On the Effectiveness of Screen Mockups in Enhancing Use Cases

Results from empirical investigations

Marco Torchiano

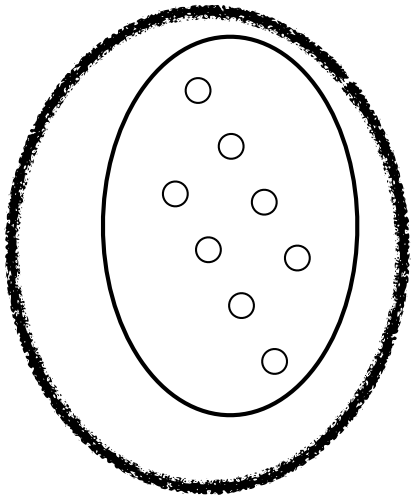
Empirical Studies

- Application of scientific method to SE
  - Define hypotheses
  - Perform experiment to test them
  - If experiment contradict, reject
  - Otherwise keep
- As more and more evidence accumulates, the hypotheses becomes a scientific theory

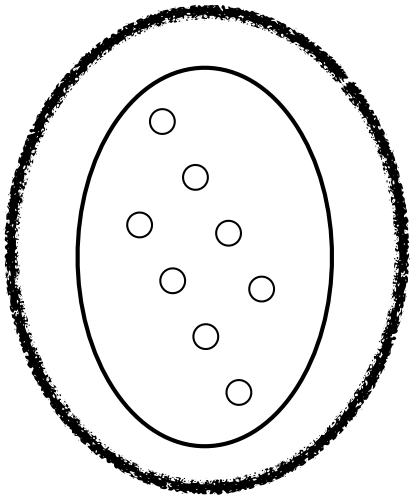
Types of studies

- Formal experiment
  - Max control, high cost
- Case study
  - Low control, low cost
- Survey
  - Low control, medium cost

Sampling



A	B
155	169
170	173
175	176
176	178
177	179
178	181
180	182
182	193



Mean:	174.12	178.87

Hypotheses on samples

- Null hypothesis
  - Both samples come from the same population
- Alternative hypothesis
  - The samples come from different populations

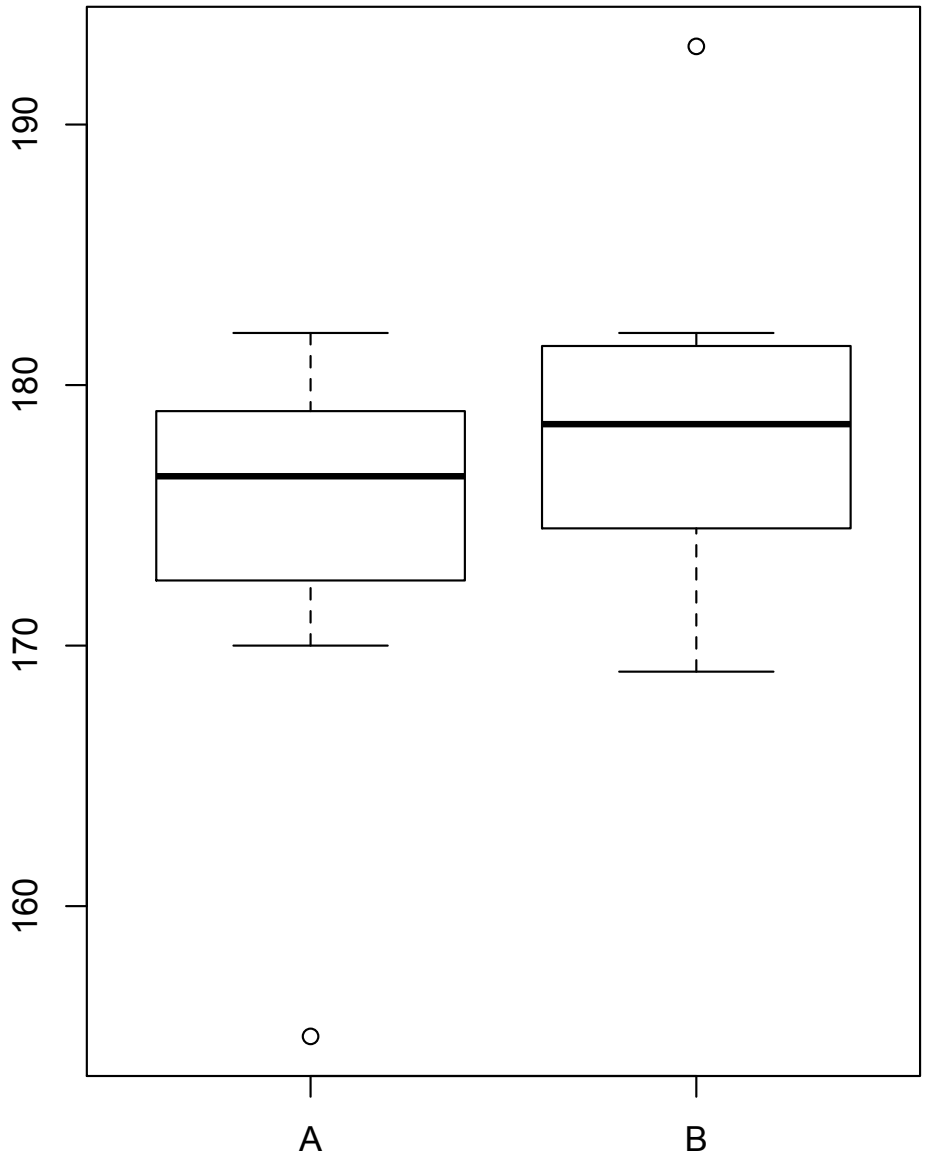
Hypothesis testing

- Assuming the null hypothesis:
  - given two samples of size N, which is the probability of finding a difference greater or equal to the observed one?

