Library Data and Student Success

Jan Fransen
Kristen Mastel
Shane Nackerud
Kate Peterson
David Peterson
Krista Soria

Library Technology Conference
Macalester College
March 14, 2012

Existing Measures

- Long history of measuring input, output, external perceptions of quality and satisfaction with library services
- Expenditures, staffing effects on retention
- Information literacy instruction
- Collections, facilities on enrollment decisions
  ...useful for management of library services, collections and resources but...

Example

- Kuh and Gonyea’s 2003
  - 300,000 students between 1984 and 2002 to the College Student Experiences Questionnaire
  - “library use does not appear to contribute directly to gains in information literacy and other desirable outcomes”

Student retention

- Haddow and Joseph 2010
  - Curtin University (4661 students)
  - Loans, workstation logins, logins to e-resources used to determine “extent of library use” with enrollment and demographic
  - “retained students showed higher levels of loans, PC logins and e-resource logins”
Using Swipe-card data

- Jones 2011
  - Georgia State University
  - Students, faculty, staff swipe before entering their library since 2002
  - Built analytics system to run queries
  - GPA, college, dorm, entry, ethnicity, gender, time, major, year in school, etc.

Library material usage

- Wong & Webb 2011
  - Hong Kong Baptist University (8,801 pairs of data)
  - Establish a mathematical correlation between student library material usage and cumulative GPA
  - Able to demonstrate impact on student learning

UK Library Impact Data Project

- 2010, University of Huddersfield
  - 700 courses (2005-2009)
  - 3 indicators of library usage (access to e-resources, book loans, access to the library)
- 2011, 8 UK institutions joined
  - 33,000 students, JISC funding
  - Grade, loans, e-resources accessed, times entered the library, school
- Focusing on non/low use and achievement

Call to Action

  - Assessment management systems
  - Develop systems to collect data on individual library user behavior
  - Record and increase library impact on student enrollment
  - Link libraries to improved student retention and graduation rates
  - Track library influences on increased student achievement
  - Demonstrate and develop library impact on student learning
Gym Bags and Mortarboards

- Student success measures
  - First Year Retention and 5 year graduation
- 5211 students in sample (2001)
- Tinto’s 1975 model of social and academic integration
- “able to demonstrate that actual usage of CRFs (campus recreational facilities) does have a positive association with academic success, even while controlling for other important academic, financial, and social fit factors.”

Layers of Data

- Office of Institutional Research Performance Data
  - Term and Cum GPA, ACT score
- Office of Institutional Research Demographics Data
  - College, Level, Major, Gender, Ethnicity, Age
- Libraries Data (13 Access Points)
  - Circulation, Digital, Instruction, Reference, and Workstation
A Word about Privacy

- In order to use OIR data, we must retain Internet ID
- For now, not aggregating anything about the library interaction other than count

<table>
<thead>
<tr>
<th>This</th>
<th>But not this</th>
</tr>
</thead>
<tbody>
<tr>
<td>Checked out X books</td>
<td>Titles</td>
</tr>
<tr>
<td>Attended X workshops</td>
<td>Which workshops</td>
</tr>
<tr>
<td>Reference interaction</td>
<td>Substance of interaction</td>
</tr>
<tr>
<td>Logged into library workstation</td>
<td>Date, location, duration</td>
</tr>
<tr>
<td>Used X digital resources of Y type</td>
<td>Which ones</td>
</tr>
</tbody>
</table>

Circulation

- Loans
  - Both new check-outs and renewals
  - Gathered by extracting data from Aleph transaction records
  - Internet ID and date of transaction
  - About 45% = Renewal data
- ILL Requests
  - Gathered by extracting data from ILLiad
  - ILLiad ID and date of transaction
  - Not all IDs were U of M Internet IDs

Digital

- Anytime someone logged into our digital resources with a U of M Internet ID
  - Database logins
  - E-Journal logins
  - E-Book logins
  - Website logins
- Due to IP based authentication, we did not track on campus usage of databases, e-journals, and e-books
  - Estimate - Missing 10-20% of our traffic
- This is only initial point of access, not actual usage

Reference

- Online reference transactions
  - Captured from QuestionPoint data
  - Some of the more difficult data to capture
  - We did not capture ref desk traffic or research consultations
- Peer Research consulting data
  - One-on-one assistance to develop research strategies
  - U of M student consultants
Instruction

- Workshop registrations
  - Captured by Drupal-based registration module
  - *Registration does not mean attendance*
- Intro to Libraries I workshop
- Intro to Libraries II workshop
- Course-integrated librarian instruction
  - Everyone registered for the course/section
  - *All students may not have been present*

Workstation

- U of M library workstation logins
  - Captured by Cybrarian application used to authenticate library users
  - *Does not include complete data from SMART Learning Commons*
- Reveals a flaw with regard to capturing “library as place”
  - Difficult to gather Internet IDs if students don’t give them to us

Library Data Layer

- 1,548,209 total transactions in all 5 categories
- 61,195 unique Internet IDs interacted with the Libraries in some identifiable way
- 37,674 people did something in only one of the five categories we measured
- 87 people did something in all five categories
- 9,324 people did only one of the 13 things we measured and *did it only once*
Questions we can’t answer alone

- How many undergraduates used the library?
- How many graduate students?
- Do some colleges use the libraries more than others?
- How many potential users are there?
- Are students who use the libraries more successful?

