FSTAR Evaluation
Testing the Efficiency of Literature Searches
Kevin M. Curtin, PhD
Jasmin Khangura
Talk Outline

- Introduction / Motivation
- Experimental Design
- Methods
- Expected Results
- Issues / Questions
Introduction / Motivation

- Introduction / Experience
  - Associate Dean for Academic Affairs for the College of Science
  - Participated in evaluations of several projects with IAFC
    - Volunteer Firefighter Recruitment and Retention
      - Four Phases in two states
      - SAFER grants
    - RESCUES
      - Researchers Creating Useable Emergency Solutions
      - Dovetailed with FSTAR earlier phase
      - Tested fact sheets against full research papers
      - Focus groups regarding concerns
    - Now with FSTAR II

- Motivation
  - One goal of FSTAR is to more effectively provide access to research regarding the fire service
    - Can we test (evaluate) whether or not that goal is being met?
    - The work described herein has been submitted to the Institutional Review Board with regard to human subjects research, and has been approved
Experimental Design

- Essentially – making comparisons in literature search performance
  - Two comparisons – within groups and across groups
  - Across groups – two populations
    - Members of the fire service
      - Broadly defined…anyone who self-identifies as a member of the fire service
      - Firefighters, staff, support personnel and others may have a legitimate need to search the literature for research regarding the fire service
    - Graduate students
      - Accustomed to performing literature searches
      - Relatively easy population to access
    - Idea is to see if there is any significant difference between populations, or if the effectiveness of the search is the same regardless of background
  - Within groups (and within entire population combined)
    - Some participants are asked to search using the FSTAR repository
    - Some participants are asked to search using their favorite browser and the internet
Methods

- Data Collection – Ongoing as we speak
  - Ask for participation and provide information seeking informed consent
  - Ask participant to search for one of three articles, either with or without FSTAR
  - Capture their search activity – SpyAgent records:
    - Computer usage time; Websites visited; Windows opened; Applications ran; Documents and files opened; Keystrokes and mouse clicks; Screenshots
Methods (II)

- **Analysis**
  - Statistical comparisons across and within populations
  - Descriptive statistics
    - How long did it take to search, and who took longer? By how much?
    - How many failed searches were there? By which group?
    - How many mouse clicks, websites visited, etc. to reach the goal
    - Compared across populations/sub-populations, and by research article
  - Inferential statistics to test whether the differences in performance (if any) are statistically significant
    - Testing whether the performance of the two populations is significantly different
    - t-test is the classic test

\[
t = \frac{\bar{X}_1 - \bar{X}_2}{s_{X_1X_2} \cdot \sqrt{\frac{1}{n}}}
\]

where

\[
s_{X_1X_2} = \sqrt{s_{X_1}^2 + s_{X_2}^2}
\]
Expected Results / Issues / Questions

Expected results
- Expect that FSTAR searches will be faster/more efficient
  - Perhaps not dramatically faster
  - There is a more direct path to the articles with less internet “noise”
- Expect more failed searches using broad internet access

Potential Issues
- Always concerned with “n” – number of observations
- Short timeline – have to get it right the first time

Questions/Comments?