ON THE IMPACT OF SEMANTIC TRANSPARENCY ON UNDERSTANDING AND REVIEWING SOCIAL GOAL MODELS

Mafalda Santos, Catarina Gralha, Miguel Goulão, João Araújo, Ana Moreira

Universidade Nova de Lisboa, Portugal

23 August, 2018

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

NOVALINCS
LABORATORY FOR COMPUTER SCIENCE AND INFORMATICS

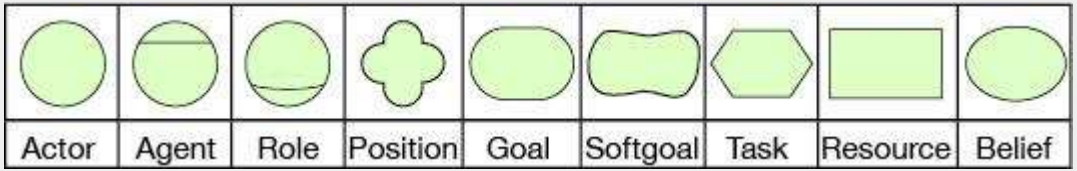
Context and Motivation

- RE success depends on the quality of the communication between requirements engineers and other stakeholders
- Communication flaws are among the most frequently reported RE problems that may lead to project failure
- Visual notations are often adopted, as they are perceived as more effective for conveying information to nontechnical stakeholders than text
- BUT, the visual syntax of SE languages has historically played a secondary role when comparing alternative visual notations

Semantic transparency

- In PoN, Moody proposed a set of principles to support the evaluation, comparison, improvement and construction of visual notations for SE.
- His proposal focused on how to visually represent a set of constructs whose semantics had been previously defined
- A core concept is cognitive effectiveness, defined as the accuracy, speed, and ease with which a representation can be processed by the human mind.
- Semantic transparency, together with the remaining 8 PoN principles, can lead to cognitive effectiveness. It is defined as “the extent to which the meaning of a symbol can be inferred from its appearance”
- Our objective is to compare the ability of stakeholders to understand and review social goal models using two concrete syntaxes

TWO *i** CONCRETE SYNTAXES, WITH DIFFERENT SEMANTIC TRANSPARENCY



Standard *i**
Semantically opaque

New *i**
Symbols with the highest semantic transparency

RESEARCH QUESTIONS

1 Does the adoption of a more semantically transparent concrete syntax improve the **accuracy, speed and ease** when performing **understanding** tasks on *i** SR models?

2 Does the adoption of a more semantically transparent concrete syntax improve the **accuracy, speed and ease** when performing **reviewing** tasks on *i** SR models?

QUASI-EXPERIMENT WITH A COMBINATION OF MEASURES

Direct

Duration
Detection time

Precision
Recall
F-measure

Indirect

Fixations
Saccades
Scanpaths
Heatmaps

Subjective

Performance
Effort
Frustration
Mental demand
Physical demand
Temporal demand

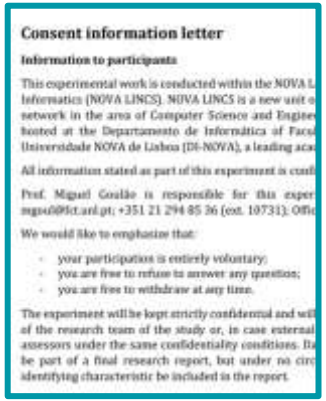
PARTICIPANTS AND EXPERIMENTAL MATERIALS

57 participants

1 eye-tracker

2 domains

READ THE CONSENT LETTER



Consent information letter

Information to participants

This experimental work is conducted within the NOVA L
Informatics (NOVA LINCS). NOVA LINCS is a new unit o
network in the area of Computer Science and Engine
hosted at the Departamento de Informática of Facul
Universidade NOVA de Lisboa (DI-NOVA), a leading aca

All information stated as part of this experiment is confi

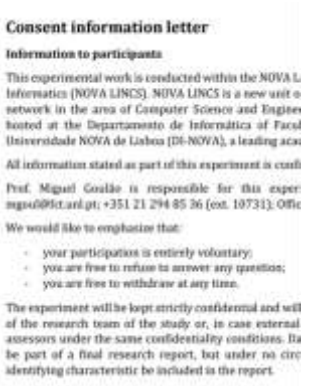
Prof. Miguel Goulão is responsible for this exper
mgoul@fct.unl.pt; +351 21 294 85 36 (ext. 10731); Offic

We would like to emphasize that:

- your participation is entirely voluntary;
- you are free to refuse to answer any question;
- you are free to withdraw at any time.

The experiment will be kept strictly confidential and will
of the research team of the study or, in case external
assessors under the same confidentiality conditions. It
be part of a final research report, but under no circ
identifying characteristic be included in the report.

WATCH A VIDEO TUTORIAL



Consent information letter

Information to participants

This experimental work is conducted within the NOVA L
Informatics (NOVA LINCS). NOVA LINCS is a new unit o
network in the area of Computer Science and Engine
hosted at the Departamento de Informática of Facu
Universidade NOVA de Lisboa (DI-NOVA), a leading aca

All information stated as part of this experiment is confi

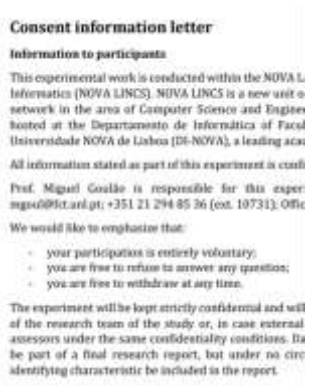
Prof. Miguel Goulão is responsible for this exper
mgoul@fct.unl.pt; +351 21 294 85 36 (ext. 10731); Offic

We would like to emphasize that:

- your participation is entirely voluntary;
- you are free to refuse to answer any question;
- you are free to withdraw at any time.

The experiment will be kept strictly confidential and will
of the research team of the study or, in case external
assessors under the same confidentiality conditions. Da
be part of a final research report, but under no circ
identifying characteristic be included in the report.

CALIBRATE THE EYE-TRACKER



Consent information letter

Information to participants

This experimental work is conducted within the NOVA L
Informatics (NOVA LINCS). NOVA LINCS is a new unit o
network in the area of Computer Science and Engine
hosted at the Departamento de Informática of Facu
Universidade NOVA de Lisboa (DI-NOVA), a leading aca

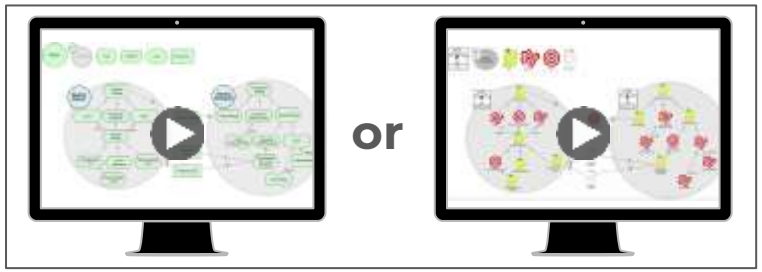
All information stated as part of this experiment is confi

Prof. Miguel Goulão is responsible for this exper
mgoul@fct.unl.pt; +351 21 294 85 36 (ext. 10731); Offic

We would like to emphasize that:

- your participation is entirely voluntary;
- you are free to refuse to answer any question;
- you are free to withdraw at any time.

The experiment will be kept strictly confidential and will
of the research team of the study or, in case external
assessors under the same confidentiality conditions. Da
be part of a final research report, but under no circ
identifying characteristic be included in the report.



