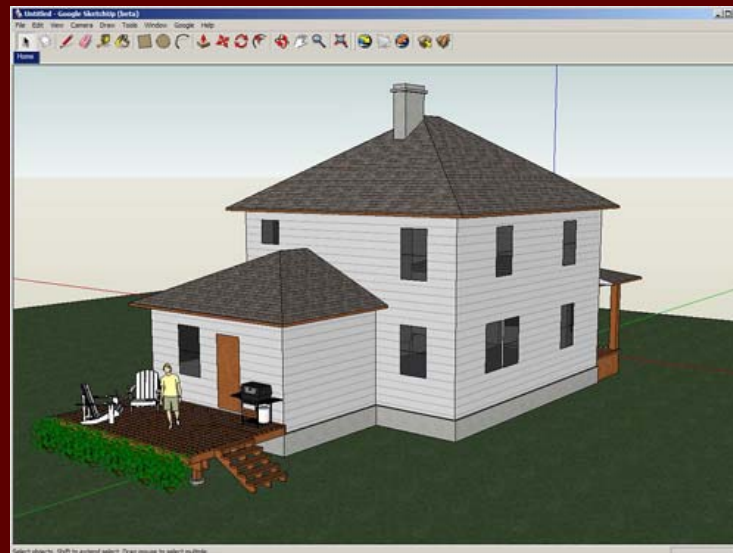


Undo and Erase Events as Indicators of Usability Problems



David Akers^{1,2}, Matthew Simpson², Robin Jeffries², Terry Winograd¹

¹Stanford University HCI Group

²Google, Inc.



Word processors



Image/video editors



3D modelers



Page layout applications

Creation-Oriented Applications



Word processors



Image/video editors



3D modelers

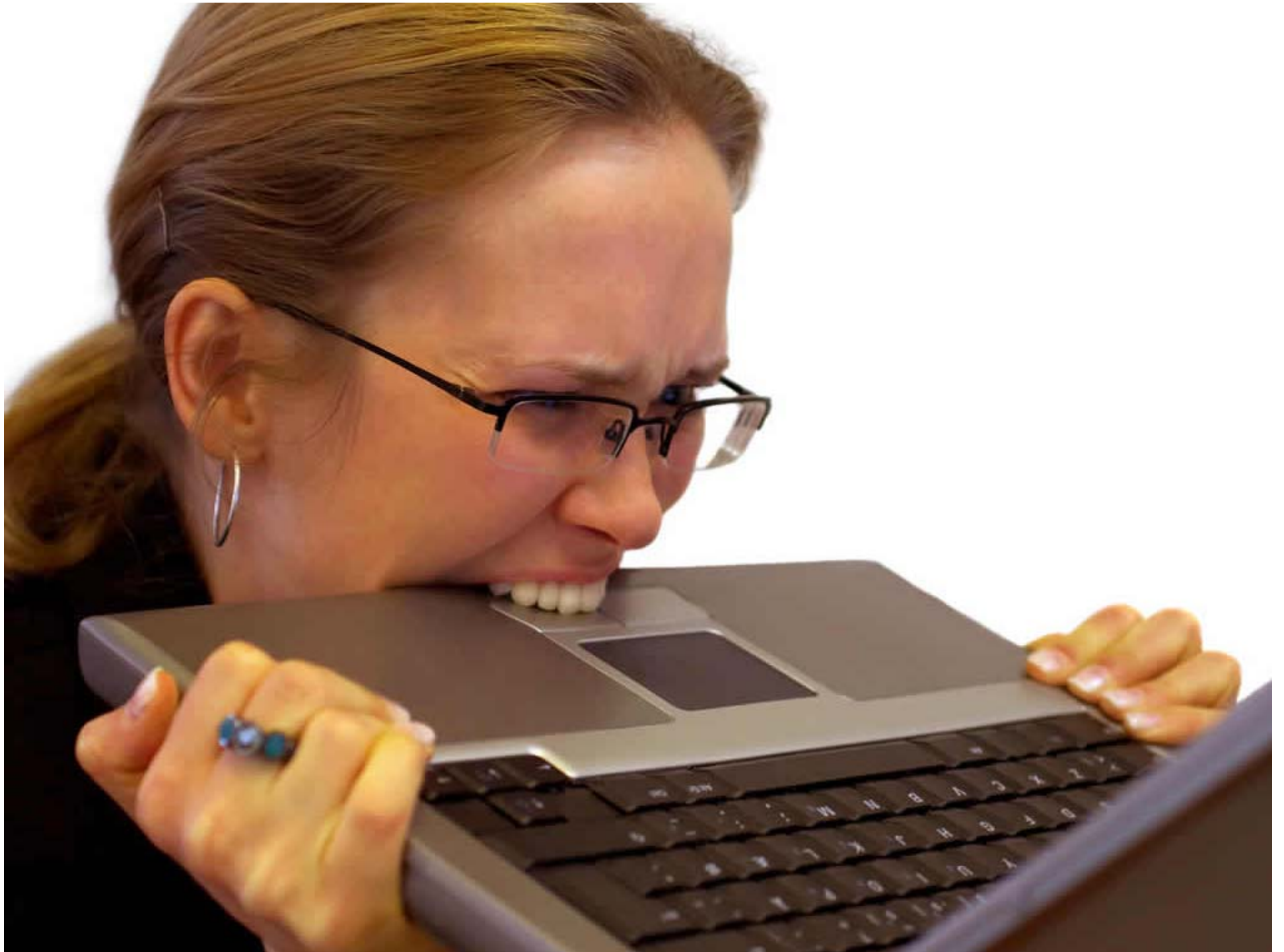


Page layout applications

**How do we evaluate the usability
of creation-oriented software?**

expensive, small scale





How can we automatically detect interaction breakdowns in creation-oriented software?

Pro

n:

Lo
un

king ev
, cance



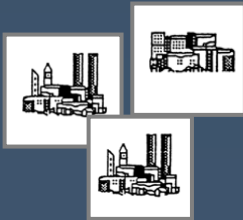
Backtracking has multiple roles...



Recovering from errors



Learning/exploring an interface



Exploring design alternatives

false
alarms?

Which method to compare against?

Patterns of user interface events

Lag Sequential Analysis (LSA) [Sackett 1978]

Fischer's cycles [Fischer 1991]

Maximal Repeating Patterns (MRP) [Siochi and Hix 1991]

Application-specific events [Swallow et al. 1997]

Expectation-Driven Event Monitoring [Hilbert and Redmiles 1998]

Physiological indicators

Skin conductance, heart rate, brain waves, etc. [Andreassi 2000]

Self-reported difficulties

User-reported critical incident technique [Hartson and Castillo 1998]

Retrospective user-reported critical incident technique [Capra 2003]

Research question

As indicators of usability problems in creation-oriented software, how do backtracking events compare to self-reported critical incidents?

