Optimizing a library website for student research: methodologies for comparing metrics between Encore and Google Scholar

Presented & Prepared by:
Lindsay Ozburn
Assistant Librarian & Assessment Coordinator

Liz Woolcott
Head of Cataloging and Metadata Services

Ryan Bushman
Graduate Research Assistant, Math and Statistics

Margaret Winward
Programmer/Analyst II, Library IT
LEARNING OBJECTIVES

Sampling Methodology, Research Design, and Analysis Process

From our presentation and subsequent proceedings, we offer:

1. A straightforward example of implementing A/B testing to evaluate user information searching preferences
2. Easy-to-understand, replicable statistical sampling methodology to enhance the validity of user testing data
3. A snapshot of the data generated from this type of experiment, how to gather them, and what you can do with it

Experiment Design
- Research question development
- Appropriate methodology selection

Sampling Approach
- Types of sampling we considered
- Best practices for sampling method selection
- Quick overview of sampling execution

Experiment Execution
- How we captured our data
- Considerations for and overview of the software
- Executing the experiment design

Analysis Process
- Data volume generated
- How we analyzed qualitative data as a group
- How we analyzed quantitative data
- Basic findings
Parts I & II

Experiment Design & Sampling Approach
Sampling
We conducted a stratified random sample with strata including STEM and Non-STEM students.

Design
- Randomized group allocation
- Pre-survey
- Interaction with Interfaces
- Post-survey

Analysis
- Statistical analysis using SAS
- Qualitative analysis using Airtable and Excel
Part III

Experiment Execution
EXPERIMENT TOOLS

01 PRE-SURVEY (QUALTRICS/LOOP11)
- Informed Consent
- Establishing searching familiarity
- Establishing academic baselines

02 INFORMATION SEARCHING TASK LIST (LOOP11)
- 10 questions, split into two groups
- Variety of item search types, including known and topical

03 SCREEN CAPTURE SOFTWARE (LOOP11)
- Remote usability testing tool.
- Online survey creation for users to complete starting on websites of your choice while recording their screen actions.

04 POST-SURVEY (QUALTRICS)
- Perceived challenges and positive qualities
- Interface preference based on item type and overall
Task Request

Interface B - Task 3

Your class acquaintance recommended a book for you. They said it was called something like teaching cues for sport skills – but you can’t remember the author’s last name. Please find the full title and author.
Task Research
Interface B - Task 3

Your class acquaintance recommended a book for you. They said it was called something like teaching cues for sport skills— but you can’t remember the author’s last name. Please find the full title and author.
Task Response

What is the full name of the book about teaching cues for sport skills and its author?

[Input Field]
Part IV

ANALYSIS PROCESS
DATA COLLECTION

Video

Heatmap

Clickstream

Data
Out of 14 tasks, respondents reported Time Management to be the most difficult.

Qualitative Analysis: Airtable

HAVE A VERY OR SOMEWHAT DIFFICULT TIME WITH TIME MANAGEMENT
Our statistical model

\[ Y_{ijk} = \mu + O_i + S_j + P_k + \varepsilon_{ijk} \]

- \( Y_{ijk} := \) Difference in average completion time
- \( \mu := \) Overall mean
- \( O_i := \) Effect due to order
- \( S_j := \) Effect due to stratum
- \( P_k := \) Effect due to prior research preference
- \( \varepsilon_{ijk} := \) Random error

**Fitting the model**

We used the PROC GLM function to fit a general linear model. This function also conducts analysis of variance to determine whether individual factors had statistically significant effects.

**Diagnostics**

When analyzing data with outliers removed, a data transformation on the response variable was necessary to meet assumptions of normality and constant variance.

**Analysis**

Our model found order and stratum to be statistically significant. Order seems to be due to participants performing faster on the second interface either due to familiarity or fatigue. When digging into the stratum variable, we found that STEM students performed nearly equally as well in both interfaces whereas non-STEM students performed better on the single-tab interface.
Part V
GENERAL CONCLUSIONS AND TAKE-AWAYS
GENERAL CONCLUSIONS AND TAKE-AWAYS

RQ1: What is the average completion time for tasks performed in each interface?
- Difference when you take into account order or strata
- Completed tasks faster on second interface used
- On average, STEM students performed tasks with similar speed regardless of the interface used, while non-STEM students performed tasks faster in the single-tab interface

RQ2: How many actions does it take to perform tasks in each interface?
- Minimal difference overall
- Weak association with fewer pages visited in single-tab interface

RQ3: What are the benefits and drawbacks participants perceived in each interface?
- Participants preferred single tab interface more frequently
- Participants who preferred single search liked simplicity of one search bar
- Participants who preferred dual tab like the option to choose search
- Filtering was difficult in both interfaces
- WebPac catalog (in dual tab) felt old and clunky
- Authentication was a deterrent in single tab
FOR QUESTIONS AND RESOURCES:


○ Loop 11

○ Qualtrics

○ Please visit our conference proceedings for all the extra goodies!

Contact us!
Lindsay Ozburn - lindsay.ozburn@usu.edu
Liz Woolcott - liz.woolcott@usu.edu
Ryan Bushman - ryan.bushman@usu.edu
Margaret Winward - margaret.winward@usu.edu