PERFORM A TASK

Consent information letter

Information to participants

This experimental work is conducted within the NOVA L
Informatics (NOVA LINCS). NOVA LINCS is a new unit o
network in the area of Computer Science and Engine
hosted at the Departamento de Informática of Facu
Universidade NOVA de Lisboa (DI-NOVA), a leading aca

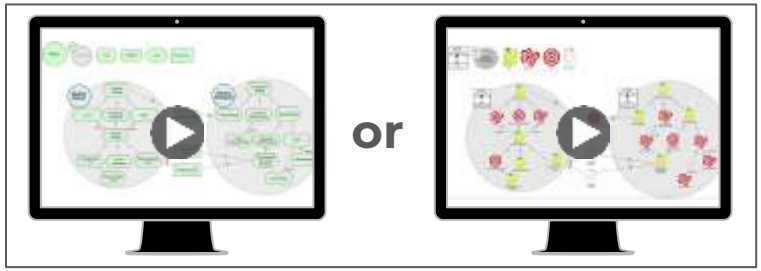
All information stated as part of this experiment is confi

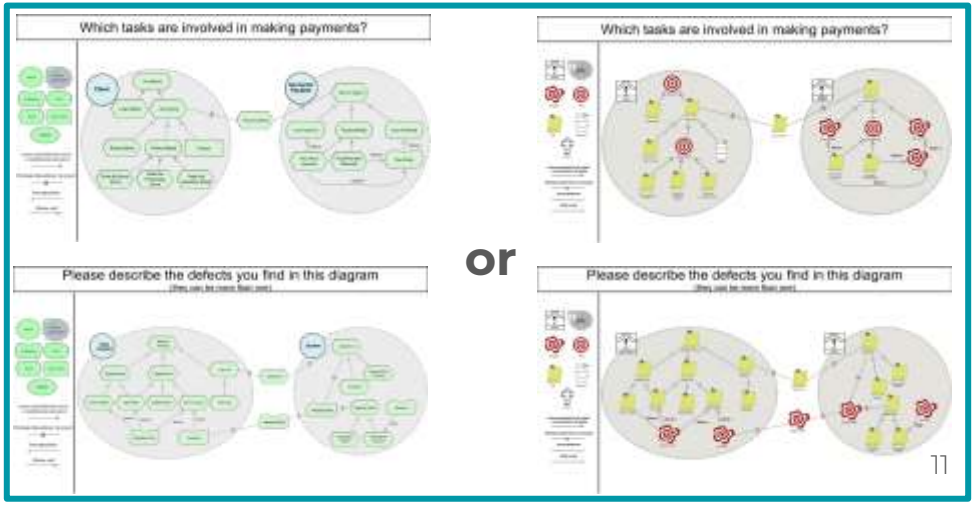
Prof. Miguel Goulão is responsible for this exper
mgoul@fct.unl.pt; +351 21 294 85 36 (ext. 10731); Offic

We would like to emphasize that:

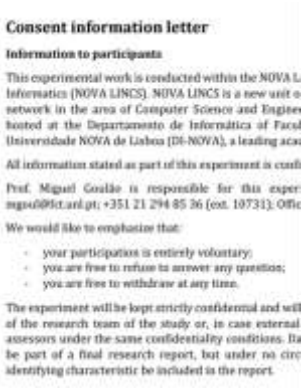
- your participation is entirely voluntary;
- you are free to refuse to answer any question;
- you are free to withdraw at any time.

The experiment will be kept strictly confidential and will
of the research team of the study or, in case external
assessors under the same confidentiality conditions. Da
be part of a final research report, but under no circ
identifying characteristic be included in the report.





ANSWER A NASA-TLX QUESTIONNAIRE


Consent information letter

Information to participants

This experimental work is conducted within the NOVA L
Informatics (NOVA LINCS). NOVA LINCS is a new unit o
network in the area of Computer Science and Engine
hosted at the Departamento de Informática of Facu
Universidade NOVA de Lisboa (DI-NOVA), a leading aca

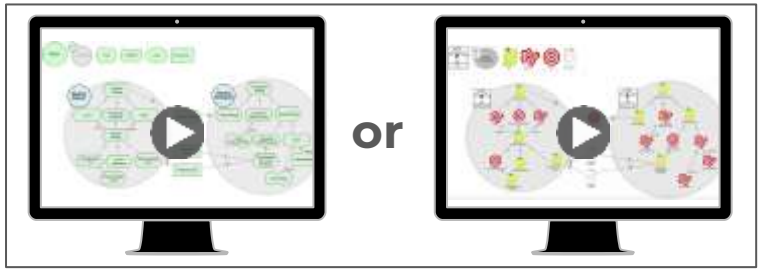
All information stated as part of this experiment is confi

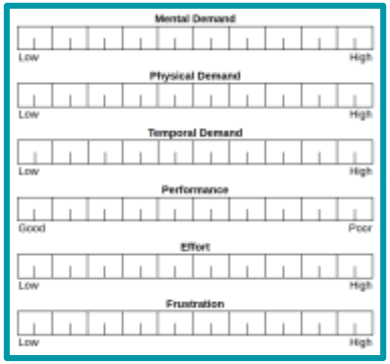
Prof. Miguel Goulão is responsible for this exper
mgoul@fct.unl.pt; +351 21 294 85 36 (ext. 10731); Offic

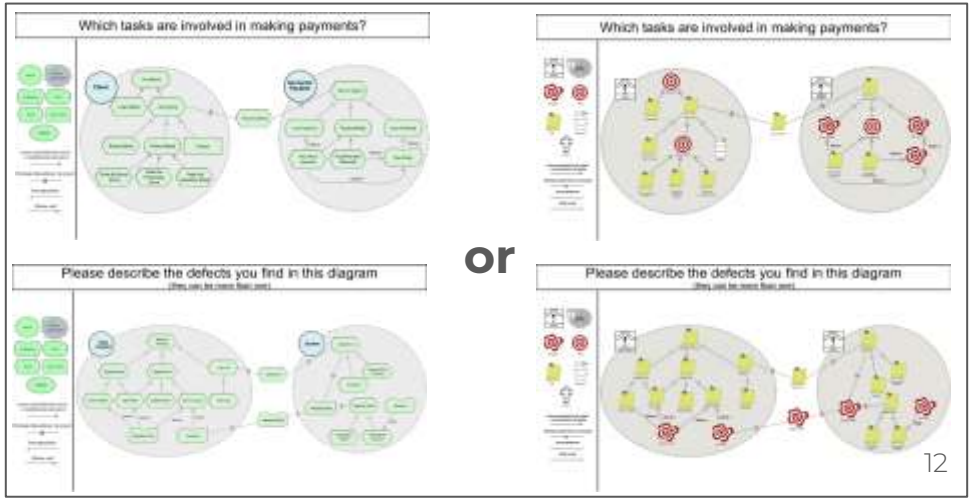
We would like to emphasize that:

- your participation is entirely voluntary;
- you are free to refuse to answer any question;
- you are free to withdraw at any time.

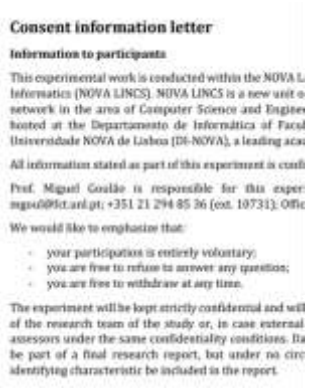
The experiment will be kept strictly confidential and will
of the research team of the study or, in case external
assessors under the same confidentiality conditions. Da
be part of a final research report, but under no circ
identifying characteristic be included in the report.







ANSWER TO DEMOGRAPHIC QUESTIONS


Consent information letter

Information to participants

This experimental work is conducted within the NOVA L
Informatics (NOVA LINCS). NOVA LINCS is a new unit o
network in the area of Computer Science and Engine
hosted at the Departamento de Informática of Facu
Universidade NOVA de Lisboa (DI-NOVA), a leading aca

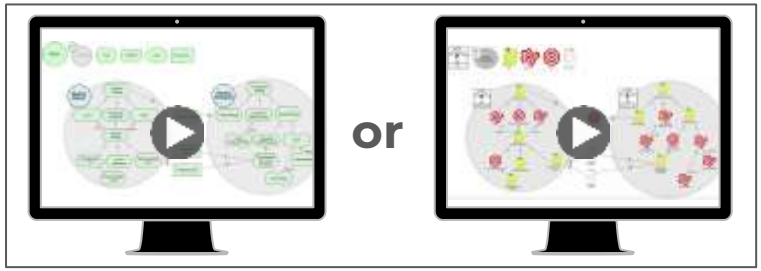
All information stated as part of this experiment is confi

Prof. Miguel Goulão is responsible for this exper
mgoul@fct.unl.pt; +351 21 294 85 36 (ext. 10731); Offic

We would like to emphasize that:

- your participation is entirely voluntary;
- you are free to refuse to answer any question;
- you are free to withdraw at any time.

The experiment will be kept strictly confidential and will
of the research team of the study or, in case external
assessors under the same confidentiality conditions. Da
be part of a final research report, but under no circ
identifying characteristic be included in the report.