Measure:

Number of usability problems indicated
Types of usability problems indicated

Outline

Motivation and Research Question

Google SketchUp

Design of the Experiment

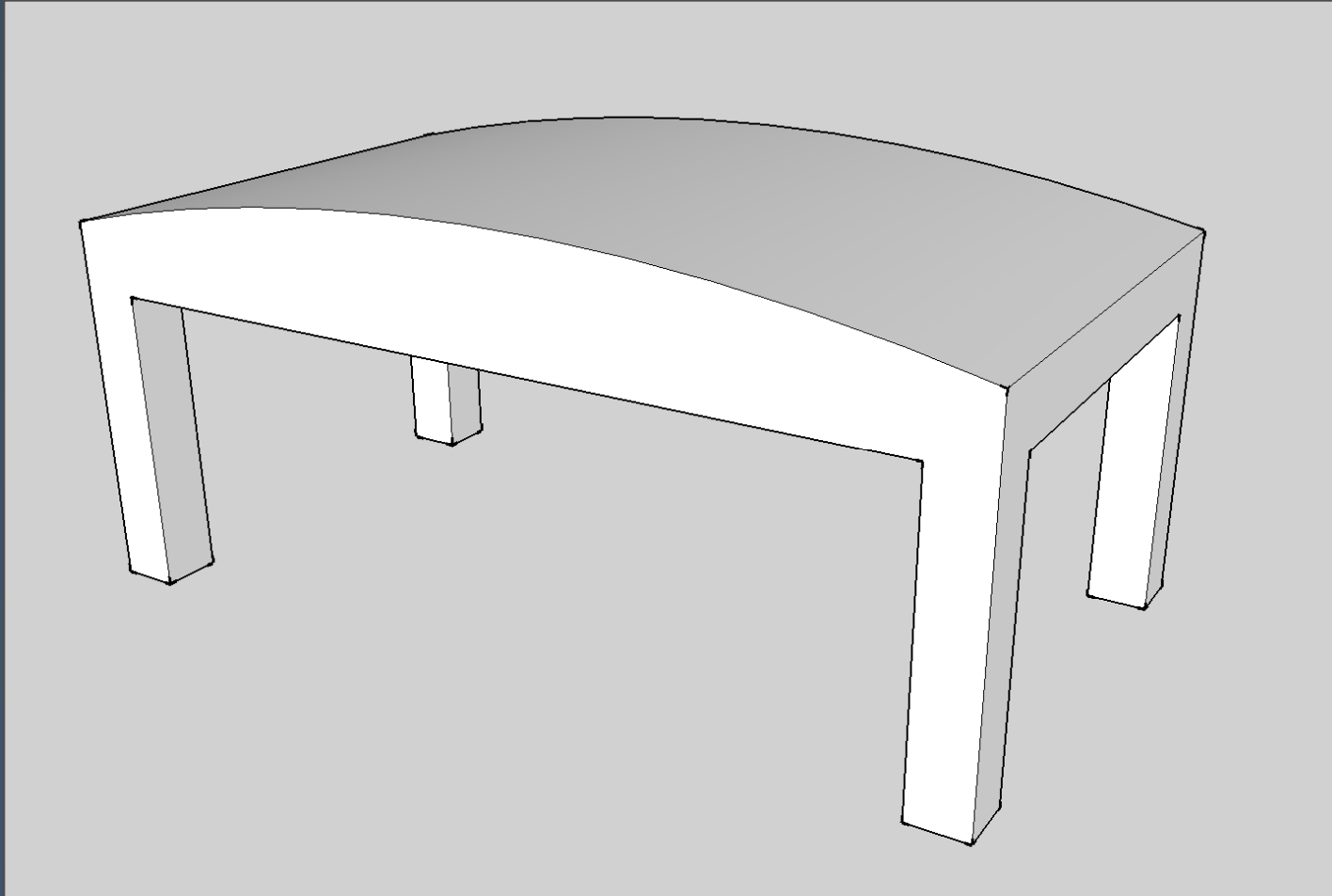
Results from the Experiment

Reflections and Future Work

Google SketchUp



SketchUp Demo



Outline

Motivation and Research Question

Google SketchUp

Design of the Experiment

Results from the Experiment

Reflections and Future Work

Instrumenting SketchUp

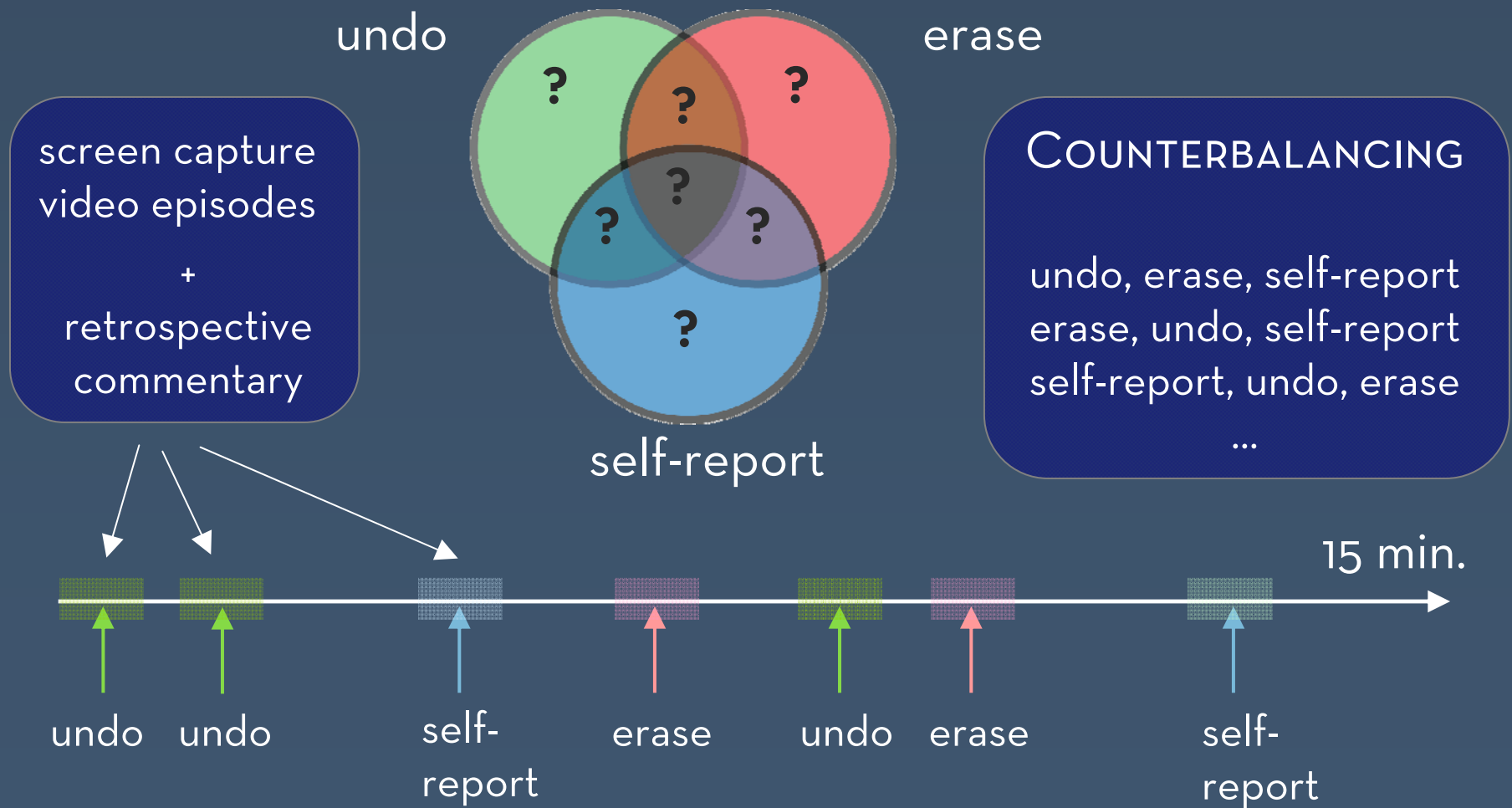
Without modifying the source code to SketchUp, we are able to record:

- Undo events

- Erase events

- Self-reported issues (“Report an Issue” button)

Experimental design (overview)