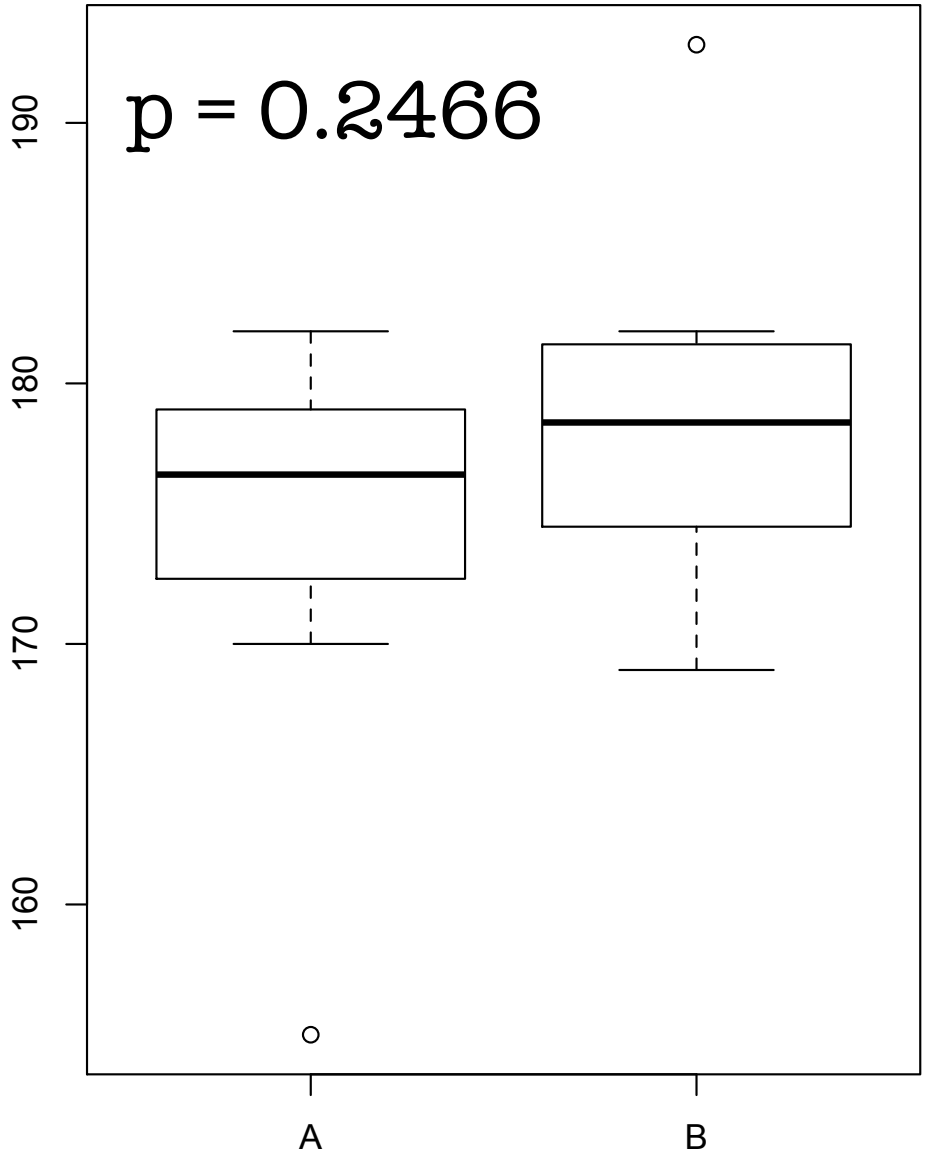
Confidence

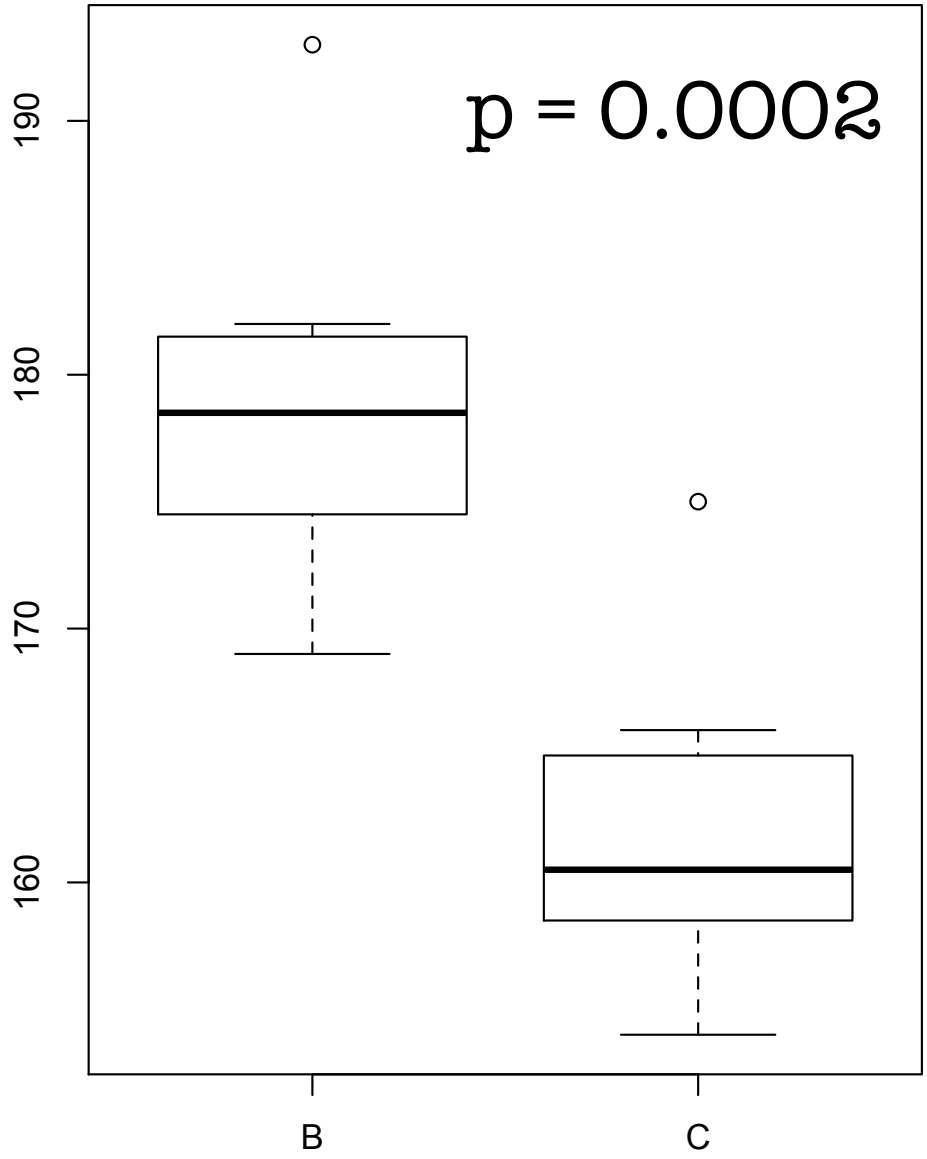
- Type I error
  - Probability of a true null hypothesis when it is true
  - Confidence level α (5%)
- Statistic tests results
  - p-value $\leq \alpha$?
- In our example: t-test p = 0.2466

Boxplot



Boxplot





The study

Use Cases