Demographic Data

*Obrigatório

Gender *

Male

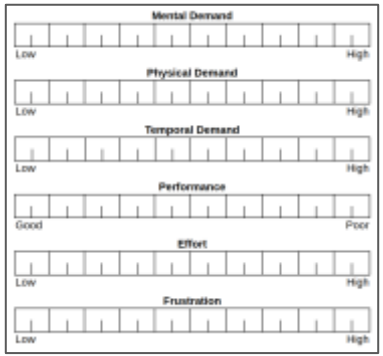
Female

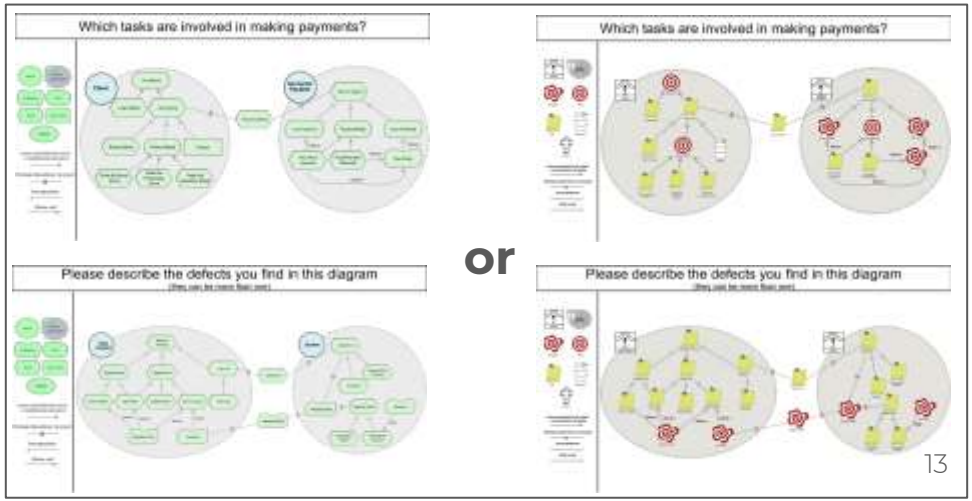
Age *

A sua resposta

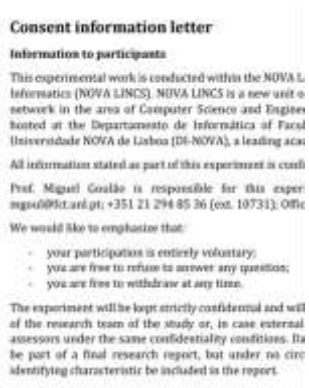
Nationality *

A sua resposta





PROTOCOL OF THE EXPERIMENT


Consent information letter

Information to participants

This experimental work is conducted within the NOVA L
Informatics (NOVA LINCS). NOVA LINCS is a new unit o
network in the area of Computer Science and Engine
hosted at the Departamento de Informática of Facu
Universidade NOVA de Lisboa (DI-NOVA), a leading aca

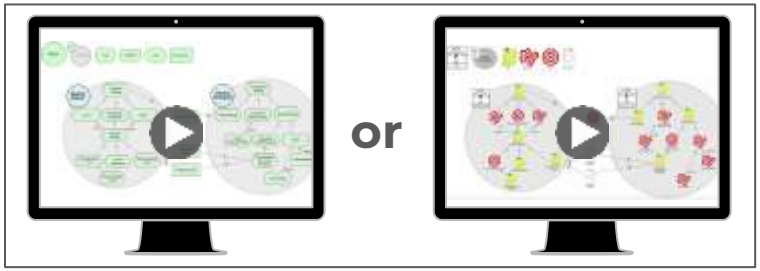
All information stated as part of this experiment is confi

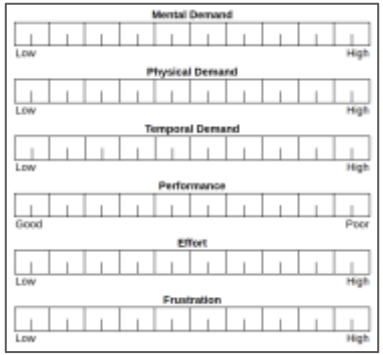
Prof. Miguel Goulão is responsible for this exper
mgoul@fct.unl.pt; +351 21 294 85 36 (ext. 10731); Offic

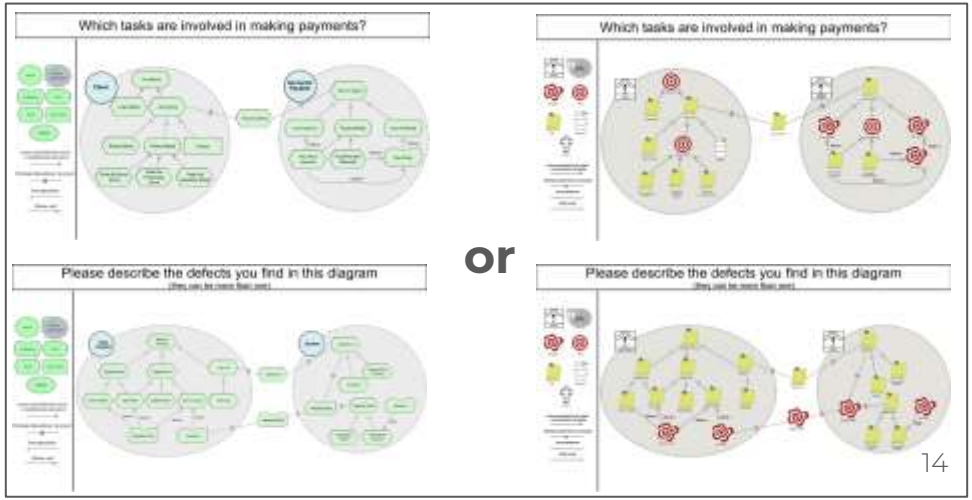
We would like to emphasize that:

- your participation is entirely voluntary;
- you are free to refuse to answer any question;
- you are free to withdraw at any time.

The experiment will be kept strictly confidential and will
of the research team of the study or, in case external
assessors under the same confidentiality conditions. Da
be part of a final research report, but under no circ
identifying characteristic be included in the report.

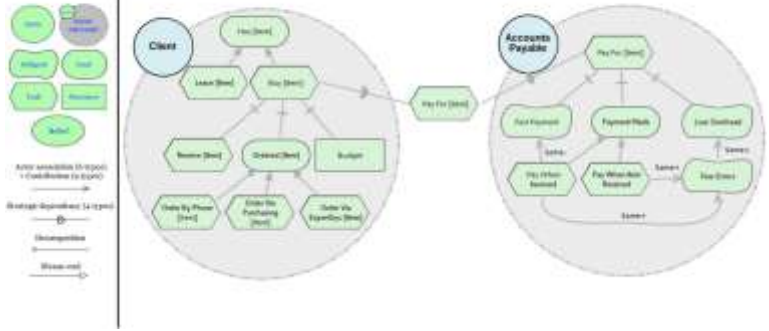




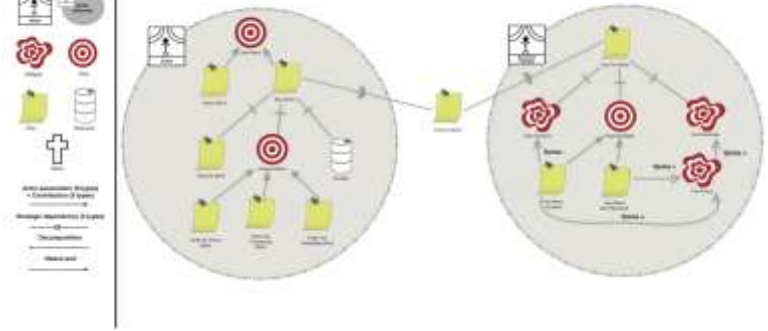


TWO UNDERSTANDING AND TWO REVIEW TASKS, BOTH WITH STANDARD *i** AND NEW *i**

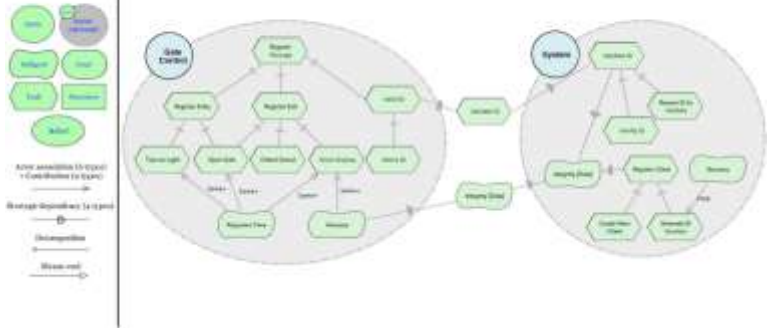
Which tasks are involved in making payments?



