POLYTECHNIQUE MONTRÉAL

FONDAZIONE BRUNO KESSLER

Ptidej

Soccerlab
SOftware Cost-effective Change and Evolution Research Lab

An Empirical Study on the Efficiency of Graphical vs. Textual Representations in Requirements Comprehension

Zohreh Sharafi, Alessandro Marchetto, Angelo Susi, Giuliano Antoniol and Yann-Gaël Guéhéneuc

ICPC'13

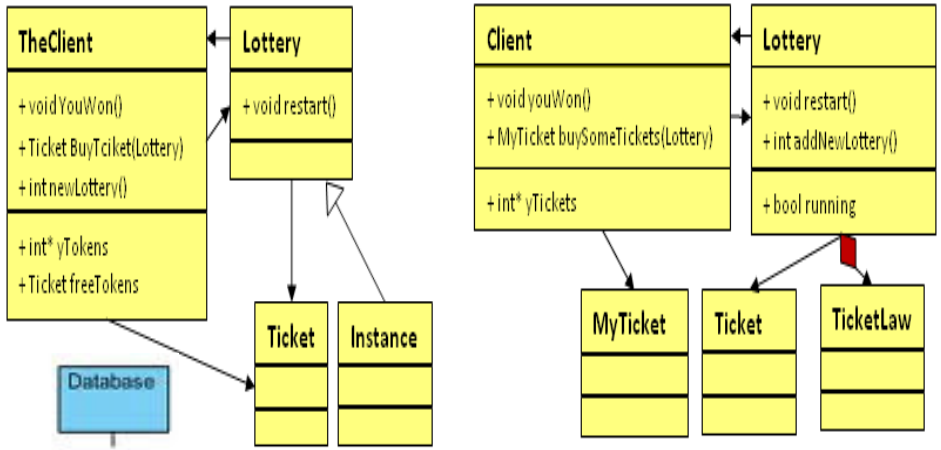
Outline

- Introduction
- Problem and Motivation
- Related Work
- Our Empirical Study
  - Design
  - Result and Analysis
  - Threats to Validity
- Conclusion

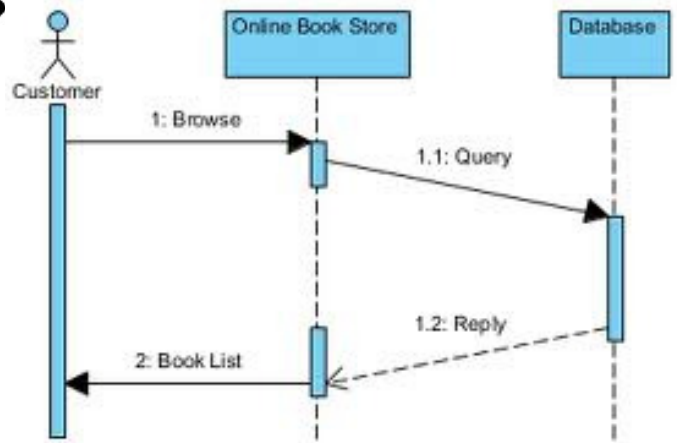
Introduction

Documentation is often only textual or graphical

Graphical Documentation



Comments
Manual pages



Not always available

Requirements

Program comprehension relies heavily on documentation

Introduction

- Is there an impact (accuracy, time) of the kind of document representation on program comprehension?
  - Requirements
    - Structured text only
    - Graphical representation only
    - Mixed textual and graphical representation
- Is there a impact of the mother language, degree of study, or gender?

Requirement Representation

	Requirement	
	Textual	**Graphical (TROPOS)**
Accuracy 	Higher accuracy or no difference?	
Time 	More time spent on textual or more on graphical?	
Effort 	Visual effort ?	

Gender Preference

- Does gender affects the preferred document representation?

- Does gender and representation affect program comprehension?

Related Works

Ottensooser et al. [3]	**Textual representations**	**Significant improvement in understanding of business processes when subjects work with textual representations**
Somervell et al. [12]	**Combination of graphical and textual representations were more efficient**	**Guidelines on the use of a combination of textual and graphical representations to improve subjects' efficiency**
Razali et al. [13]	**Graphical formal specification vs. a purely textual formal specification**	**Combination of semi-formal and formal notations improves the subjects' accuracy**
Heijstek [2] et al.	**Graphical and textual notations for software architecture**	**No difference in accuracy, more experienced subjects mostly preferred a textual representation**

[2] W. Heijstek, T. Kuhne, and M. R. V. Chaudron, "Experimental analysis of textual and graphical representations for software architecture design," in Proceedings of the International Symposium on Empirical Software Engineering and Measurement, IEEE Computer Society, 2011, pp. 167–176.

[3] A. Ottensooser, A. Fekete, H. A. Reijers, J. Mendling, and C. Menictas, "Making sense of business process descriptions: An experimental comparison of graphical and textual notations," Journal of Systems and Software, vol. 85, no. 3, pp. 596–606, March 2012.