- Simple way to capture in textual form and define requirements from the end user point of view
- Define goal-oriented sets of interactions between external actors and the system

USE CASE: Insert Coin

Level: User-Goal

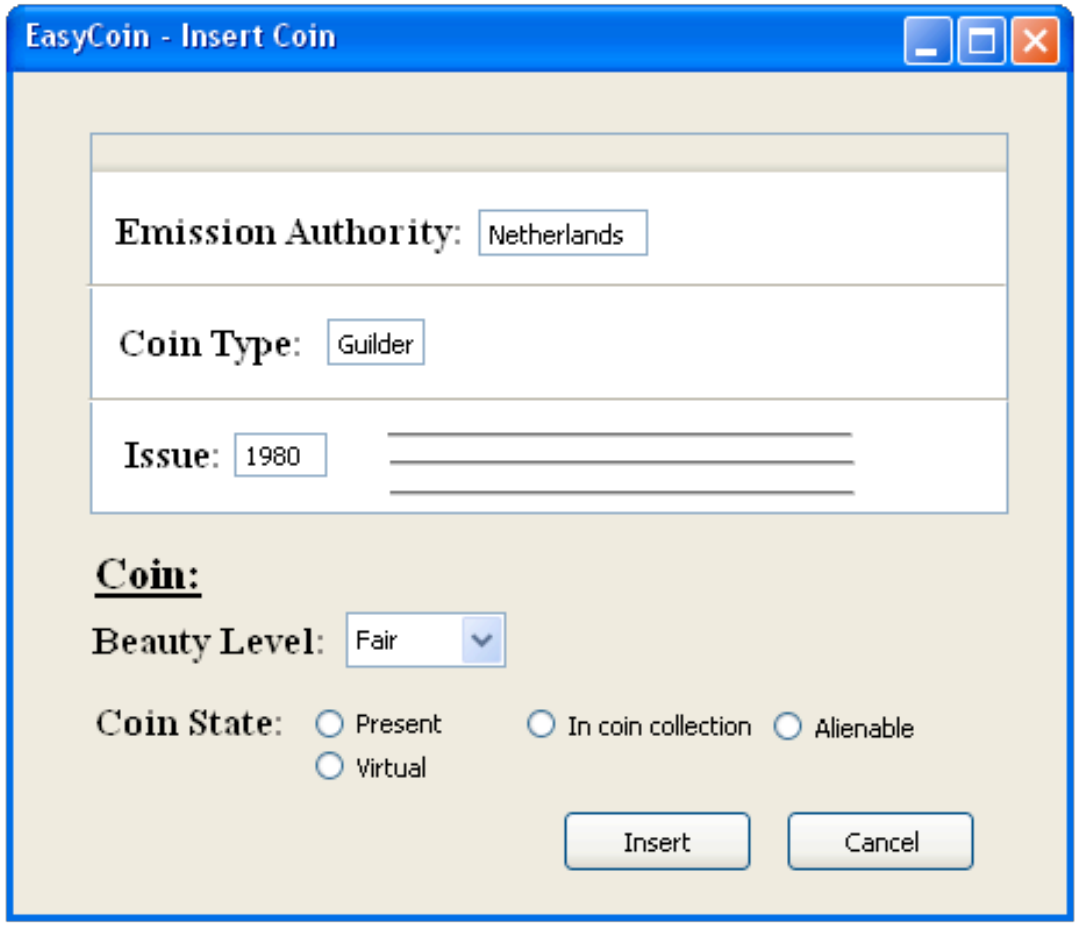
Intention in context: the collector wants to insert a coin in the collection

