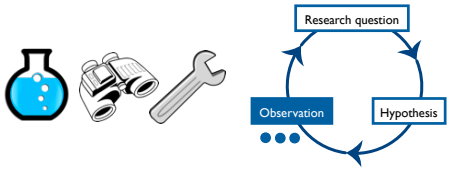


HCI Research Literacy II

Experimental Research Applied to a Text Entry Research Project

So You've Invented a New Keyboard Layout?

- Scenario: You have designed a new keyboard layout, and you want to know how good it is
- Strategy: compare it with existing techniques
- Basic research questions
  - How fast is it?
  - How accurate is it?
- In-class exercise:
  What are independent (IV) and dependent variables (DV)?

Measures (DV)

- Speed
- Accuracy

Trade-off

- Qualitative feedback
  - Comfort
  - Device impressions
  - Report as anecdotes or quotes
- In-class exercise:
  How would you make an operational definition of speed?

Speed Measures: Words per Minute

$$\mathrm{WPM} = \frac{|T| - 1}{S} \times 60 \times \frac{1}{5}$$

$|T|$ Length of the transcribed string

-1 Timing begins after the first character was pressed

S Duration in seconds

$\frac{1}{5}$ Estimated length of a word: 5 characters including spaces (Yamada, 1980)

\+ Easiest measure, you just need a watch

– Disregards errors in the final text

- Alternative: insist on the user correcting all errors, but may lead to user frustration

– Disregards the process of entering

- E.g., It doesn't matter how many times you pressed the backspace key.

Speed Measures: Keystrokes per Second (KSPS)

$$\text{KSPS} = \frac{|IS| - 1}{S}$$

$|IS|$ Length of the input stream (all characters including backspaces)

+ Reflects the process during text entry (every keystroke counts)

– May not reflect real use

- E.g., a fast but error-prone keyboard may have a high KSPS

Accuracy Measures: Keystrokes per Character (KSPC)

$$\text{KSPC} = \frac{|IS|}{|T|}$$

$|IS|$ Length of the input stream

$|T|$ Length of the transcribed string

+ Simplicity

– No distinction between backspaced characters that are initially correct vs. those that are initially incorrect

- Check (Wobbrock, 2007) for discussion of other measures

In-depth: "Measurement of Text Entry Performance." Wobbrock, 2007

Other Variables

- How should I treat other variables: age, gender, finger lengths, hand size, etc.?
  - Include those that make sense as IVs ⇒ more experimental conditions!

Leave as random ⟷ Control

Reflects variation in natural use

Higher confidence to infer causality in the results

↑ external validity

↑ internal validity

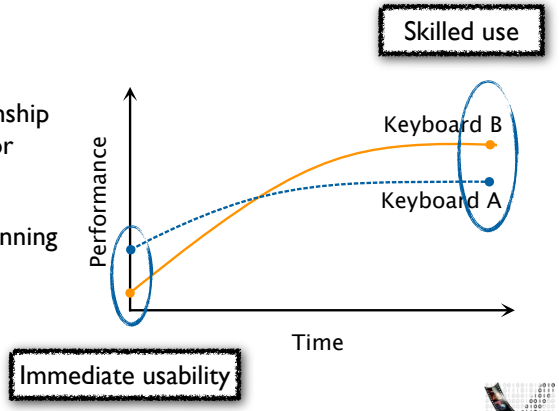
Internal vs. External Validity

- A study has internal validity if it produces a single, unambiguous explanation for the relationship between two variables
- External validity refers to the extent to which we can generalize the results to people, settings, times, measures, and characteristics other than those used in that study
- Always a trade-off, strike an appropriate balance depending on the goal of your research

Definitions from (Gravetter and Forzano, 2012)

Effect from Learning

- Learning curve: relationship between experience (or time) and performance
- Rapid raise at the beginning follow by a plateau



Experimental Design

- Usually preferred: within-group design
  - Minimizes confounding effects from the behavioral differences between participants
- Sometimes, we need a between-groups design
  - E.g., when testing whether a keyboard favors users with right-handedness over those with left-handedness
  - When there are interferences between conditions, e.g., different keyboard layouts on the same hardware

Choosing the Task

- Copy text
  - Exclude behaviors that may compromise the measures, e.g., pondering what to write
  - Allows identifying error because the content is known
  - Can control the distribution of letters and words
- Create own text
  - Mimics typical usage
- Compromise: Read and memorize a short sentence before entering

Choosing the Text

- English phrase set: MacKenzie and Soukoreff (CHI 2003)
- 500 phrases in moderate length, easy to remember, and representative of the target language
- Ignore case and enter all characters in lowercase.

+ Allows replication

- Examples:

```
there will be some fog tonight
round robin scheduling
time to go shopping
frequently asked questions
```

there will be some fog tonight

there w_

Coming Up Next…

- Lab: Dissecting the evaluation section of a text entry research paper
- Next week: Research Literacy III: Reading the results section
- Assignment Zero…

Assignment Zero: Writing a Review for Dummies

- Write a review about the evaluation section for one of these papers:
  - Typing on Flat Glass[1] (Findlater et al., CHI '11) — Even-number groups
  - The 1Line Keyboard[2] (Li et al., UIST '11) — Odd-number groups
- Required reading for background:
  - Evaluation of Text Entry Techniques[3] (MacKenzie, 2007) — REQUIRED
- Peer grading
  - In groups of 3, select one of the papers
  - Individually review the evaluation sections in the paper
  - Grade each other's review
  - Structured review form and grading form will be posted online
  - Submission: 3 × original reviews and 6 × peer grading feedback
  - Deadline: Tuesday, April 23rd, 2013 before 12:00 noon

1 http://dl.acm.org/citation.cfm?id=1979301
2 http://dl.acm.org/citation.cfm?id=2047257
3 http://www.yorku.ca/mack/chapter4.html

Typing on Flat Glass
Available at: http://dl.acm.org/citation.cfm?id=1979301

The 1line Keyboard