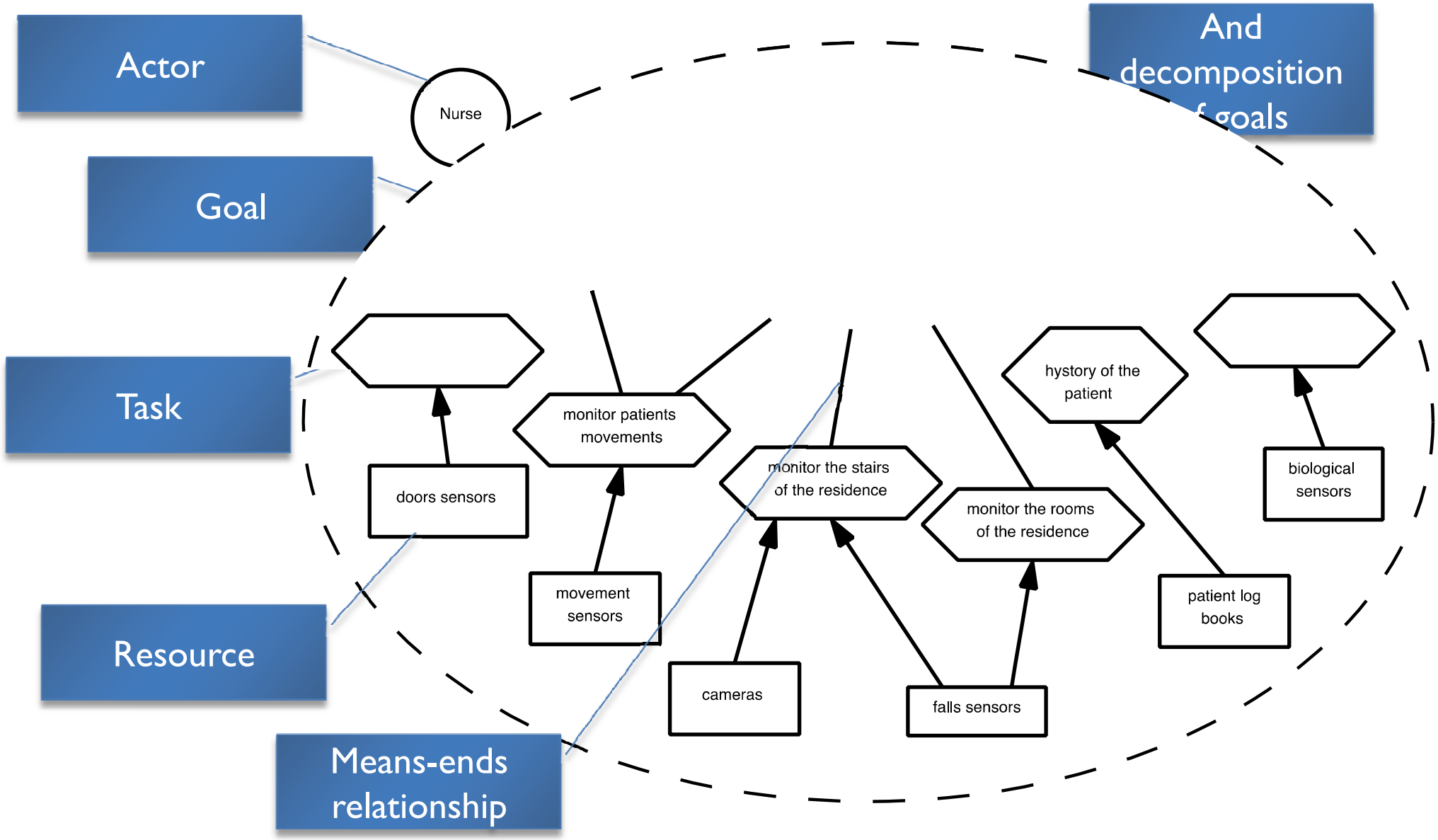
[12] J. Somervell, C. M. Chewar, and D. S. Mccrickard, "Evaluating graphical vs. textual secondary displays for information notification," in Proceedings of the ACM Southeast Conference, ACM press, 2002, pp. 153–160

[13] C. F. Snook and R. Harrison, "Experimental comparison of the comprehensibility of a UML-based formal specification versus a textual one," in Proceedings of 11th International Conference on Evaluation and Assessment in Software Engineering, ACM Press, 2007, pp. 955–971.

TROPOS

- TROPOS is a goal-oriented requirements modeling approach based on concepts such as:
  - *Actor* - typically representing a domain stakeholder
  - *Goal* - representing a state of affairs desired by the actor
  - *Task*- representing set of activities which operationalizes goals
  - *Resource* - which is an element (such as information, device, database, …) whose presence is needed to support the satisfaction of goals or the execution of a task
- And relationships such as:
  - *AND/OR decomposition* of goals and tasks into sub-goals and sub-tasks
  - *Means-ends* to describe the relationship between a goal and the task that fulfill it
- Each concept or relationship has a visual counterpart

TROPOS: goal diagrams



Our Empirical Study

Our Goal: Design and perform an experiment to investigate the impact of requirement representation on comprehension accuracy, time and strategy.

High Level Research Question: Does the document representation impact time or accuracy in program understanding tasks ?

Perspective:

- Developers
- Researcher

Detailed Research Questions

- **RQ1**: Does the type of requirement representations (graphical vs. textual) **impact** the developers' **effort, time, and answer accuracy** in requirements comprehension tasks?
- **RQ2**: Does the structure of the representations lead developers to use **specific task-solving strategies** (top-down vs. bottom-up) during requirements comprehension tasks?
- **RQ3**: Given a graphical and textual representation of a requirements comprehension task, **is there any preferred representation** by the subjects?

Detailed Research Questions

- **RQ2**: Does the structure of the representations lead developers to use specific task-solving strategies (top-down vs. bottom-up) during requirements comprehension tasks?

Detailed Research Questions

- **RQ3**: Given a graphical and textual representation of a requirements comprehension task, is there any preferred representation by the subjects?