Recruitment

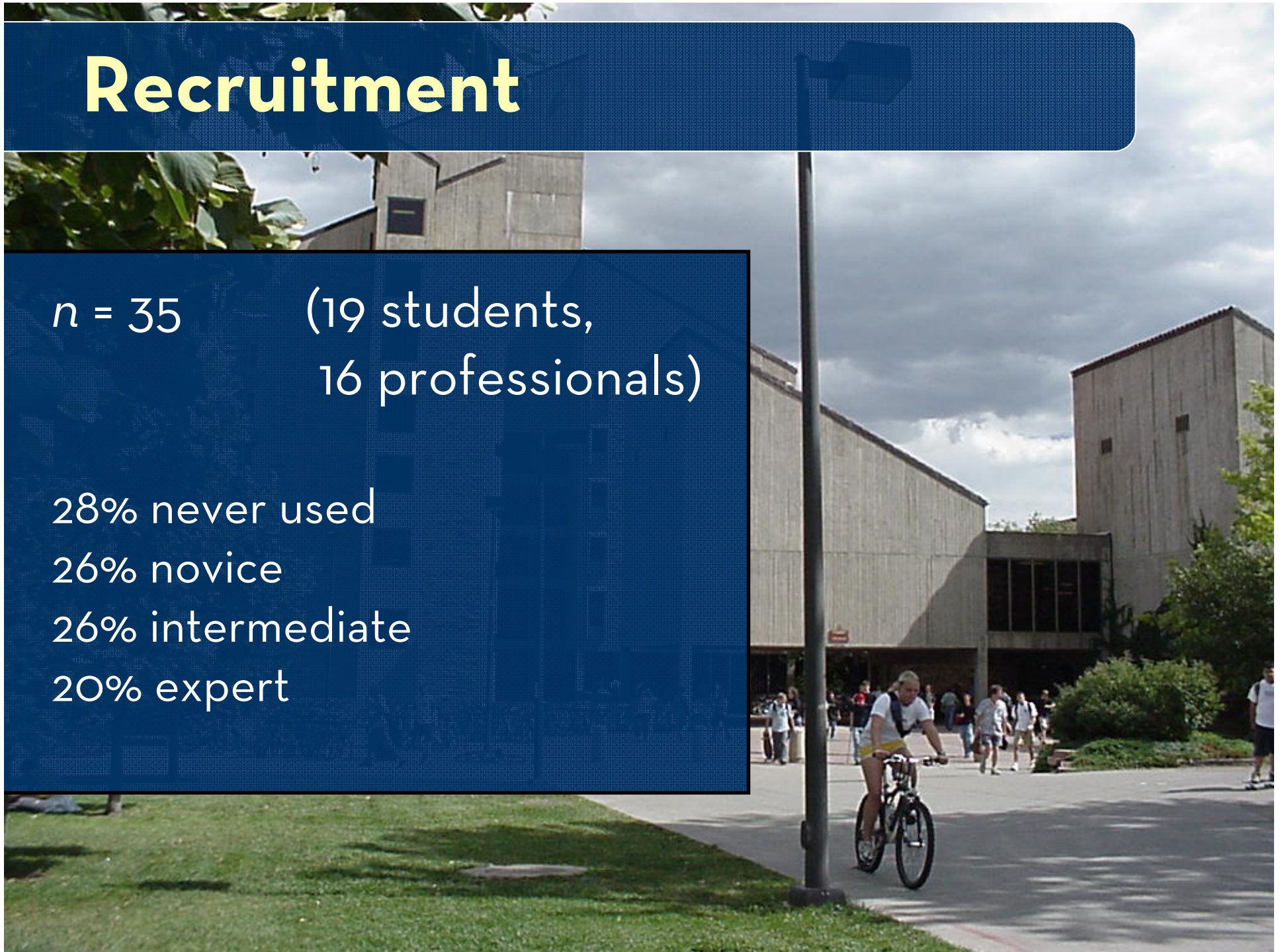
$n = 35$ (19 students,
16 professionals)

28% never used

26% novice

26% intermediate

20% expert



Experimental setup



Experimental protocol

Training in SketchUp

15 min

Training in self reporting

20 min

Practice

10 min

Modeling tasks

15 min

Retrospective commentary

30 min

SketchUp New User Series

Video 1: Concepts of SketchUp

Video 2: Drawing Shapes

Video 3: Modify Tools

Experimental protocol

Training in SketchUp

15 min

Training in self reporting

20 min

Practice

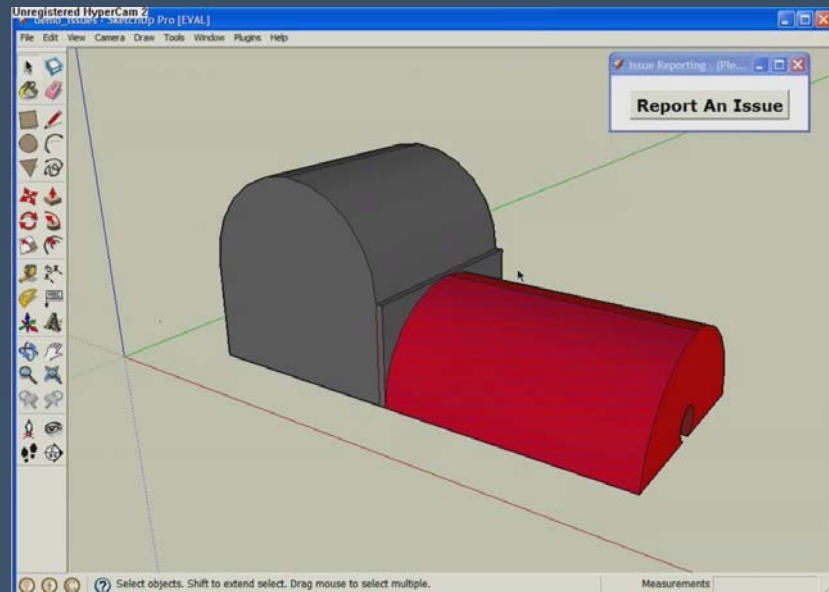
10 min

Modeling tasks

15 min

Retrospective commentary

30 min



(modeled after Hartson and Castillo [1998])

Experimental protocol

Training in SketchUp

15 min

Training in self reporting

20 min

Practice

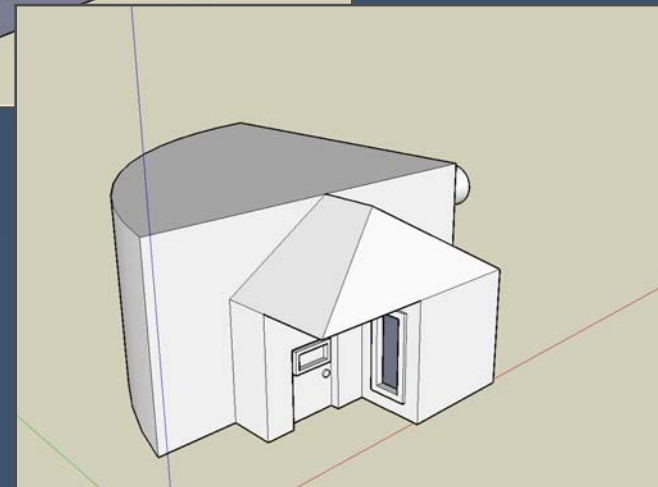
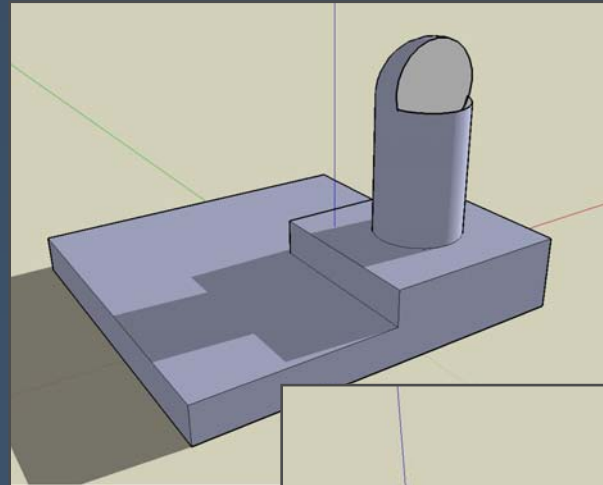
10 min