Primary actor: coin collector

Precondition: a non-void list of issues is selected

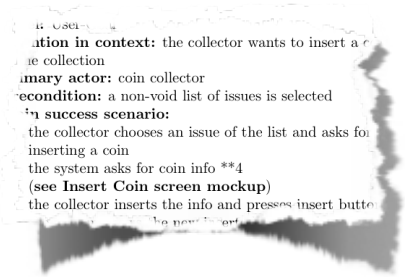
Main success scenario:

1. the collector chooses an issue of the list and asks for inserting a coin
2. the system asks for coin info * *
   (see Insert Coin screen mockup)
3. the collector inserts the info and presses insert button
4. the system shows the new inserted coin to the collector and the Use Case ends with success

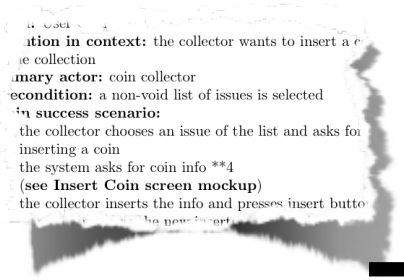


Screen Mockup

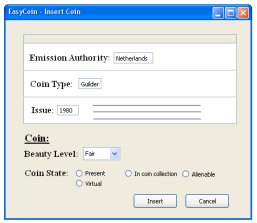
Research question



vs



+



Do screen mockups provide a more effective way to increase the comprehension of functional requirements w.r.t. use cases alone?

Hypotheses

H_{1O} The presence of screen mockups in Use Cases does not significantly improve the comprehension level of software requirements.

H_{eO} The presence of screen mockups in Use Cases does not significantly reduce the effort to comprehend software requirements.

Independent variables

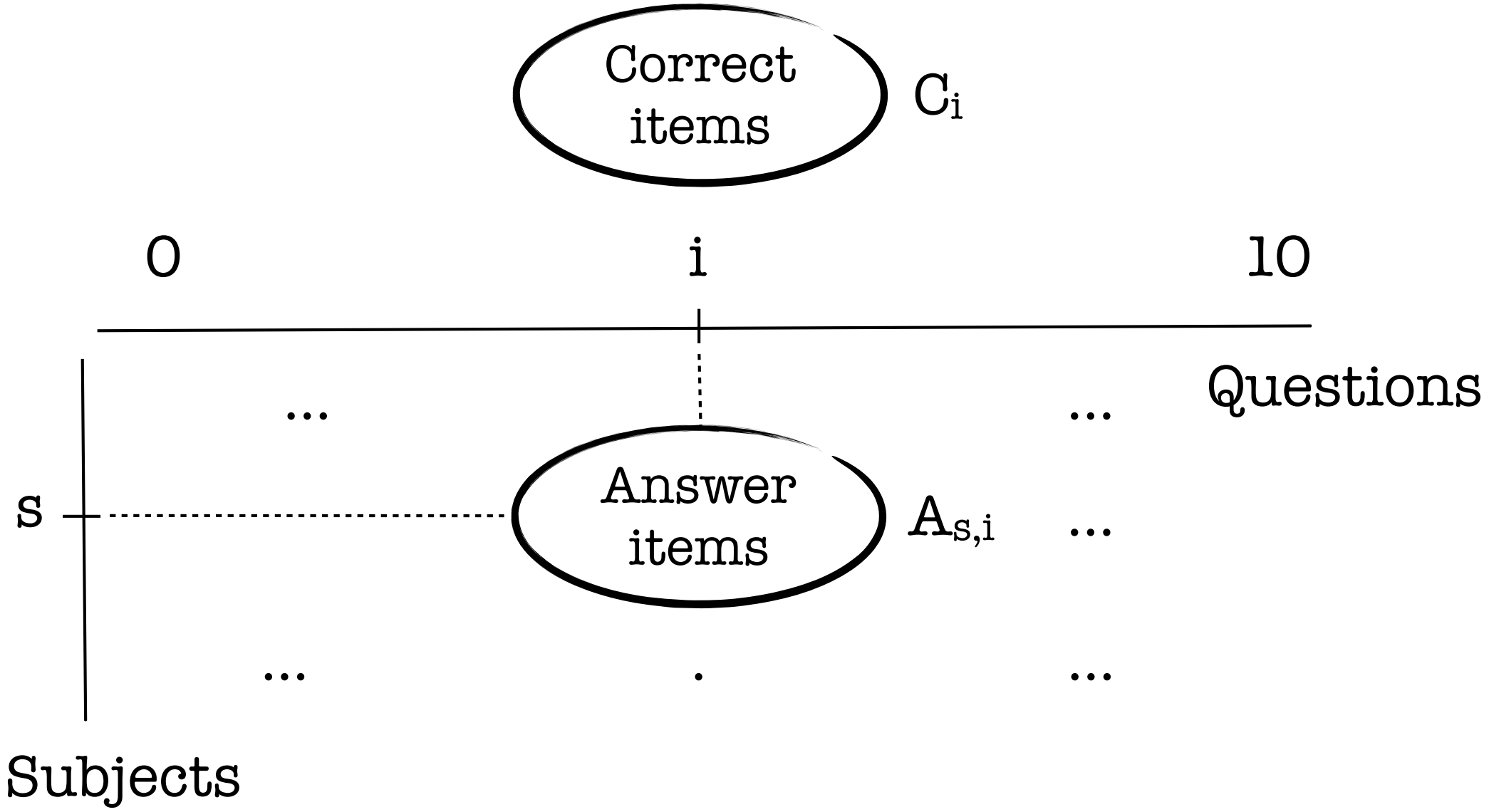
- Treatments
  - T: Purely textual use case
  - S: Above plus screen mockups
- Objects
  - AMICO: condominium mgmt
  - Easy Coin: coin collection catalog