Experiment Design

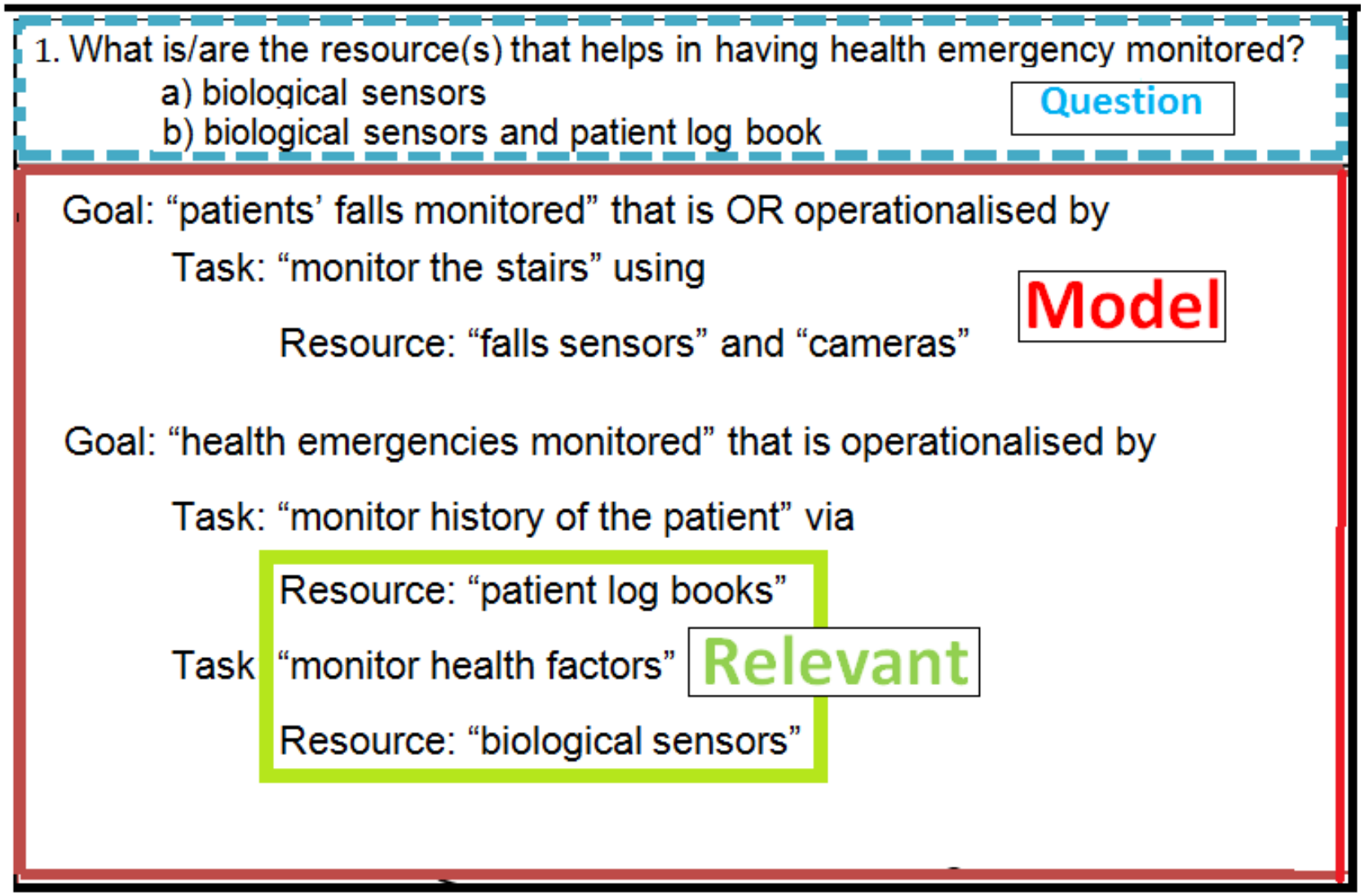
Goal	Study the impact of requirement representation
Independent variables	1. Document representation a) Graphical, b) Textual, c) Both;
Dependent variables	1. Accuracy 2. Required time 3. Effort - Visual Effort
Mitigating variables	1. Study level 2. English language proficiency 3. Mother language 4. Gender: male (M) or female (F)

Subjects' Demography				
Academic background			**Gender**	
Ph.D.	**M.Sc.**	**B.Sc.**	**Male**	**Female**
15	11	2	16	12