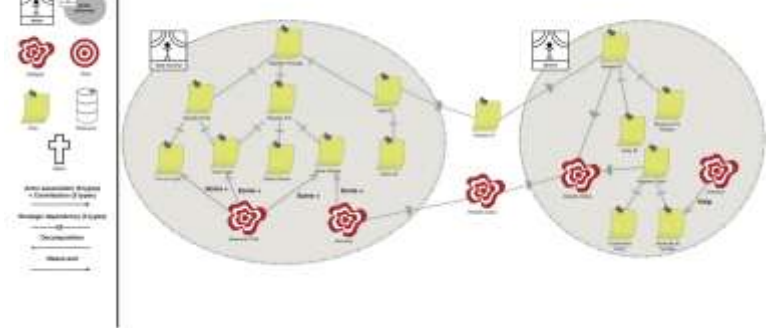
Which tasks are involved in making payments?



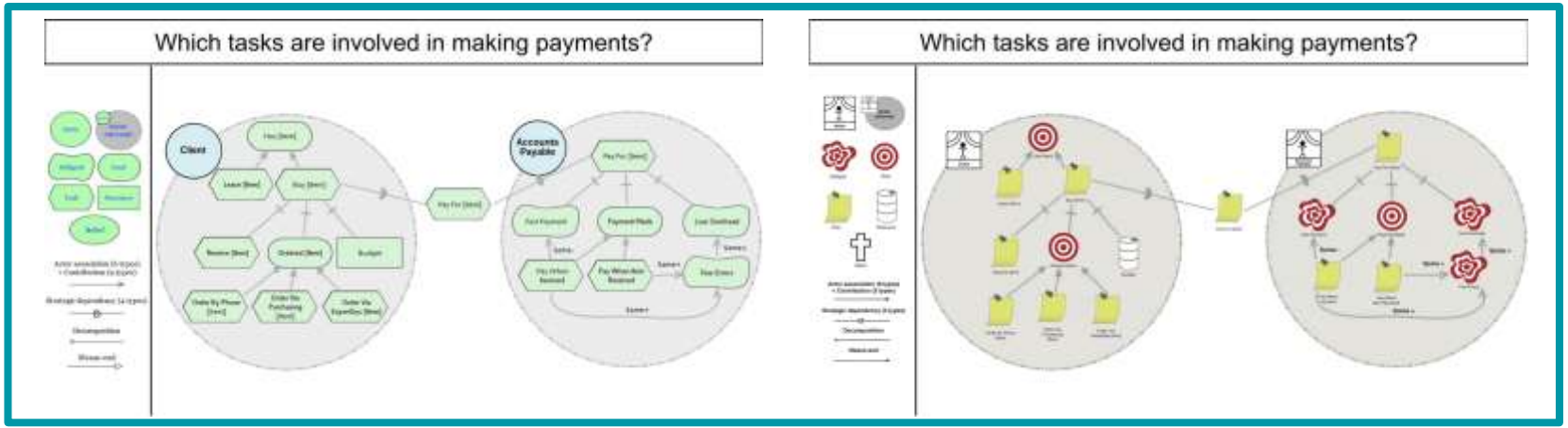
Please describe the defects you find in this diagram
(they can be more than one)

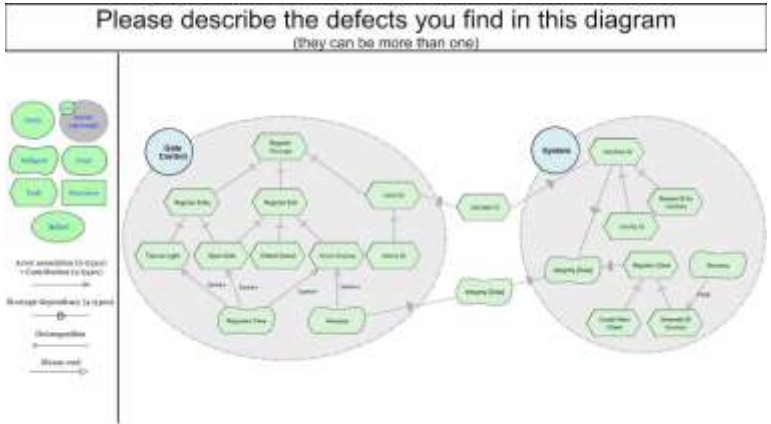


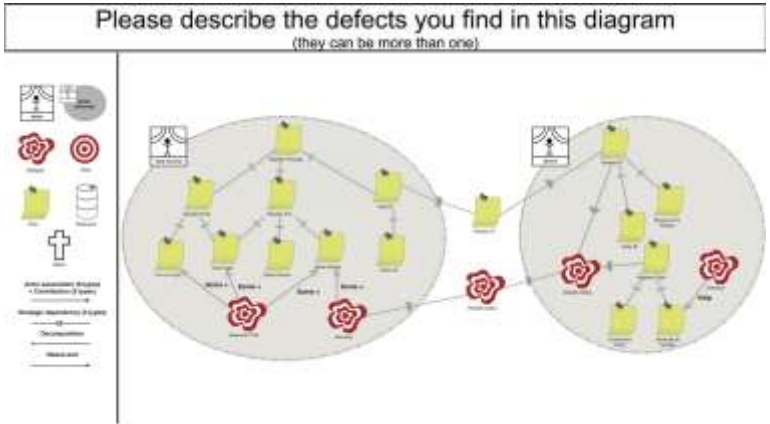
Please describe the defects you find in this diagram
(they can be more than one)