OIR Demographics Layer

- Office of Institutional Research
  - OIR collects and analyzes data to provide information for institutional planning, policy formation, and decision-making
- Key library data numbers:
  - 1,548,209 total transactions in 5 categories
  - 61,195 unique Internet IDs

Layers of Data

- Office of Institutional Research Demographics Data
  - College, Level, Major, Gender, Ethnicity, Age
- Libraries Data (13 Access Points)
  - Circulation, Digital, Instruction, Reference, and Workstation

- 77% of Undergrads made use of the Libraries during the Fall Semester 2011

- 85% of Grad Students made use of the Libraries during the Fall Semester 2011 (including professional schools)
Layers of Data

Office of Institutional Research Performance Data
- Term and Cum GPA, ACT score

Office of Institutional Research Demographics Data
- College, Level, Major, Gender, Ethnicity, Age

Libraries Data (13 Access Points)
- Circulation, Digital, Instruction, Reference, and Workstation

Undergrad Instruction Usage

Undergrad Reference Usage

Undergrad Cumulative GPA as of Fall 2011
Inferential Analyses

- First-year students (non-transfer, $n = 5,368)$
- Examined two outcomes: first-semester grade point average and retention
- Many ways to slice the data:
  - Any use of the library
  - Type of library use
  - Frequency within type of library use
First Analyses

- Measures:
  - Use of library (71.3%)
  - Demographics:
    - Gender (F = 47.8%)
    - Race/ethnicity (SOC= 18.4%)
    - Pell grant (22.3%)
    - Veteran status (.6%)
    - First-generation (25.9%)
  - College environment:
    - Freshmen seminar (27.8%)
    - Access to Success program (8.8%)
    - Dorm (85.2%)
  - Prior academics
    - ACT/SAT scores (M = 27.49)
    - AP credits (n = 3137, M = 8.73)

- Analyses:
  - Multiple linear regressions (gpa)
  - Logistic regression (retention)

GPA Results

- Controlling for demographics, college environment, and academic variables:
  - Using the library one time was associated with a \(0.23\) increase in students’ gpa holding other factors constant
  - A one-unit increase in types of use was associated with a \(0.07\) increase in gpa
Additional GPA Results

• Controlling for the same variables, we examined using different types of sources at least once (dummy-coded):
  – Course integrated instruction: -.11
  – Database: .14
  – EJournal: .10
  – Loan: .11

*note: 12 outliers removed

Additional GPA Results

• Controlling for the same variables, we examined using different types of sources by frequency (so, a one-unit increase is associated with...):
  – Course integrated instruction: -.08
  – Database: .01
  – E-Journal: .004
  – Workstation: .006
  – Reference: .08

Retention Results

• Controlling for the same variables, we examined retention:
  – Students who used the library at least once were 1.54 times more likely to re-enroll
  – For every one-unit increase in the types of library use, students were 1.1 times more likely to re-enroll
Additional Retention Results

• Controlling for the same variables, we examined retention:
  – Students who had “Intro to Libraries 2” library instruction were 7.58 times more likely to re-enroll
  – A one-unit increase in database uses was associated with students being 1.03 times more likely to enroll

How

• Start Small
  – What are you collecting? What’s easy?
  – Loan, E-Resources, Workstation + ID
  – Maybe start with demographics
• Contact your “Office of Institutional Research”
  – Might be one person
• Or gain access yourself

How

• ACRL Value of Academic Libraries
  – http://www.acrl.ala.org/value/
• UK Library Impact Data Project study toolkit

Next steps

• More analysis by Libraries and OIR
• Keep going in the Spring Semester
  – Do we need more data and analysis going forward?
  – Do we collect the same data? Do we need to make an attempt to gather things we aren’t gathering now?
Questions?

Technical details? Contact us!

Contact information

- Jan Fransen (fransen@umn.edu)
- Kristen Mastel (meye0539@umn.edu)
- Shane Nackerud (snackeru@umn.edu)
- Kate Peterson (katep@umn.edu)
- David Peterson (pete2786@umn.edu)
- Krista Soria (botx0006@umn.edu)

Resources

  http://www.acrl.ala.org/value/