80, \ r^2=0.64 \]

\[ R^2 = 0.3606 \]

**Humphrey**

COUNTER & Affinity Strings

\[ n=29; \ r=0.60, \ r^2=0.36 \]
Case Study: Humphrey School of Public Affairs

n=29; r=0.60, r²=0.36

n=29; r=0.71, r²=0.50
Case Study: Humphrey School of Public Affairs

Humphrey Eigenfactor Score & Cited By

\[ R^2 = 0.4796 \]

Humphrey Impact Factor & Cited By

\[ R^2 = 0.1691 \]

\[ n=29; \, r=0.69, \, r^2=0.48 \]

\[ n=29; \, r=0.41, \, r^2=0.17 \]
Case Study: Nursing

\[ R^2 = 0.8096 \]

**Nursing**

SFX & Affinity Strings

\[ n=97; r=0.90, r^2=0.81 \]

\[ R^2 = 0.4754 \]

**Nursing**

COUNTER & Affinity Strings

\[ n=97; r=0.71, r^2=0.48 \]
Case Study: Nursing

n=97; r=0.51, r²=0.25

n=97; r=0.62, r²=0.40
Case Study: Nursing

n=97; r=0.80, r²=0.63

n=97; r=0.60, r²=0.36
Value Formula

\[(\text{Cost} \div \text{SFX}) + (\text{Cost} \div \text{Cites}) \div 2\]
Conclusions

• SFX click-throughs, combined with Affinity String data, provide a “good enough” picture of usage
• SFX click-throughs are more predictive of citation behavior than COUNTER downloads
• Eigenfactor Scores are more predictive of citation behavior than Impact Factors
• There is marked variation among disciplines
## Next Steps

<table>
<thead>
<tr>
<th>2009 Downloads</th>
<th>Title</th>
<th>2010 downloads</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>962</td>
<td>Nursing Research</td>
<td>1258</td>
<td>Nursing Research</td>
</tr>
<tr>
<td>826</td>
<td>J of Advanced Nursing</td>
<td>1231</td>
<td>J of Advanced Nursing</td>
</tr>
<tr>
<td>837</td>
<td>BMJ</td>
<td>859</td>
<td>American J of Nursing</td>
</tr>
<tr>
<td>602</td>
<td>American of Nursing</td>
<td>844</td>
<td>BMJ</td>
</tr>
<tr>
<td>575</td>
<td>American J of Public Health</td>
<td>539</td>
<td>Pediatrics</td>
</tr>
<tr>
<td>447</td>
<td>J Nursing Administration</td>
<td>535</td>
<td>J Nursing Administration</td>
</tr>
<tr>
<td>434</td>
<td>JAMA</td>
<td>463</td>
<td>Lancet</td>
</tr>
<tr>
<td>427</td>
<td>Lancet</td>
<td>441</td>
<td>Medical Care Research &amp; Review</td>
</tr>
<tr>
<td>410</td>
<td>Health Affairs</td>
<td>412</td>
<td>MedSurg Nursing</td>
</tr>
<tr>
<td>311</td>
<td>J Clinical Nursing</td>
<td>387</td>
<td>JAMA</td>
</tr>
</tbody>
</table>
Acknowledgements

George Swan, data wrangler

Jan Fransen and Shane Nackerud, Student Learning Outcomes Project, U Minnesota Libraries

Amy West, Data Librarian

Sanford Weisberg, Professor, Statistical Consulting Clinic, and Grace Jackson, Teaching Assistant

Abra Brisbin, Assistant Professor, U Wisconsin - Eau Claire Math Department
Questions?

• Katherine Chew  
  chewx002@umn.edu

• Jim Stemper  
  stemp003@umn.edu

• Caroline Lilyard  
  lily@umn.edu

• Mary Schoenborn  
  hawki003@umn.edu