Modeling tasks

15 min

Retrospective commentary

30 min



Experimental protocol

Training in SketchUp

15 min

Training in self reporting

20 min

Practice

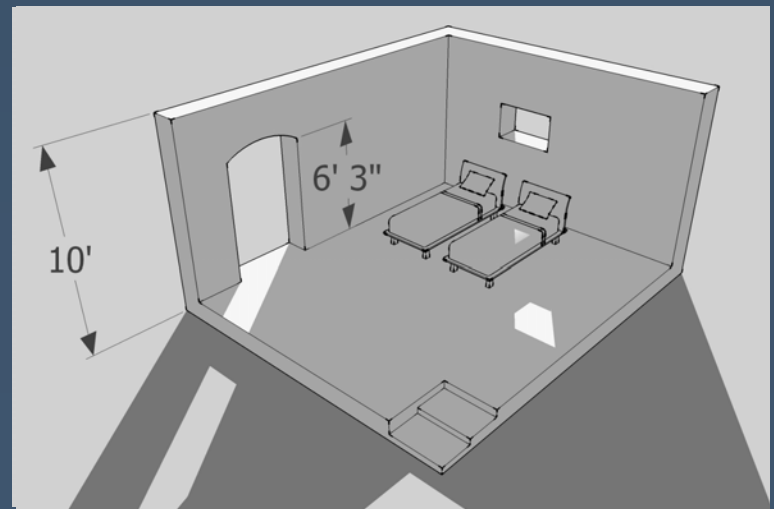
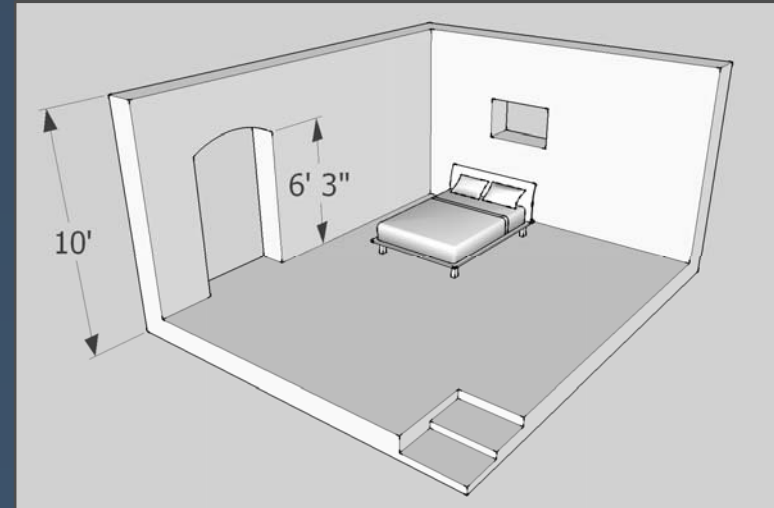
10 min

Modeling tasks

15 min

Retrospective commentary

30 min



Experimental protocol

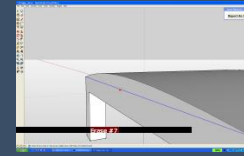
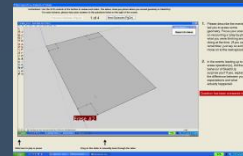
Training in SketchUp
15 min

Training in self reporting
20 min

Practice
10 min

Modeling tasks
15 min

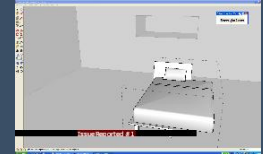
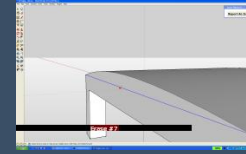
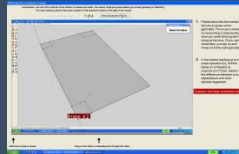
Retrospective commentary
30 min



- 1) Describe the events that led you to [undo/erase/report an issue].
- 2) In the events leading up to this, were you surprised by the software? If so, how were you surprised?
- 3) Did you find a way around the problem? If so, how?

Experimental protocol

Training in SketchUp
15 min



Training in self reporting
20 min

Practice
10 min

Modeling tasks
15 min

Retrospective commentary
30 min

- 4) Did you report this as an issue?
- 5) If you did not report this as an issue, why do you think that you did not?

paired retrospectives

Outline

Motivation and Research Question

Google SketchUp

Design of the Experiment

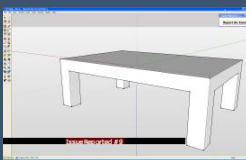
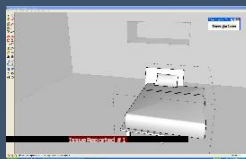
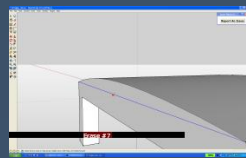
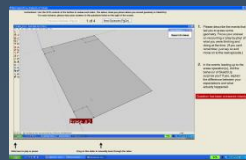
Results from the Experiment

Reflections and Future Work

From episodes to problems

Episodes (353)

Problem Instances (215)



FALSE ALARMS

50% of erase episodes

cosmetic edges leftover
from modeling process

temporary construction
lines

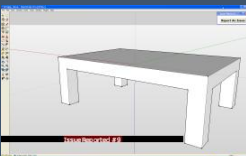
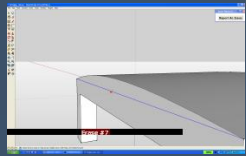
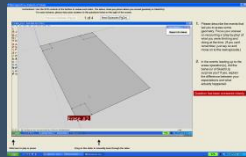
2% of self-report
episodes

accidentally pressing
'report' button

0% of undo episodes

From episodes to problems

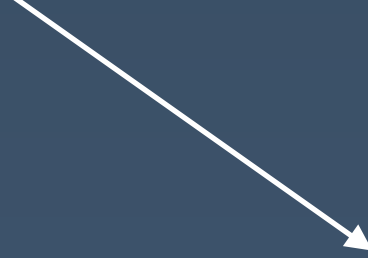
Episodes (353)



Problem Instances (215)



Problems (95)



Problem severity

Three SketchUp experts rated each problem for its:

estimated frequency in an average application session

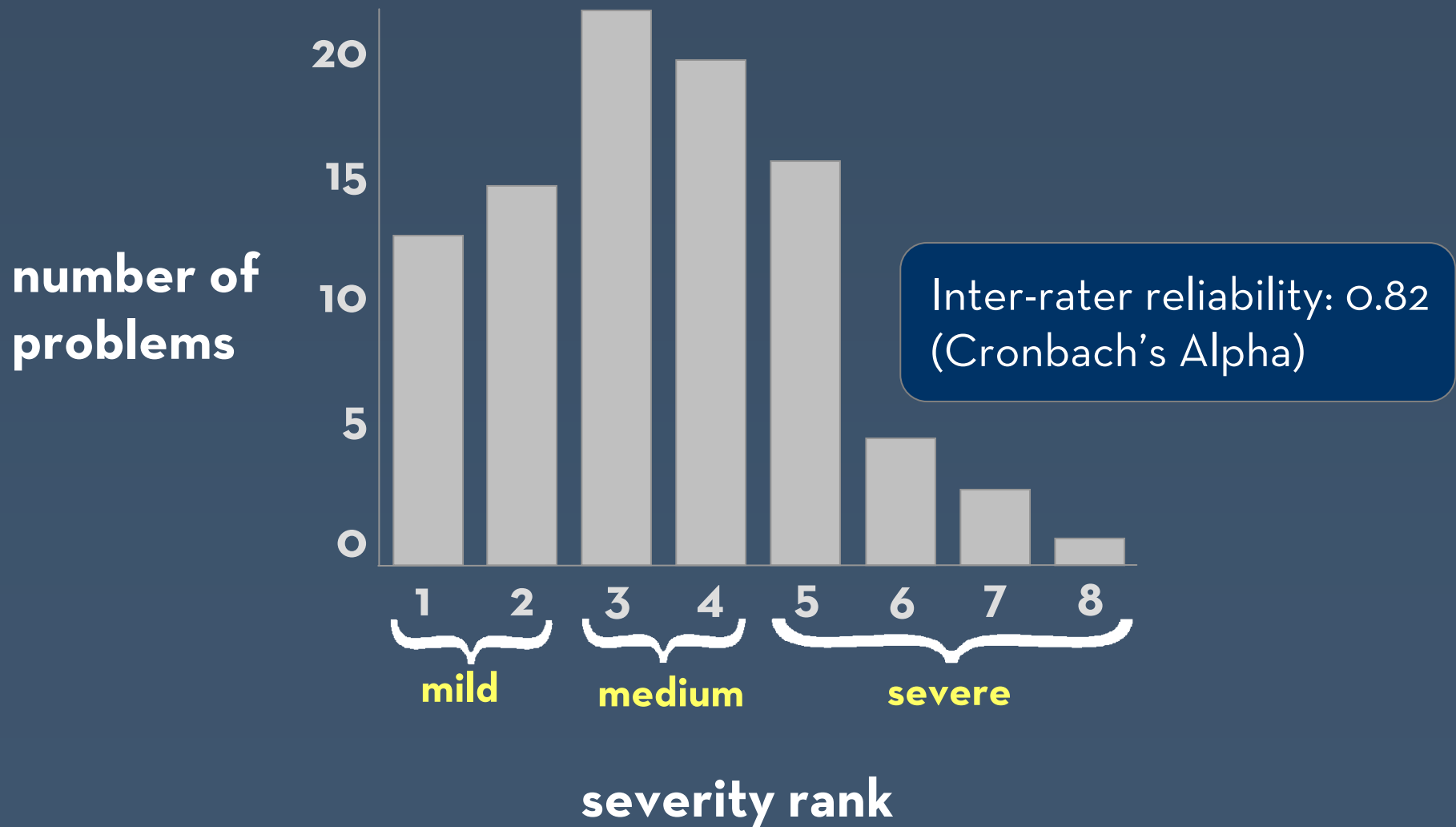
and

estimated impact and persistence on the user

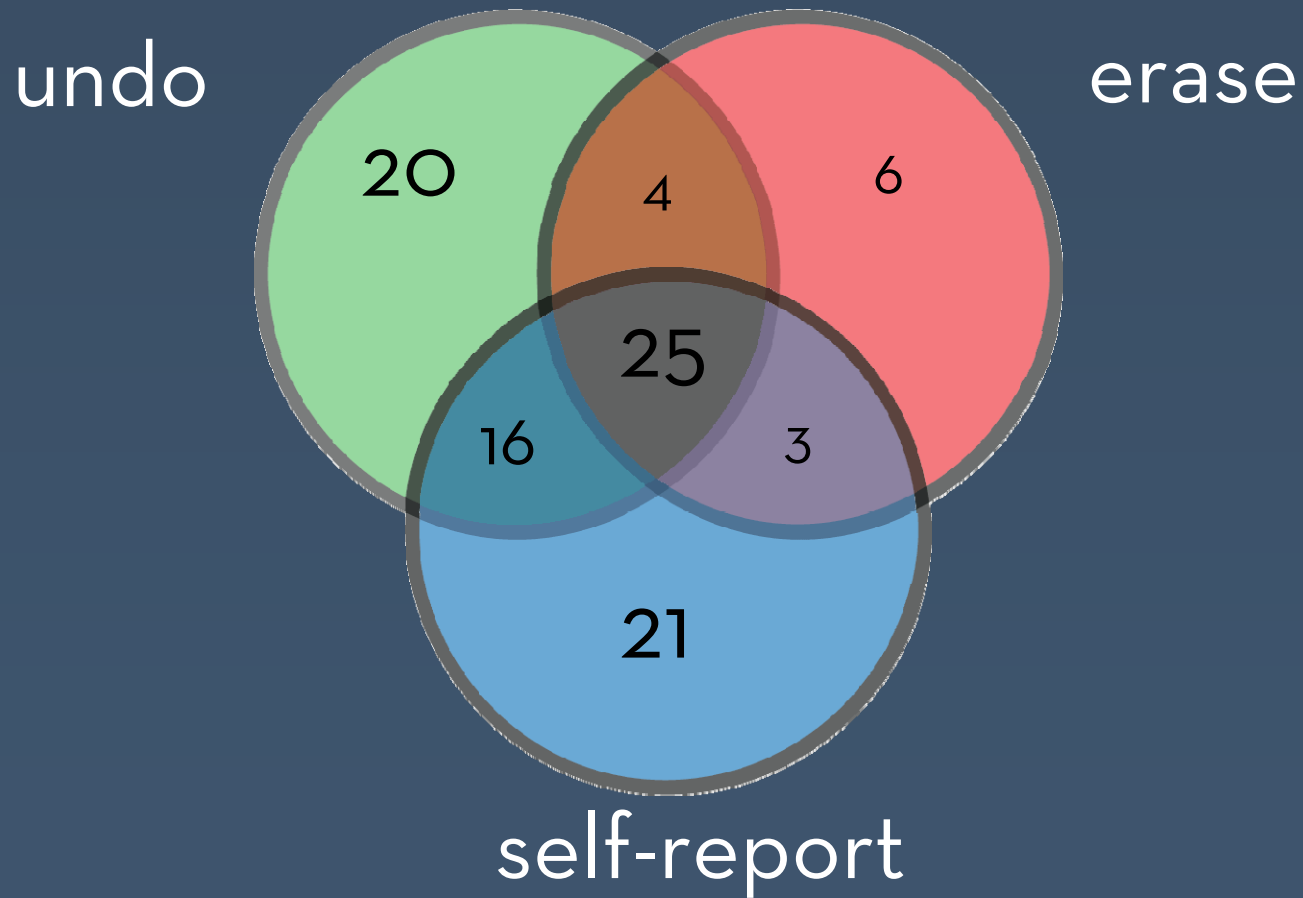


severity rank (1-8)