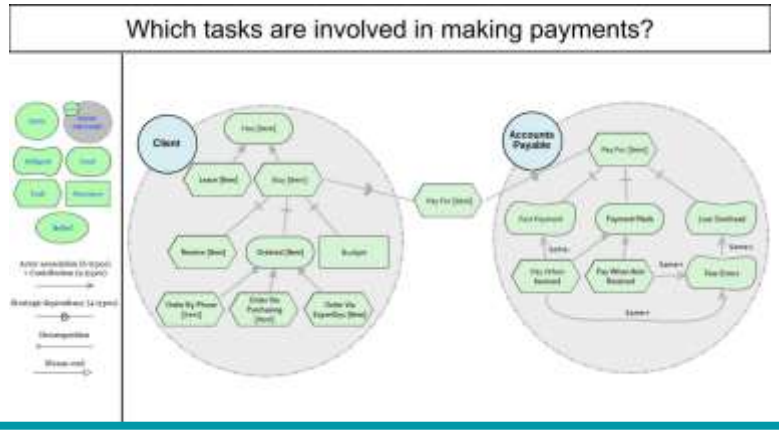
TWO UNDERSTANDING AND TWO REVIEW TASKS, BOTH WITH STANDARD *i** AND NEW *i**

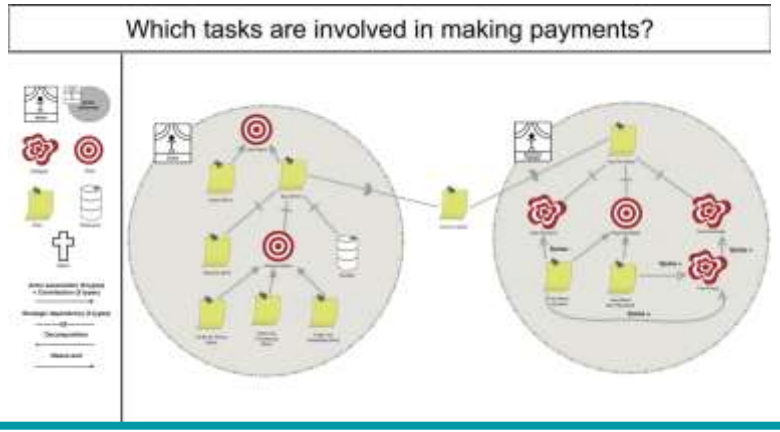


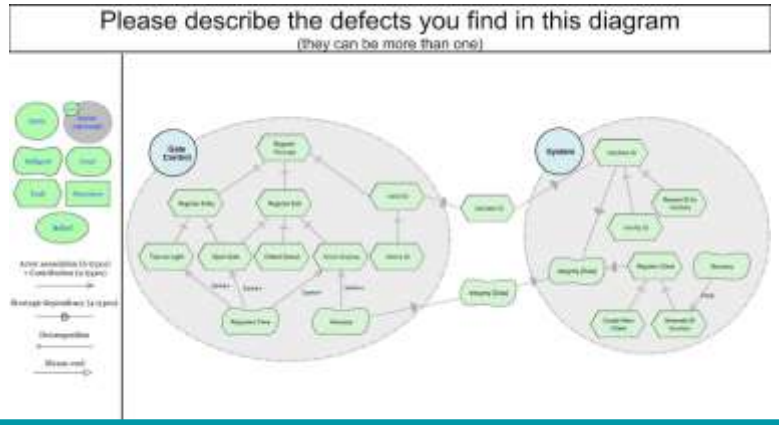


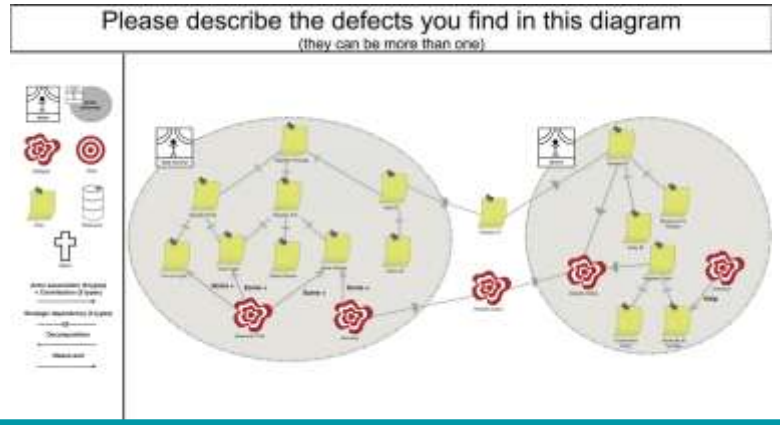


TWO UNDERSTANDING AND **TWO REVIEW** TASKS, BOTH WITH STANDARD *i** AND NEW *i**

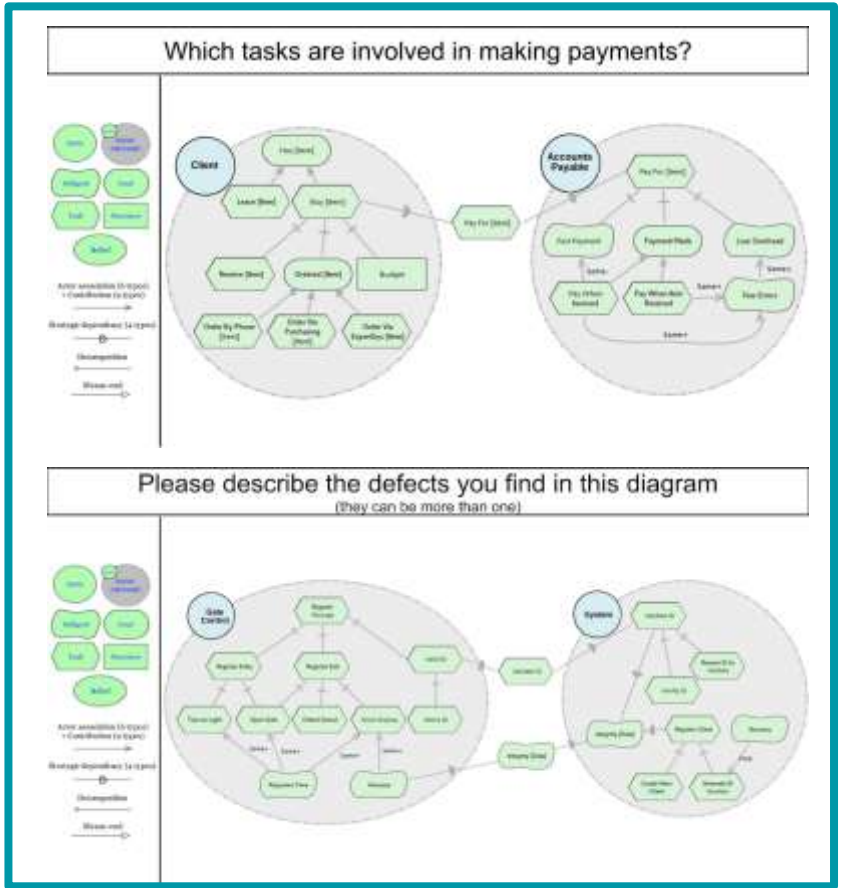


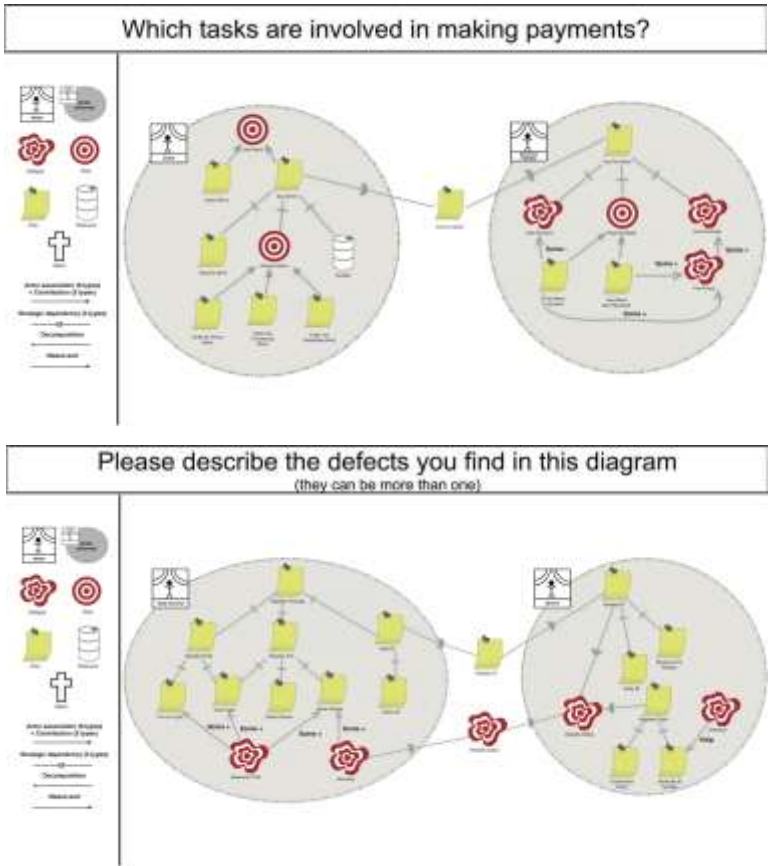






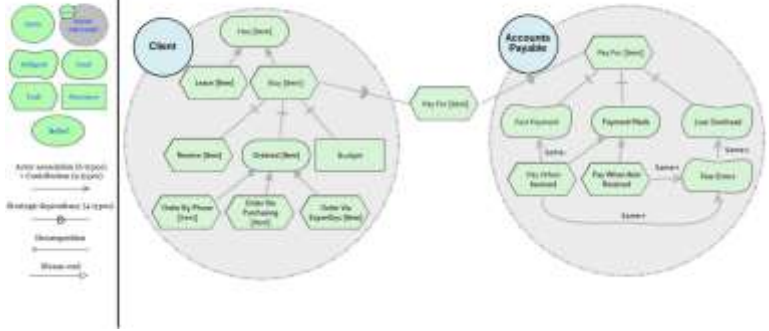
TWO UNDERSTANDING AND TWO REVIEW TASKS, BOTH WITH **STANDARD *i**** AND NEW *i**



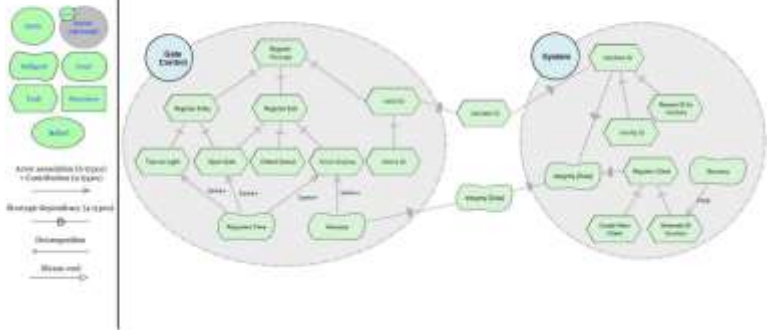


TWO UNDERSTANDING AND TWO REVIEW TASKS, BOTH WITH STANDARD *i** AND **NEW *i****

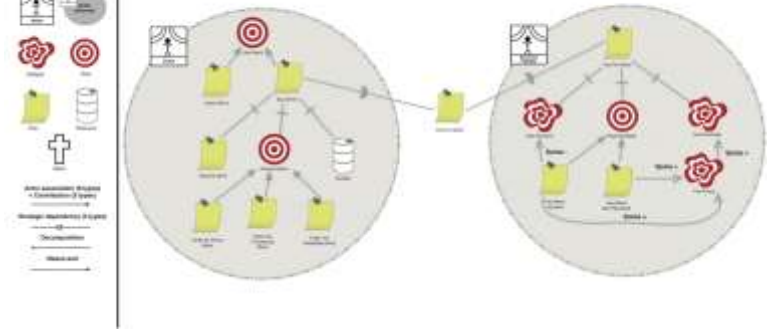
Which tasks are involved in making payments?