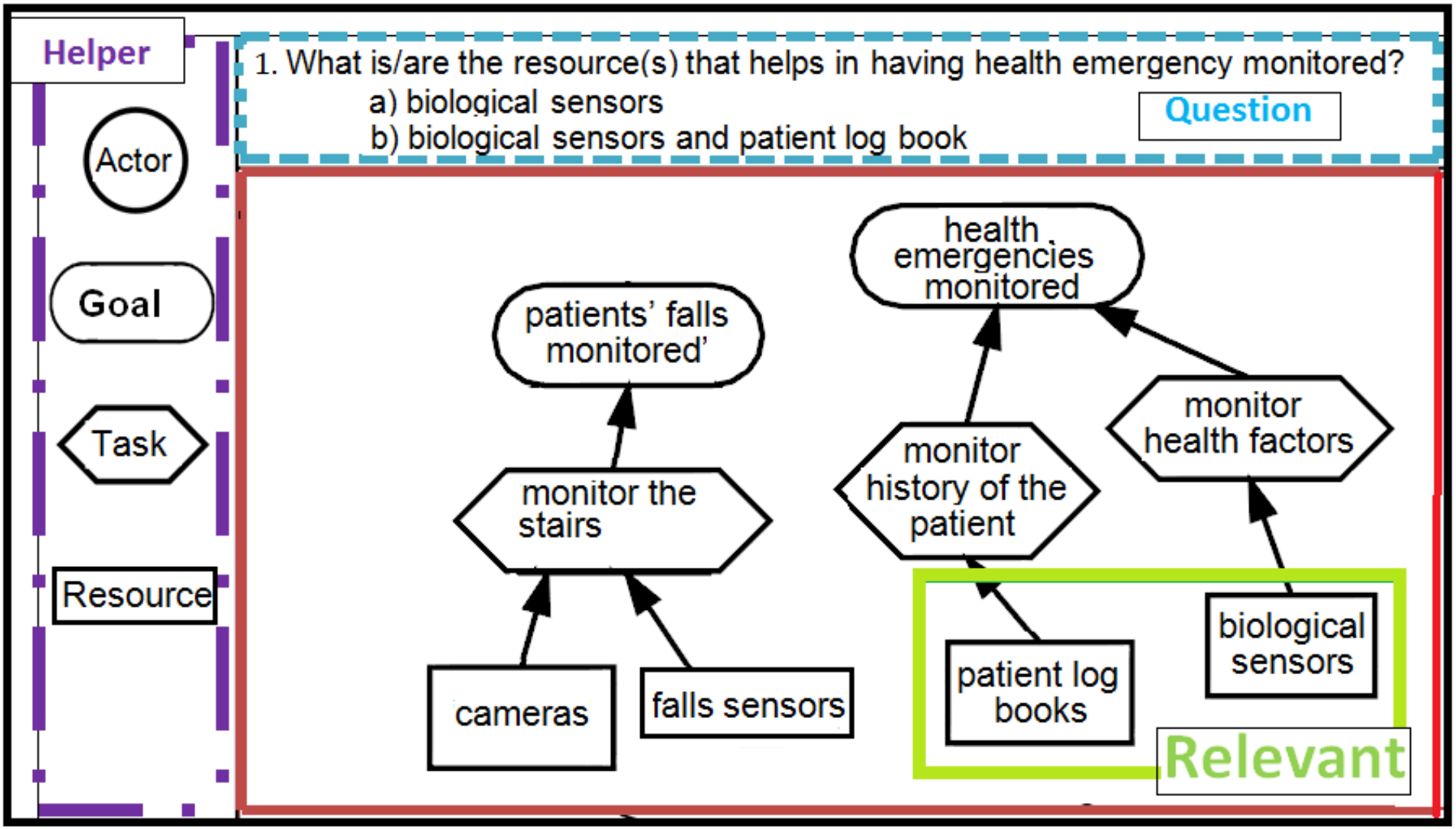
Experiment Design

Subjects' Demography				
Academic background			Gender	
Ph.D.	M.Sc.	B.Sc.	Male	Female
15	11	2	16	12

Experiment Design



Experiment Design



Experiment Design

- FaceLAB
  - Video-based
  - Two camera
  - One infrared
- Non-intrusive
  - No goggles
  - No wires
  - No sensing device

Result and Analysis: Visual Effort

- Visual effort
  - Calculated from eye-tracking data.
  - Calculated based on the amount of visual attention
    - less attention → less time → less effort
- Visual attention triggers the mental processes
- Two types of eye gaze data
  - Fixation
  - Saccade
- We use fixation to calculate effort

Result and Analysis: Visual Effort

- Convex hull: the smallest convex sets of fixations that contains all of a subject's fixations*
  - Measure Average Fixation Duration (AFD) via convex hull as effort proxy
    - Smaller convex hull → close fixations → less effort

* J. H. Goldberg and X. P. Kotval, "Computer interface evaluation using eye movements", 1999.

Result and Analysis: Visual Effort

MODEL

Actor: Nurse

Goal: "emergency detection" is AND decomposed in the:

Goal: "patients escapes controlled" that is OR operationalised by

Task: "monitor the social residence doors" using

Resource: "doors sensors"

Task: "monitor patients movements" using

Resource: "movement sensors"

Goal: "patients falls monitored" that is OR operationalised by

Task: "monitor patients movements" using

Resource: "movement sensors"

Task: "monitor the stairs of the residence" using

Resource: "falls sensors" and "cameras"

Task: "monitor the rooms of the residence" using

Resource: "falls sensors"

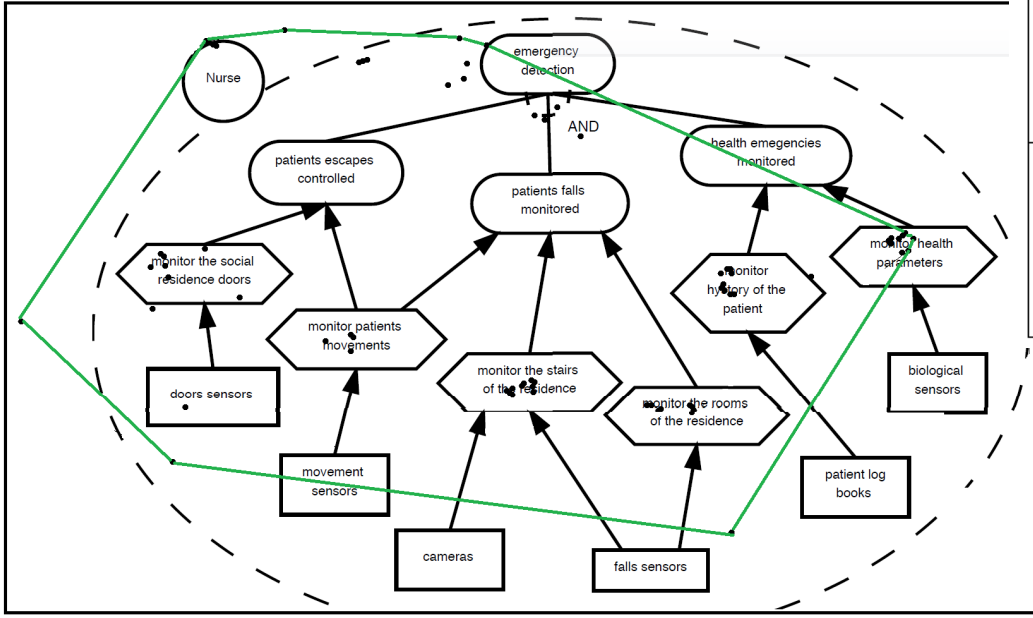
Goal: "health emergencies monitored" that is operationalised by