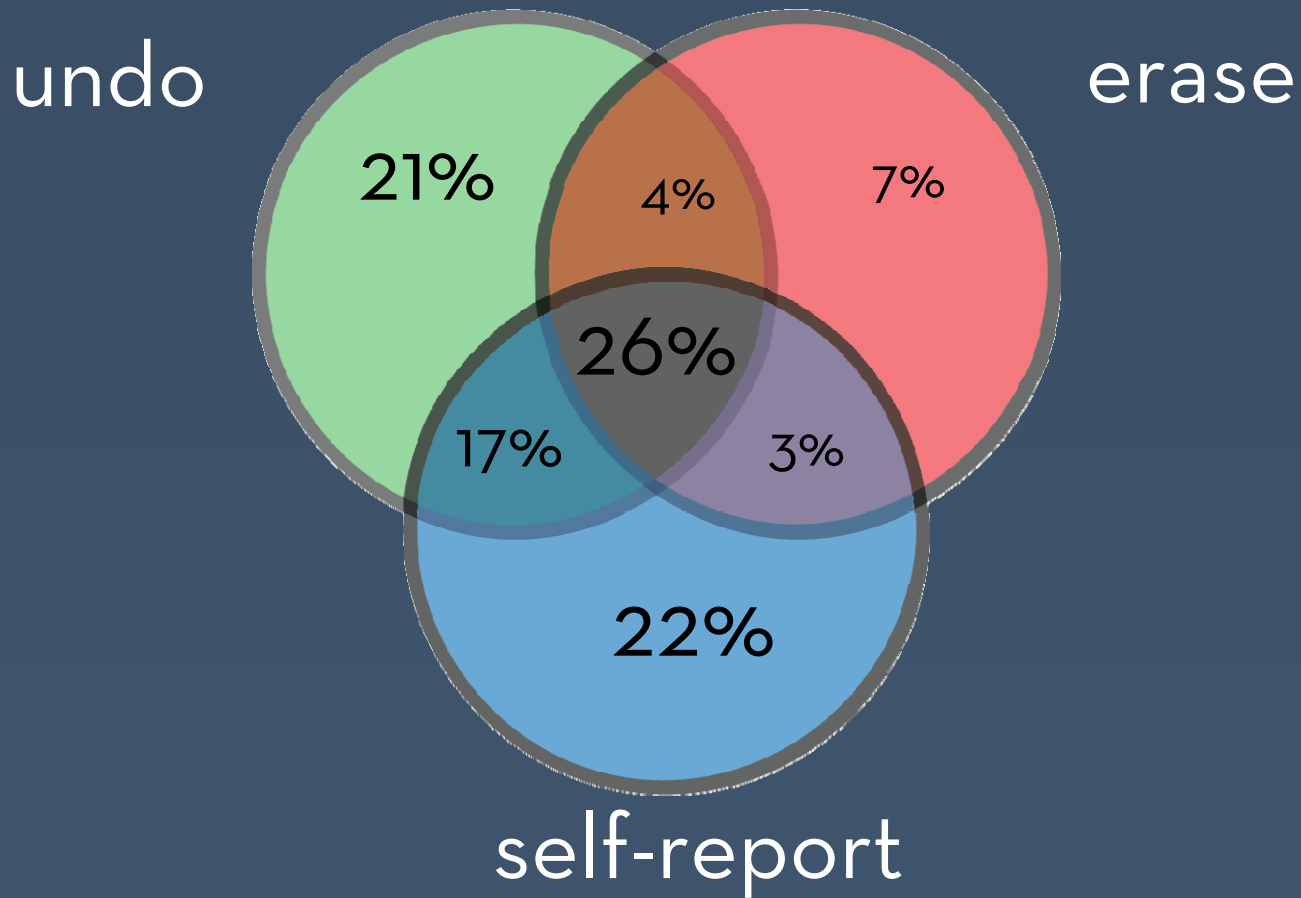
Problem severity histogram



Problem counts



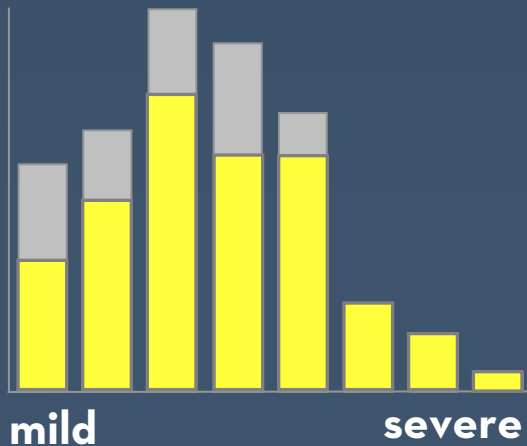
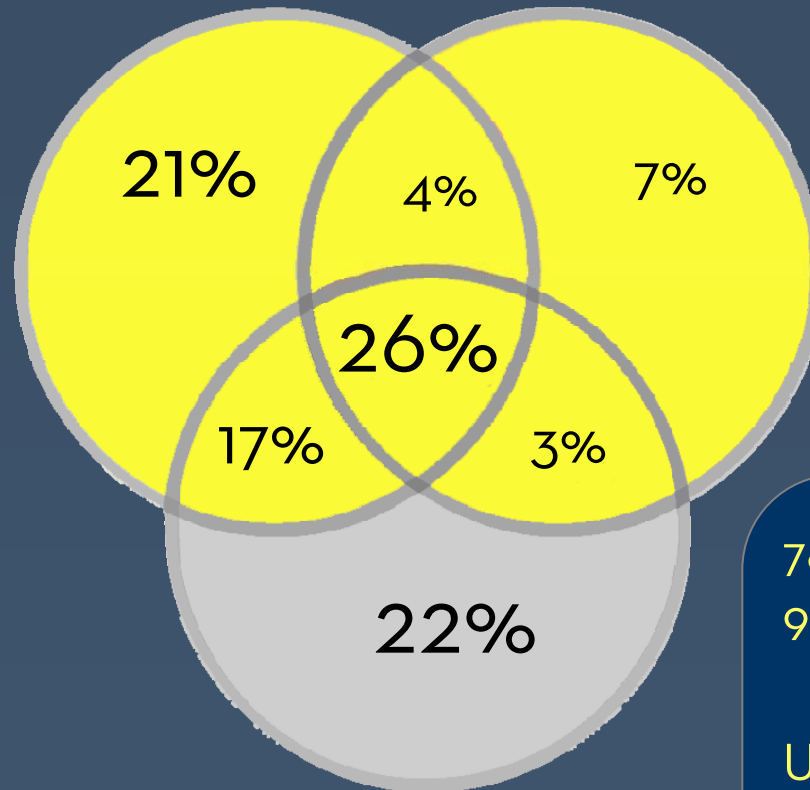
Problem counts



Problems found by undo/erase

undo

erase



SEVERITY HISTOGRAM

self-report

78% of all problems
92% of severe problems

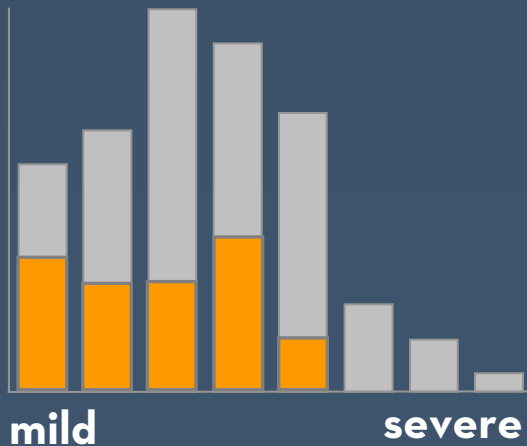
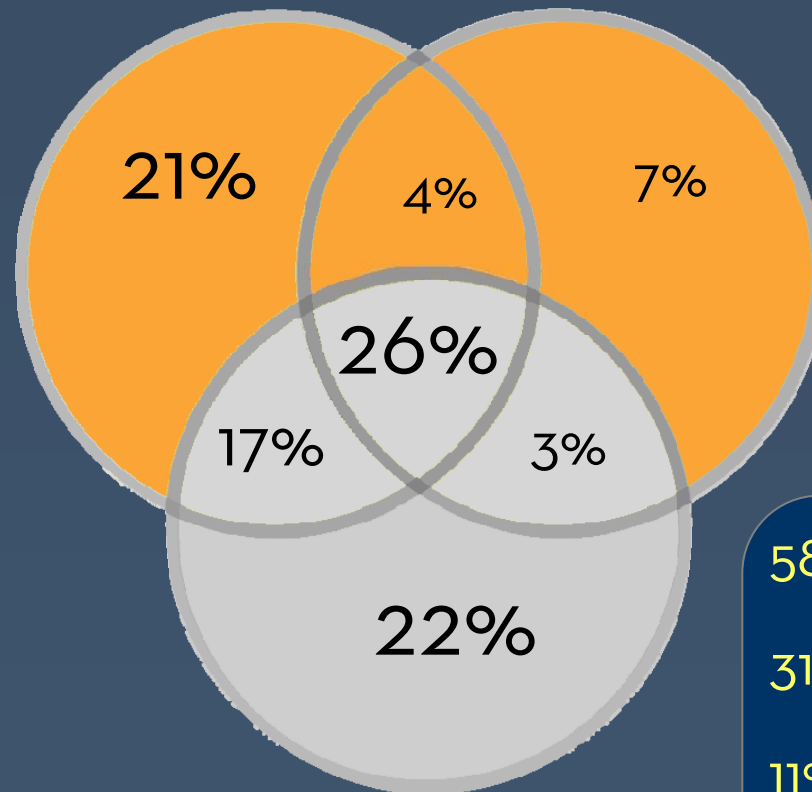
Unexpected undo/erase