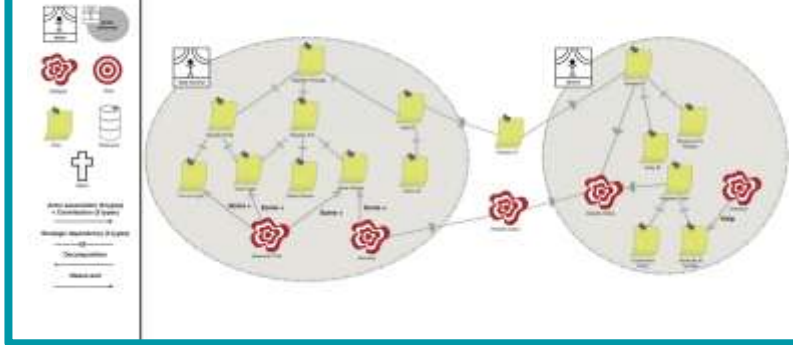
Please describe the defects you find in this diagram
(they can be more than one)



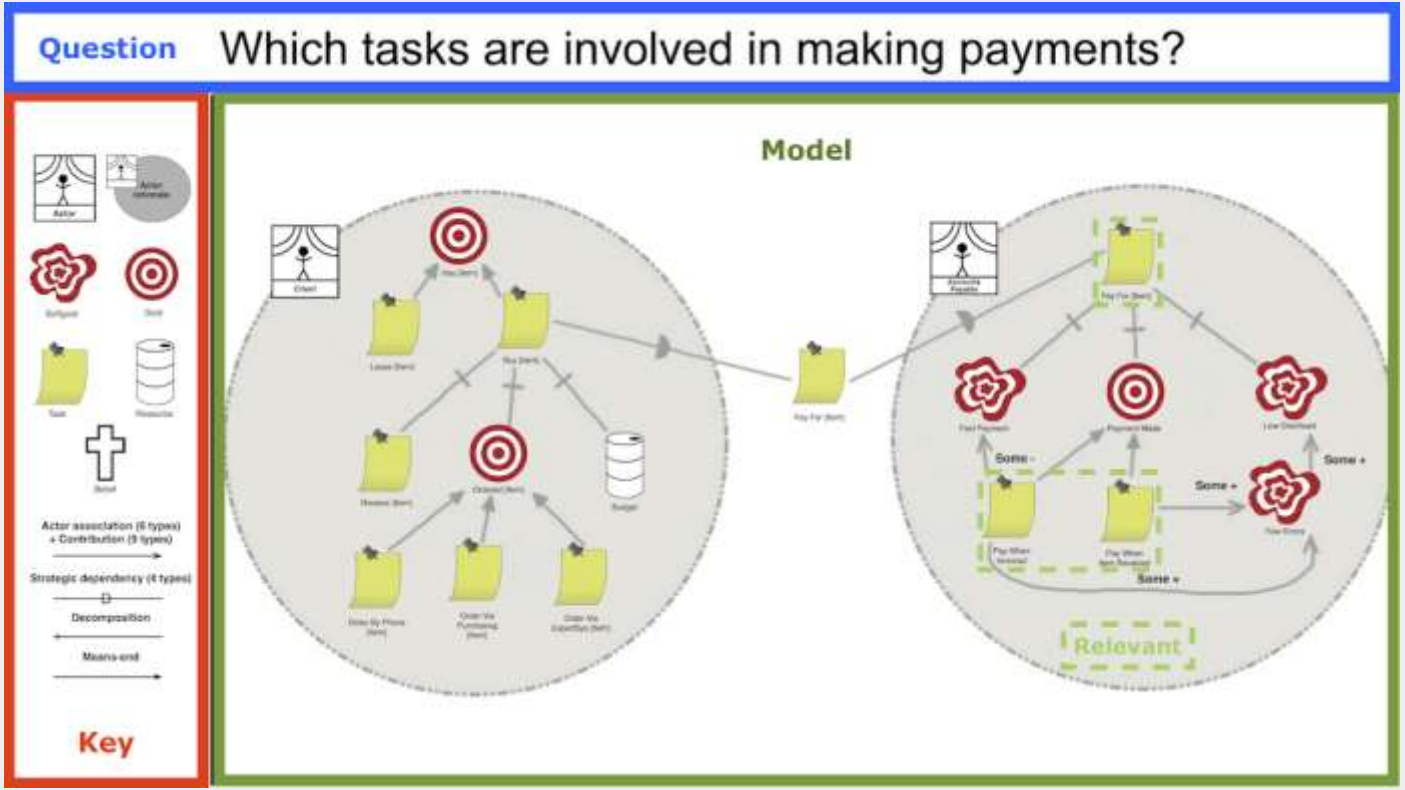
Which tasks are involved in making payments?

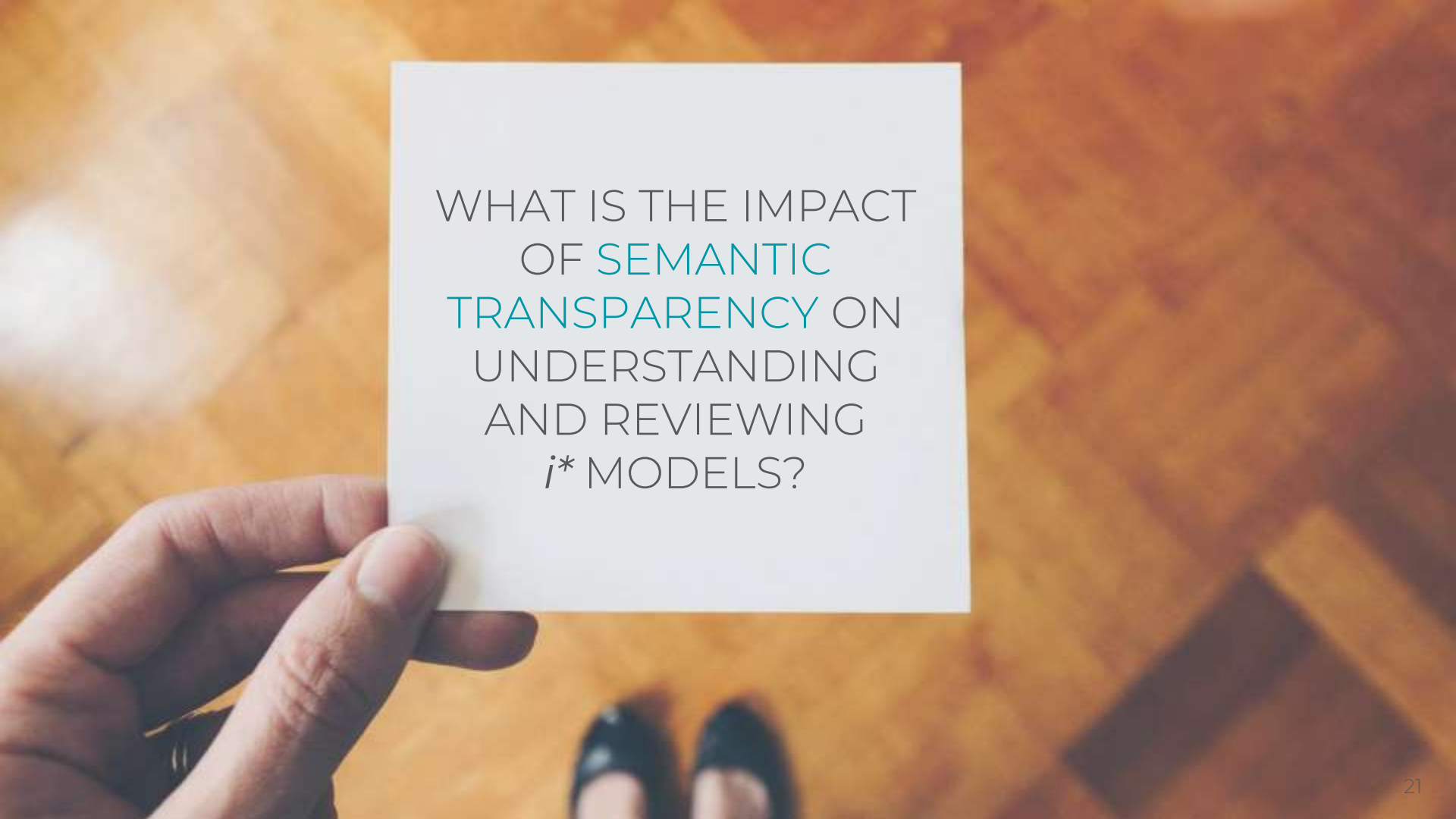


Please describe the defects you find in this diagram
(they can be more than one)