Task: "monitor history of the patient" via

Resource: "patient log books"

Task: "monitor health parameters"

Resource: "biological sensors"



Numbers of eye fixations

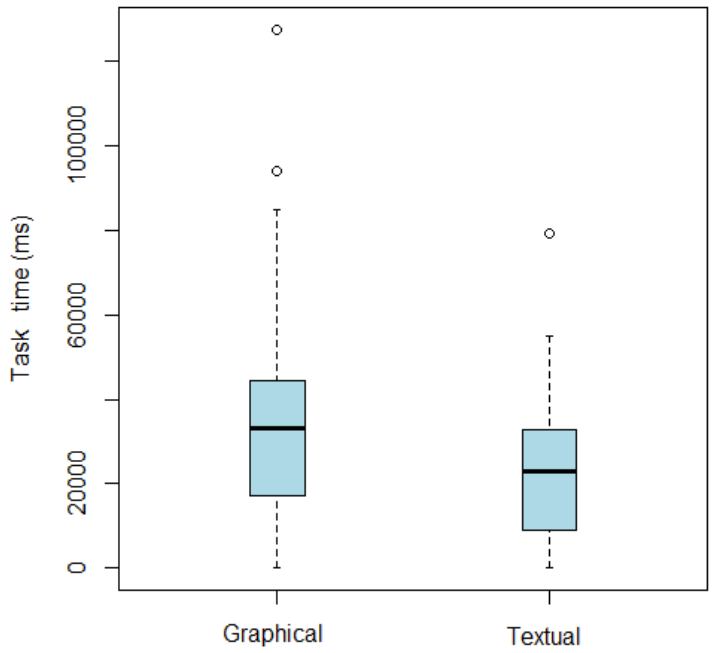
$$FC(Q) = \sum_{a \in \text{ tasks, all answers}} f(a) \quad (1)$$

$$FR(correct) = \frac{\sum_{a \in \text{ correct answer}} f(a)}{\sum_{a \in \text{ correct answer} \cup \text{ distracters}} f(a)} \quad (2)$$

$$FR(distracters) = \frac{\sum_{a \in \text{ distracters}} f(a)}{\sum_{a \in \text{ correct answer} \cup \text{ distracters}} f(a)} \quad (3)$$

Result and Analysis: RQ1

	Accuracy %	
	Correct	**Wrong**
Graphical	97%	3%
Textual	98%	2%
Mixed	96%	4%

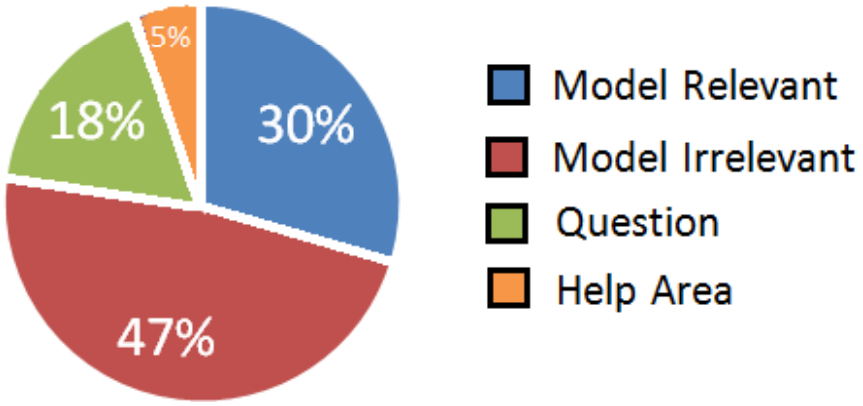


 Accuracy	No significant difference
 Time	There is a significant difference

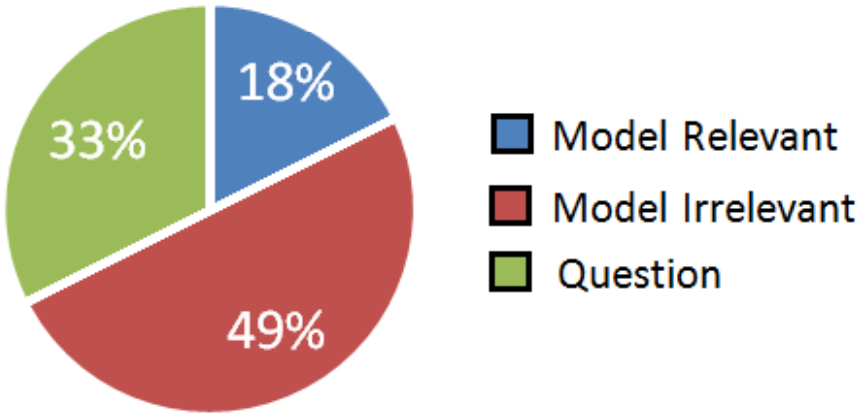
Result and Analysis: RQ1

- Different model imply different areas of focus

The percentage of time that our subjects spent on different AOIs for Graphical model.