Detection rate not
considered

Problems missed by self-report

undo

erase



SEVERITY HISTOGRAM

58% blamed self

31% too minor to report

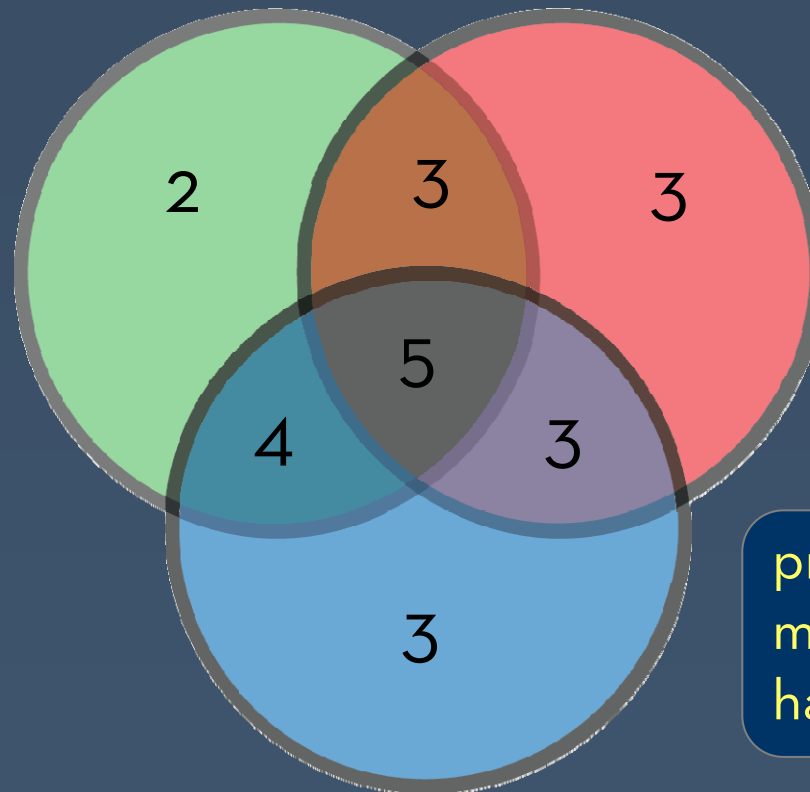
11% forgot to report

→ attribution theory?

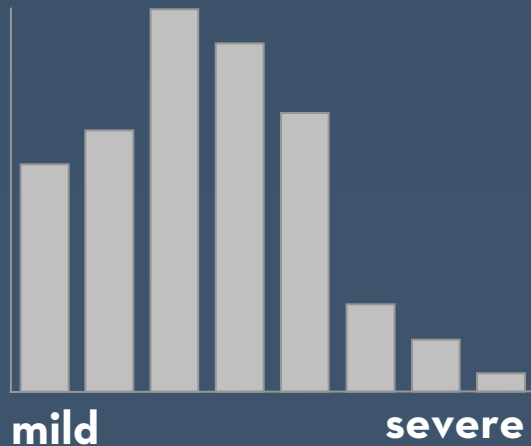
Median severity of problems

undo

erase



problems detected by more than one method have higher severity.