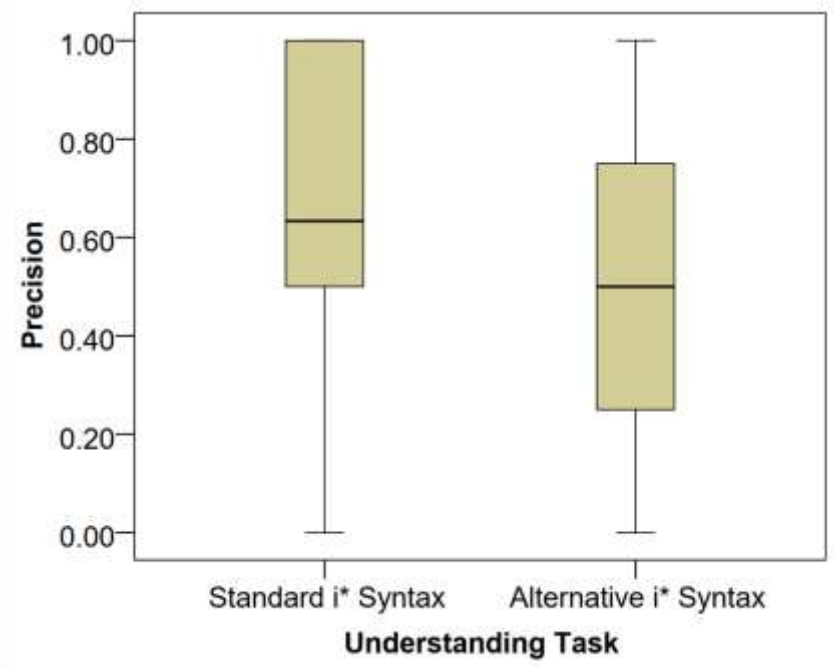


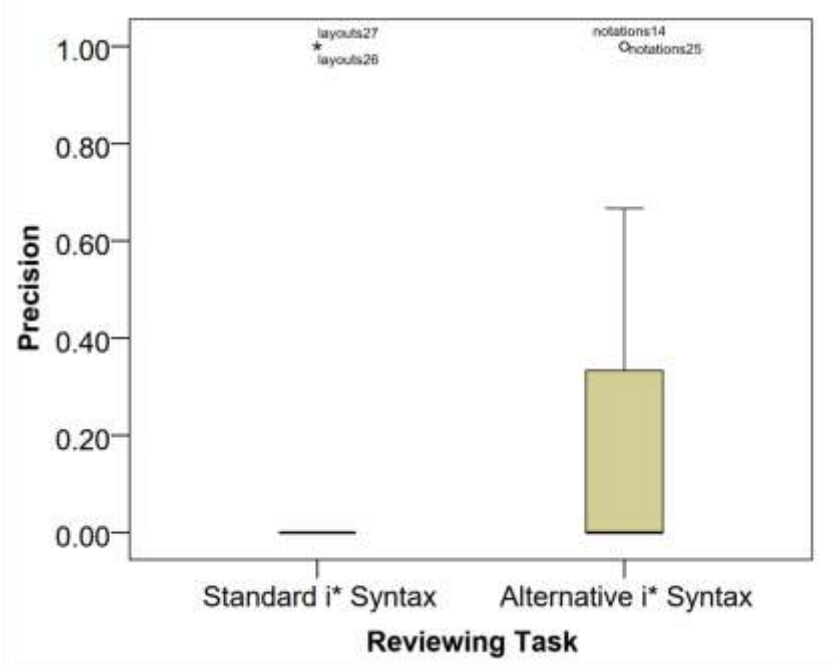
AREAS OF INTEREST



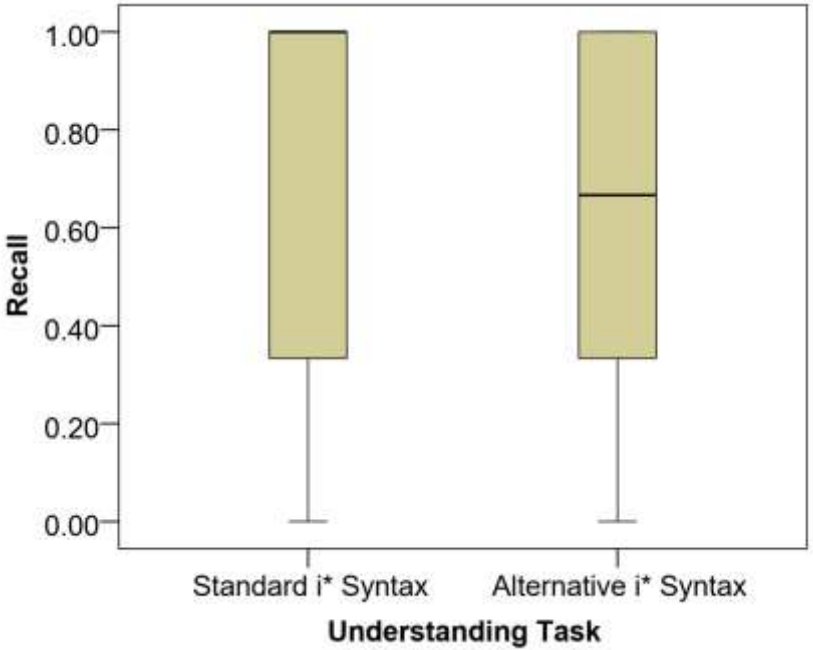

WHAT IS THE IMPACT OF SEMANTIC TRANSPARENCY ON UNDERSTANDING AND REVIEWING *i** MODELS?

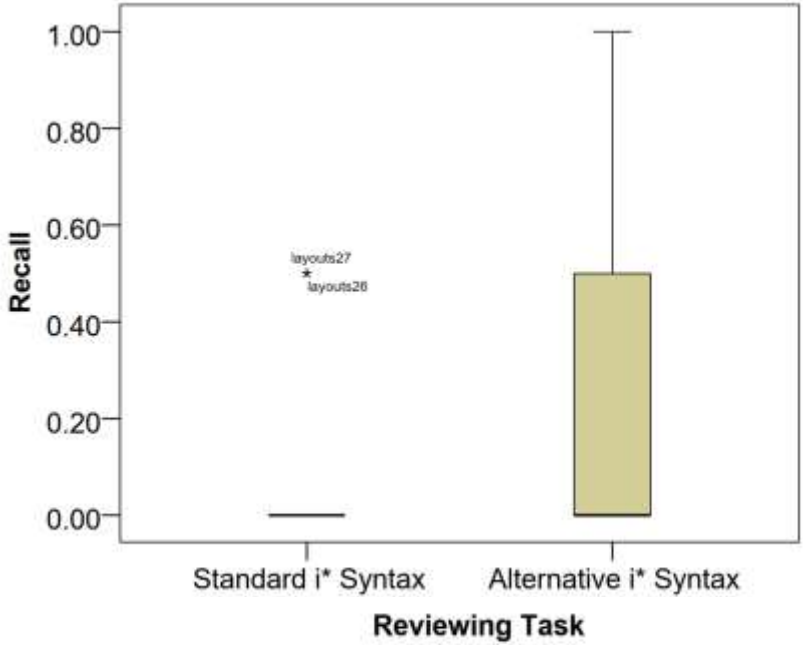
Precision is higher for understanding tasks, but there is no statistically significant difference between concrete syntaxes