Design

- Counterbalanced

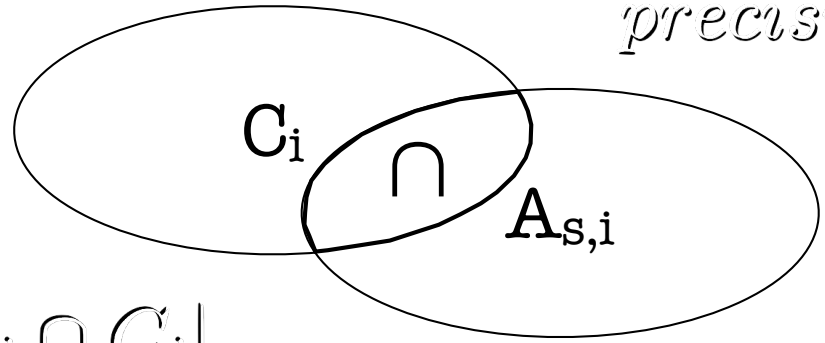
	Group1	Group2	Group3	Group4
Task1	EasyCoin S	EasyCoin T	AMICO T	AMICO S
Task2	AMICO T	AMICO S	EasyCoin S	EasyCoin T

Comprehension



Measures

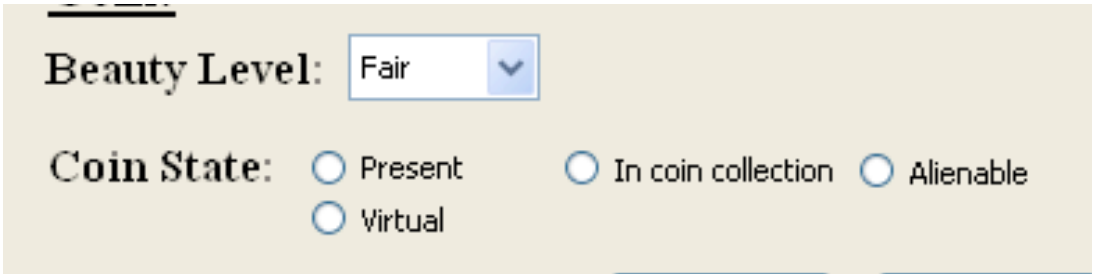
$$precision_{s,i} = \frac{|A_{s,i} \cap C_i|}{|A_{s,i}|}$$



$$recall_{s,i} = \frac{|A_{s,i} \cap C_i|}{|C_i|}$$

$$F-Measure_{s,i} = \frac{2 \cdot precision_{s,i} \cdot recall_{s,i}}{precision_{s,i} + recall_{s,i}}$$

Comprehension Questions



- Example:
  - A coin in EasyCoin contains the following information: beauty level and coin state. Report an example of coin state.