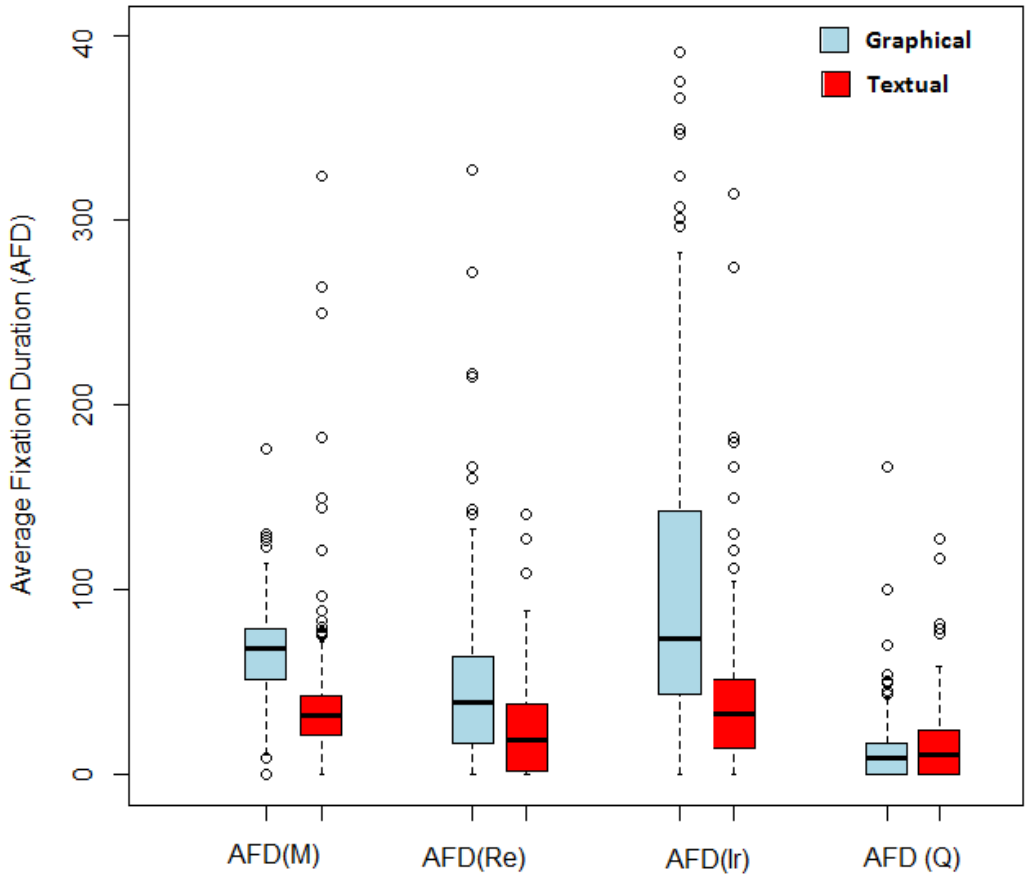


The percentage of time that our subjects spent on different AOIs for Textual model.