SEVERITY HISTOGRAM

Outline

Motivation and Research Question

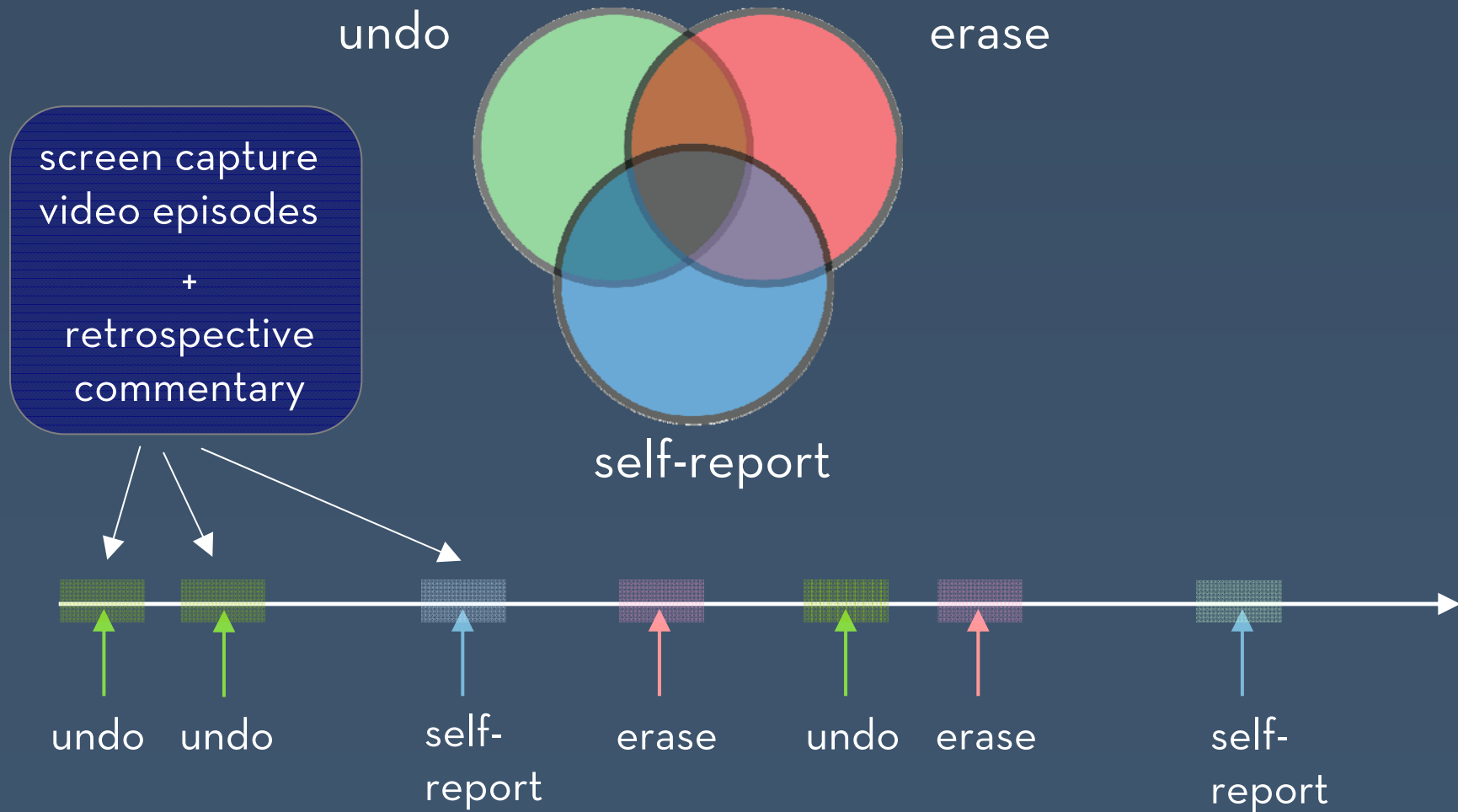
Google SketchUp

Design of the Experiment

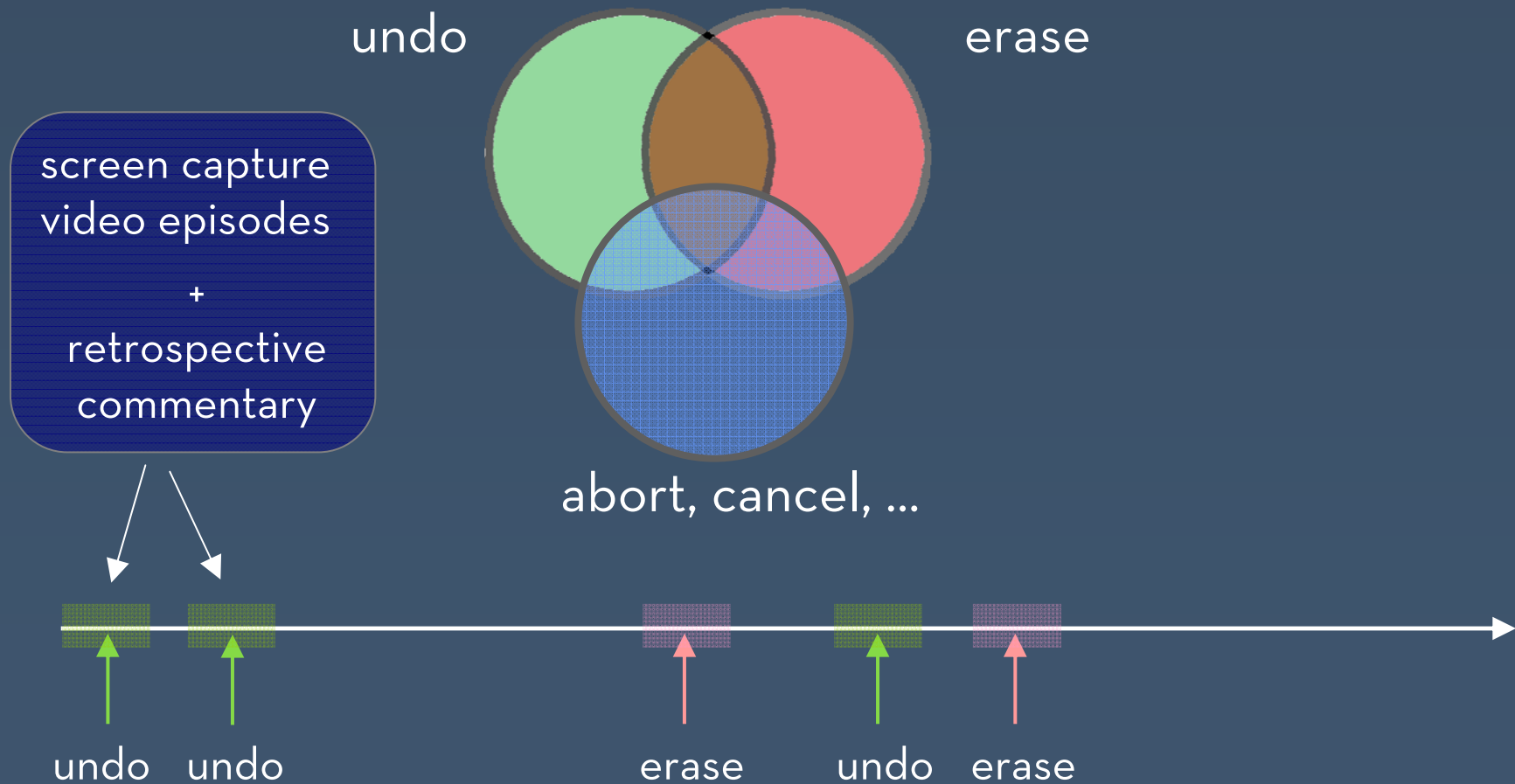
Results from the Experiment

Reflections and Future Work

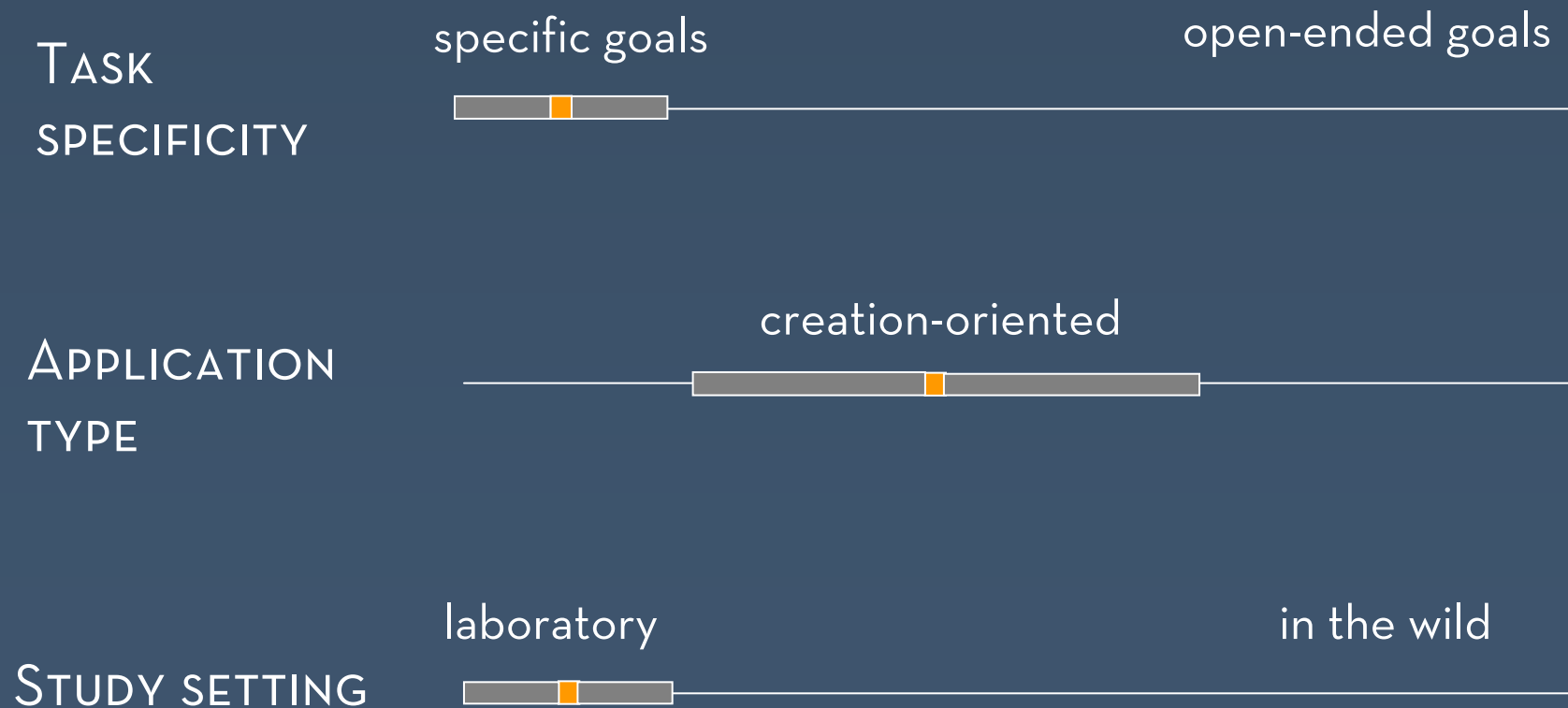
Experimental design



Backtracking Analysis



Extending Backtracking Analysis

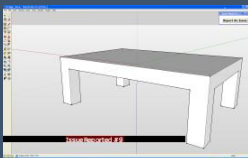
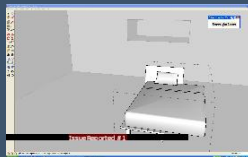
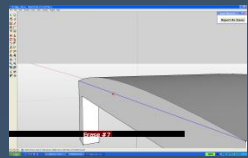
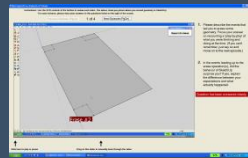


The evaluator effect

Episodes (353)

Problem Instances (215)

Problems (95)



Summary



Undo and erase events are useful indicators of usability problems in Google SketchUp, compared to self-reporting.

We introduce a new usability evaluation method, backtracking analysis.

Backtracking analysis is not meant to replace traditional usability testing.

Acknowledgments

Thanks to Bryce Stout, Charles Akers, and Xin Fu for their generous feedback on the study design and the paper.

Thanks to Google for funding this research.

David Akers
dakers@stanford.edu