Other factors

Question cognitive area

- Development
- I/O
- Domain

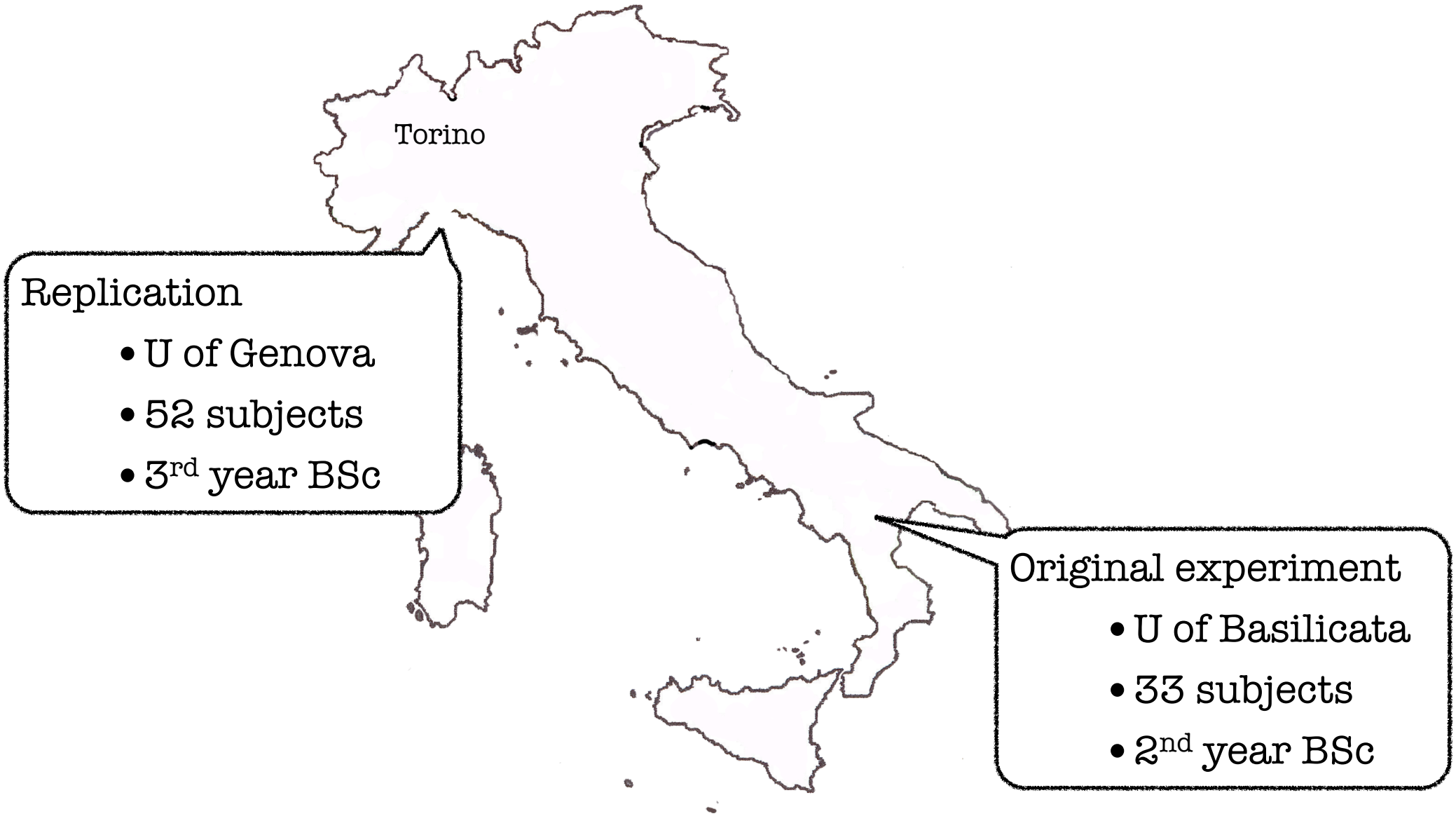
Information sources

- Use Case
- Use Case Diagram
- Screen Mockup
- Previous knowledge / web
- Glossary

Other Hypotheses

H_{SRO} The proportion of questions where screen mockups is the source of information used to answer is equal or lower than the average proportion of information sources.

H_{SPO} The proportion of questions where screen mockups is the source of information used to answer is not the greatest.