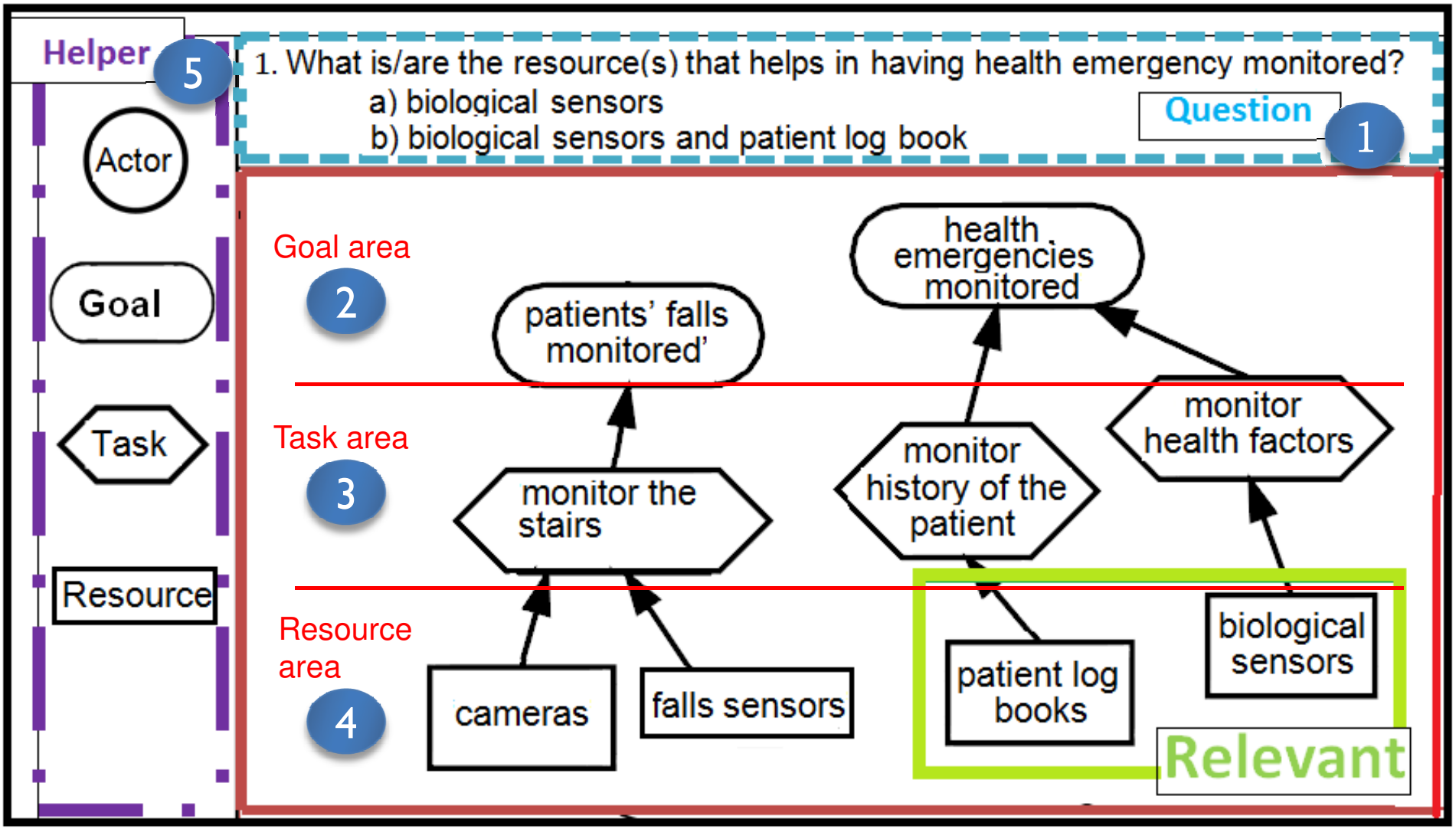
Result and Analysis: RQ1

- There is a significant difference in visual effort though Cohen-d is from medium up
  - AFD(Q) -- Average Fixation Duration -- is borderline 0.07 !



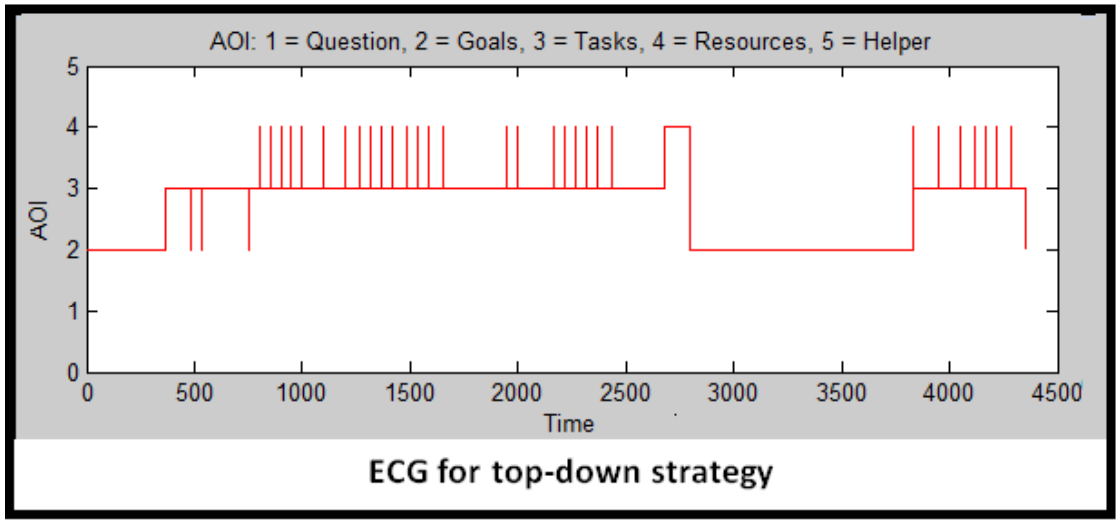
M: Model
Re: Relevant
Ir: Irrelevant
Q: Question

Result and Analysis: RQ2

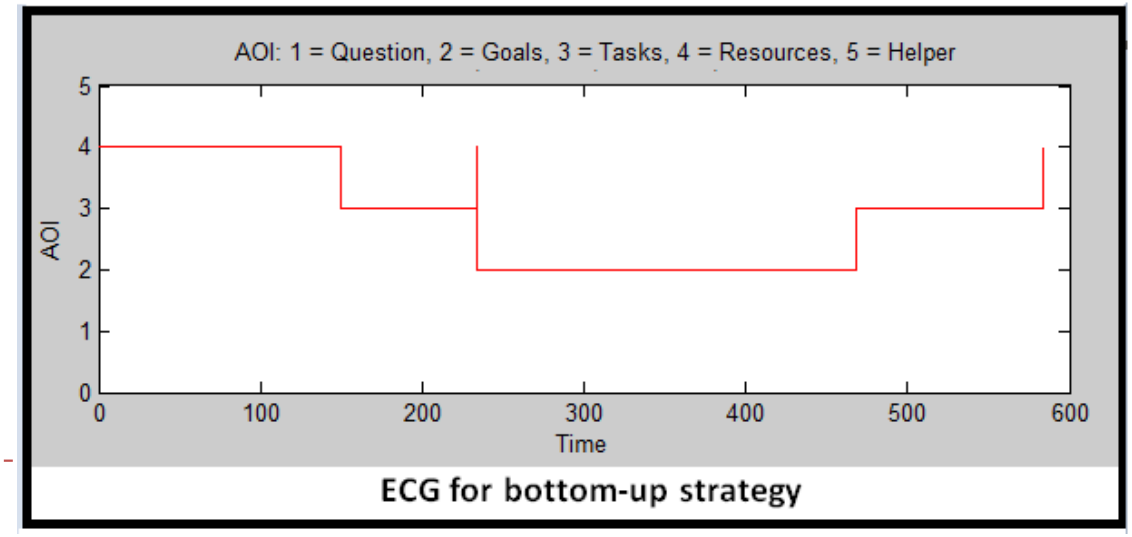


Result and Analysis: RQ2

- The structure of our document makes subject use a top-down (goals to resources) or bottom up (resources to goals) strategy



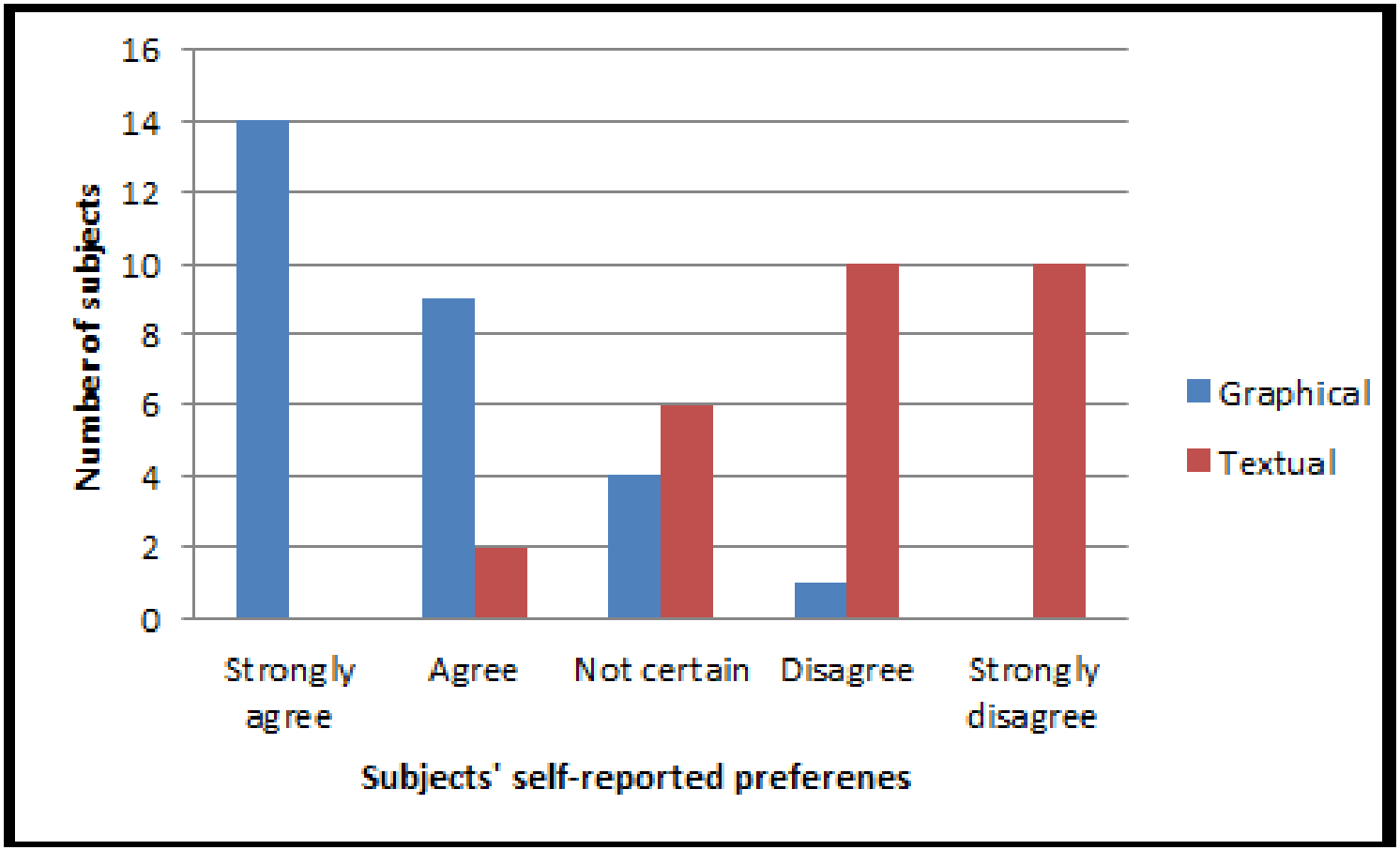
ECG for top-down strategy



ECG for bottom-up strategy

Result and Analysis: RQ3

- Subject Prefer Graphic Notation



Threats to validity

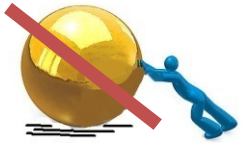
- Internal validity
  - Random ordering of stimuli
  - Provide comfortable environment
- External validity (generalisation of the results)
  - Students as subjects
  - "Only" 28 subjects

Conclusion

- Requirement representation has an impact


Accuracy


Time


Effort


Preference

- Language distance has no impact
  - Closer to English is "better"


Accuracy


Time

- Graphical representation is preferred but requires greater effort
- Gender has not impact

Conclusion

- Design and perform an eye-tracking experiment
- Investigate the impact of document representation on program comprehension
- Examine accuracy, time, effort, and preference