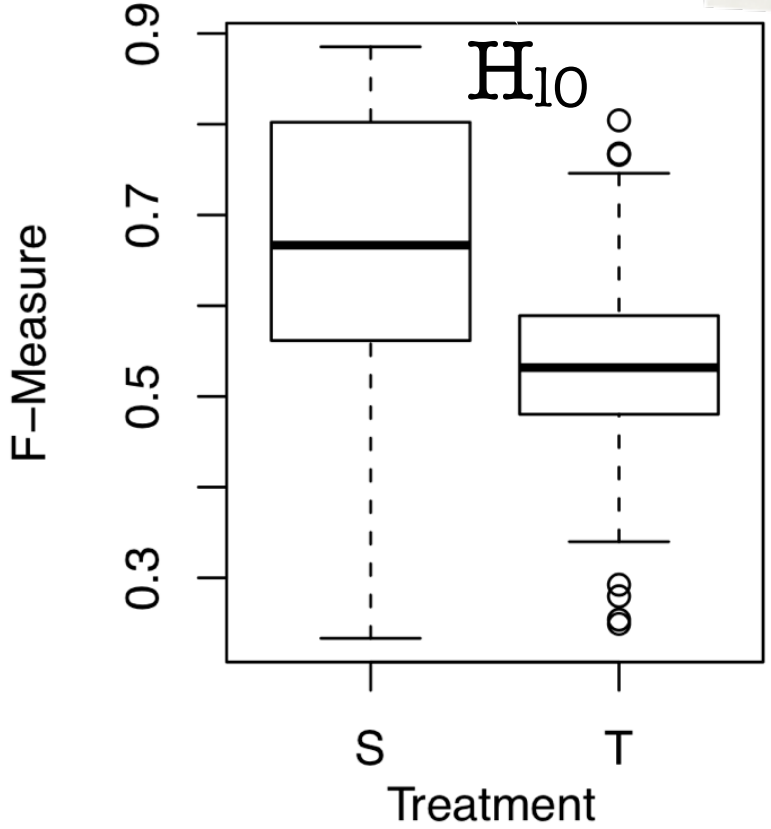


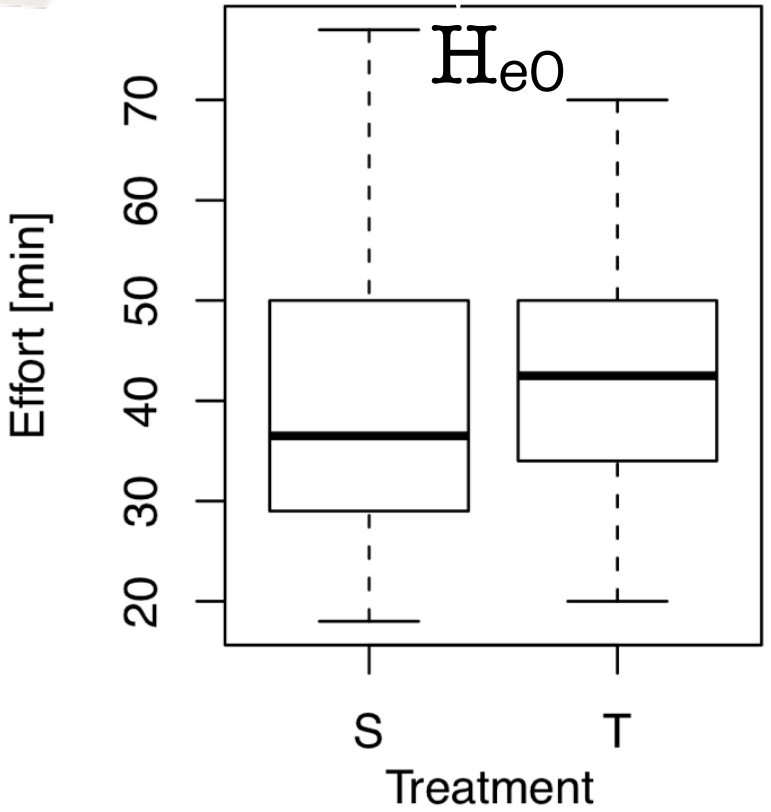
Our Experiments

Replication Results

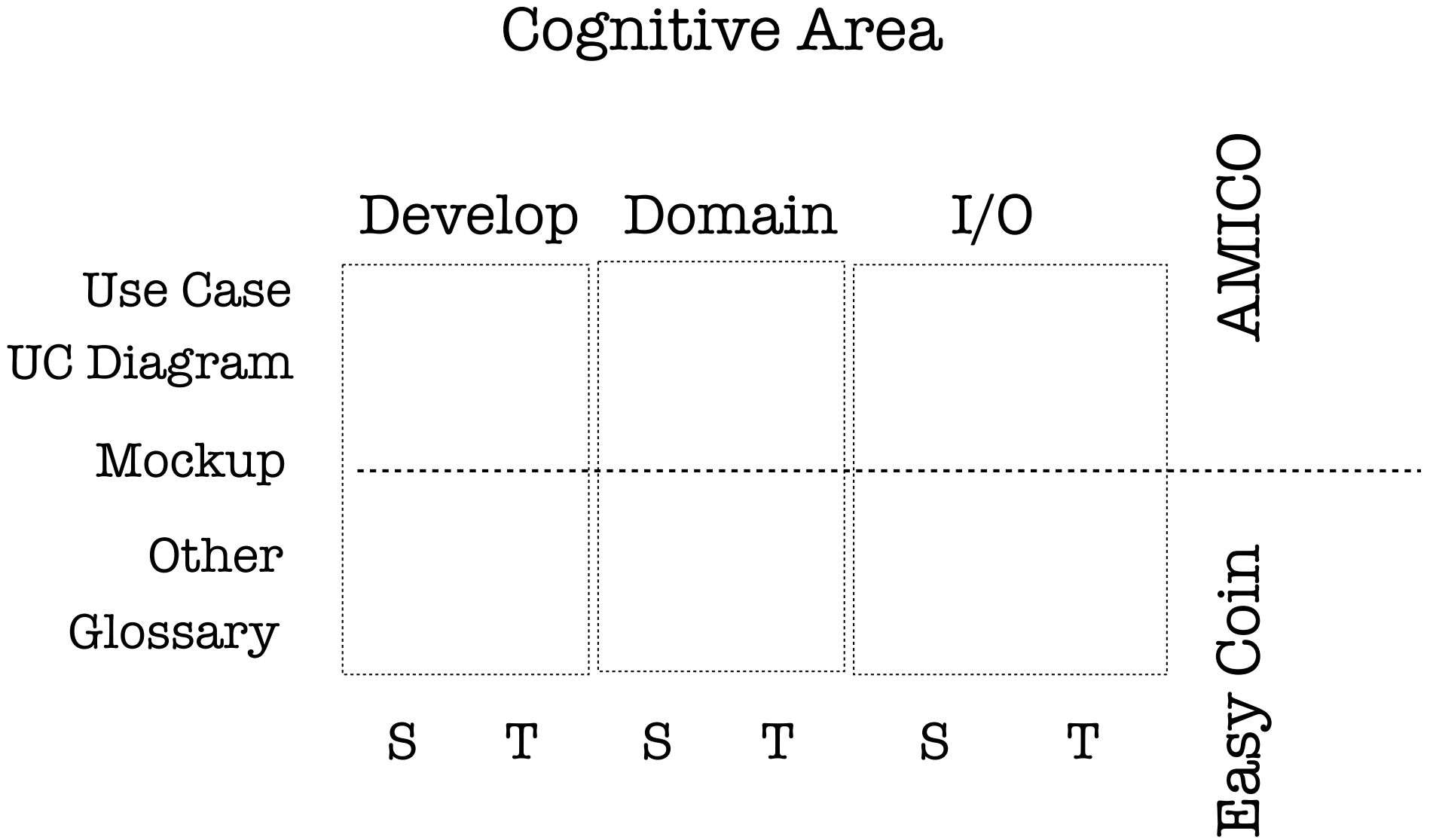
Mann-Whitney
p = 2*10-6

Mann-Whitney
p = 0.1

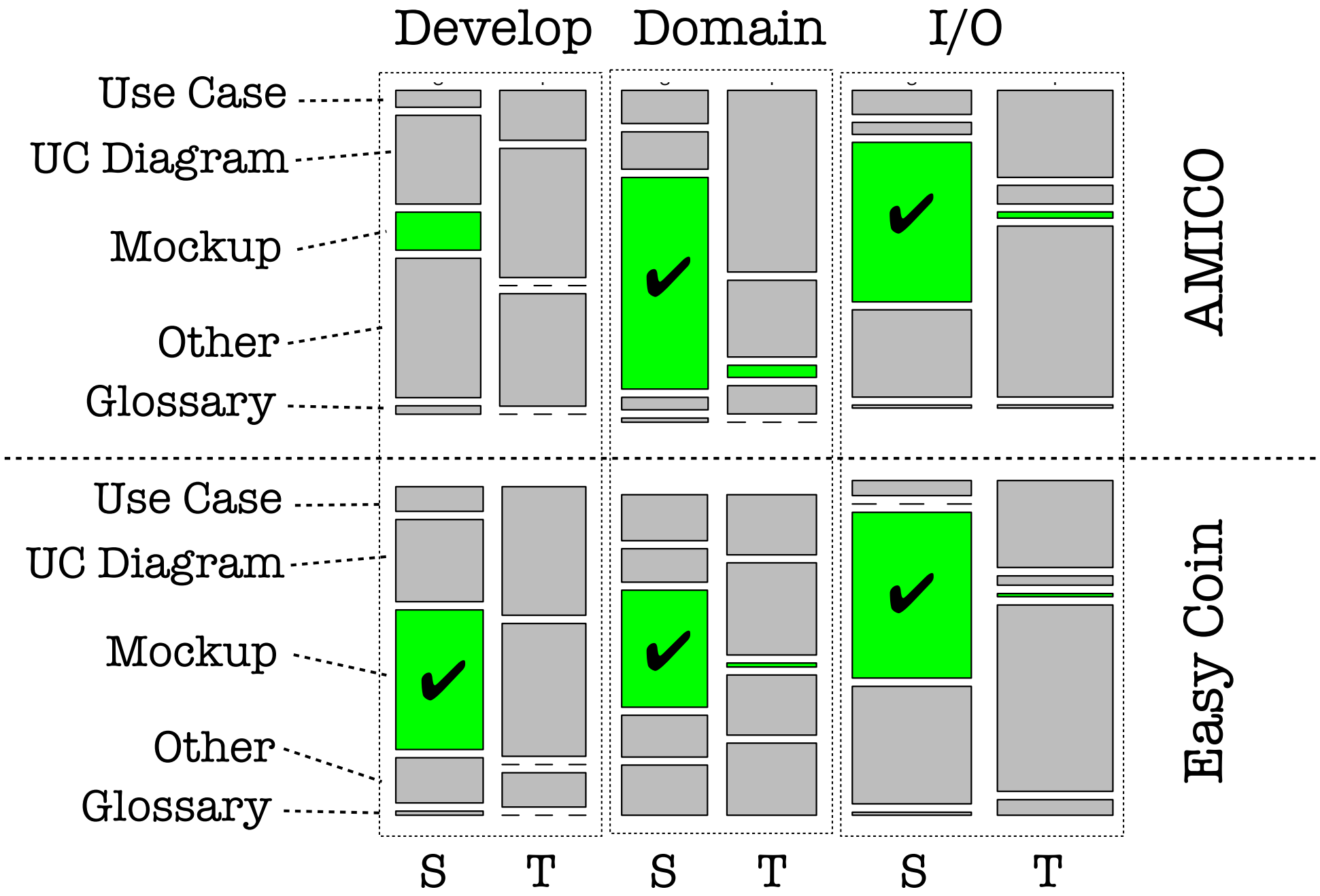




Level & Effort



Information sources



Information sources

Future work

- The effects of changing the domain and the complexity of the tasks
- The influence of the factor “familiarity”
- The influence of Object types
- Will benefits of screen mockups remain consistent across different categories of subjects
  - e.g. graduated students, Ph.D. students, and professional developers

Conclusions

- Significant improvement in comprehension level:
  - +25% (+46% original study)
  - Mockups are a relevant source
- No difference in terms of effort
- Domain familiarity plays a roles
  - increases level improvement
  - decreases mockup relevance as domain-related source of information

Thanks for listening.