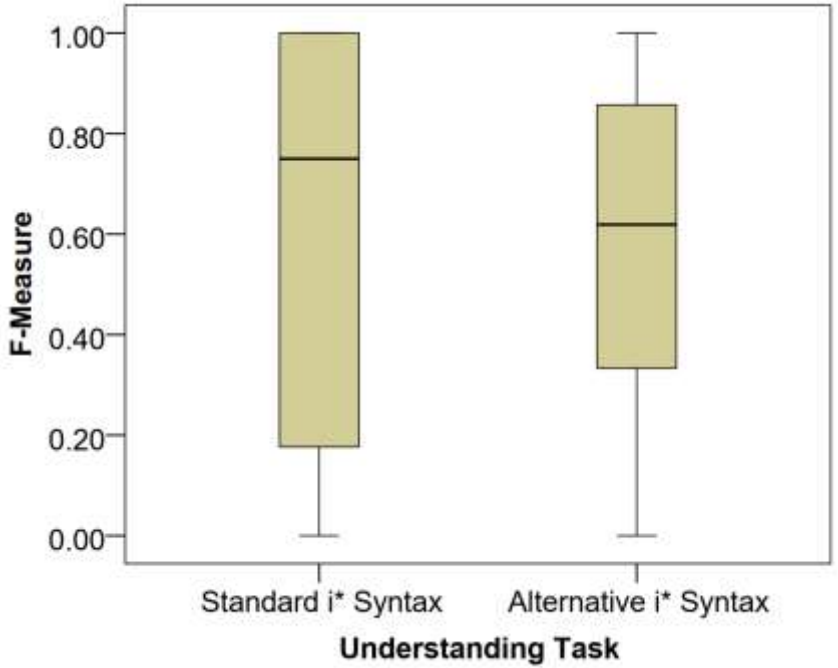


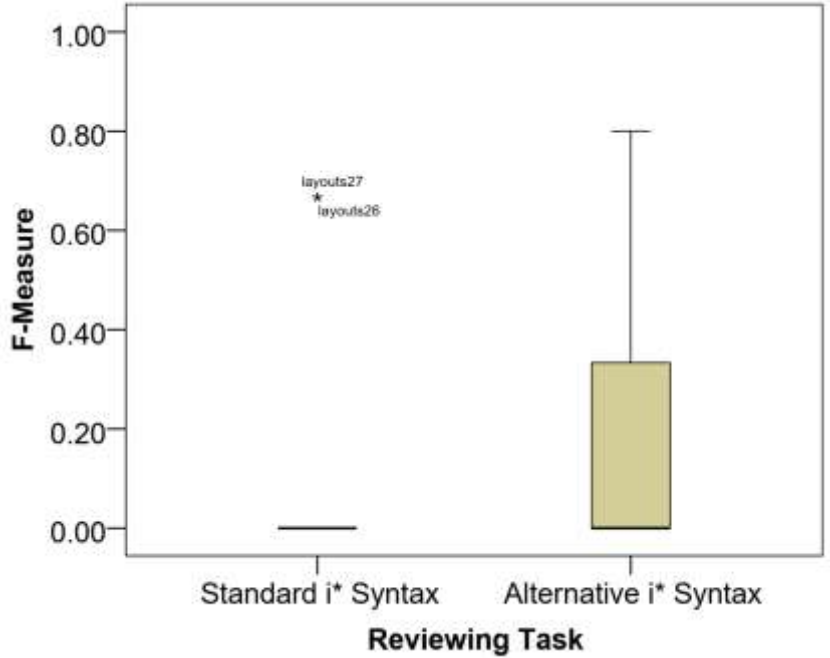
Recall is better for understanding tasks, but there is no statistically significant difference between concrete syntaxes



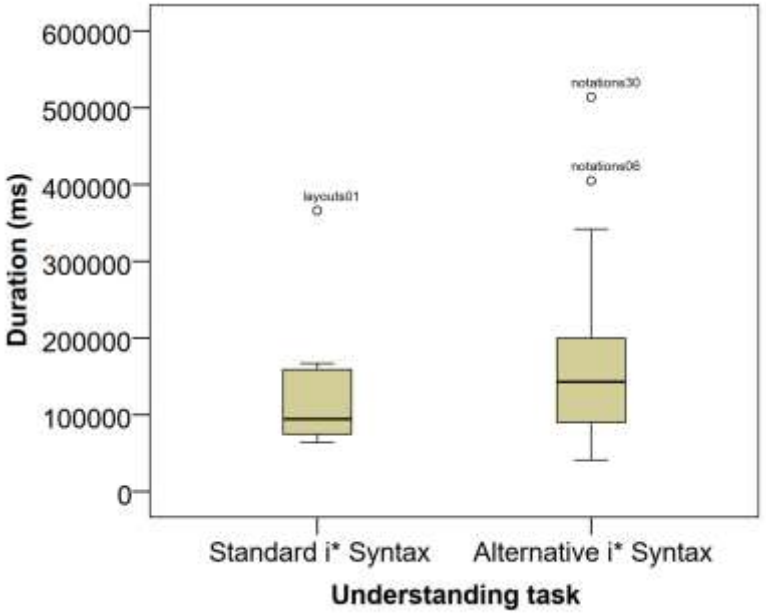


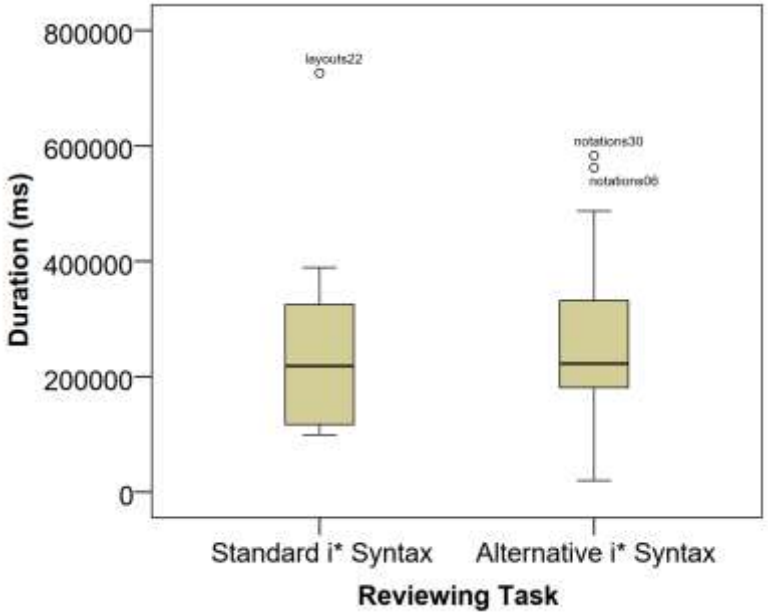
F-Measure is higher for understanding tasks, but there is no statistically significant difference between concrete syntaxes



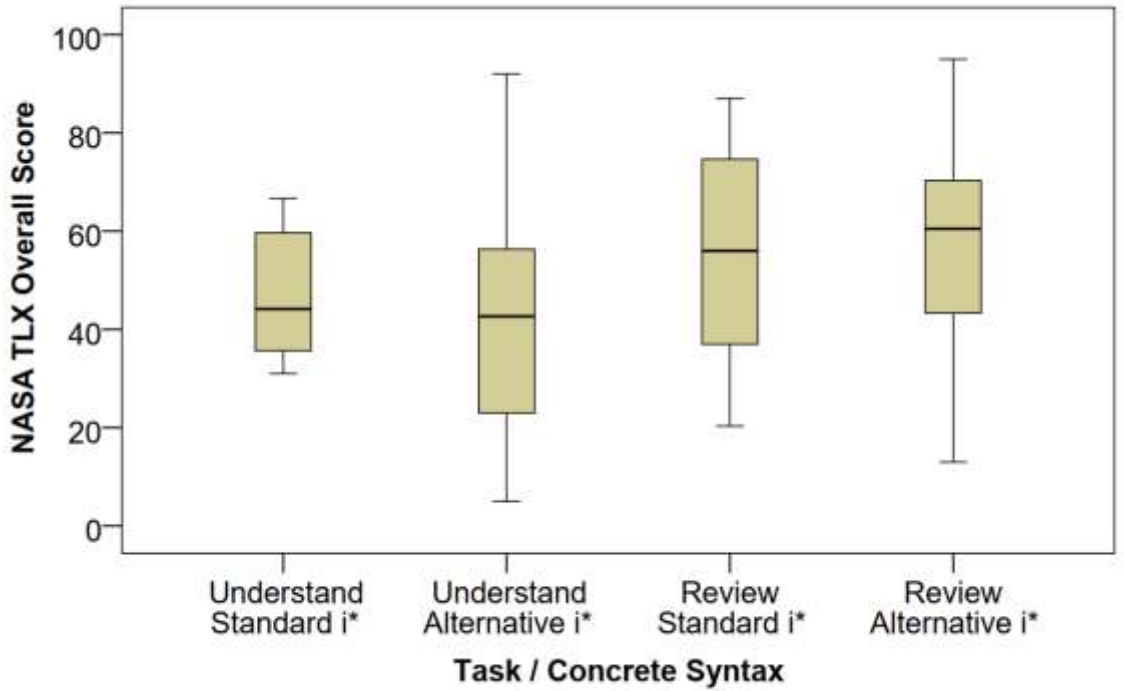


There is no difference in terms of duration, between concrete syntaxes for both tasks



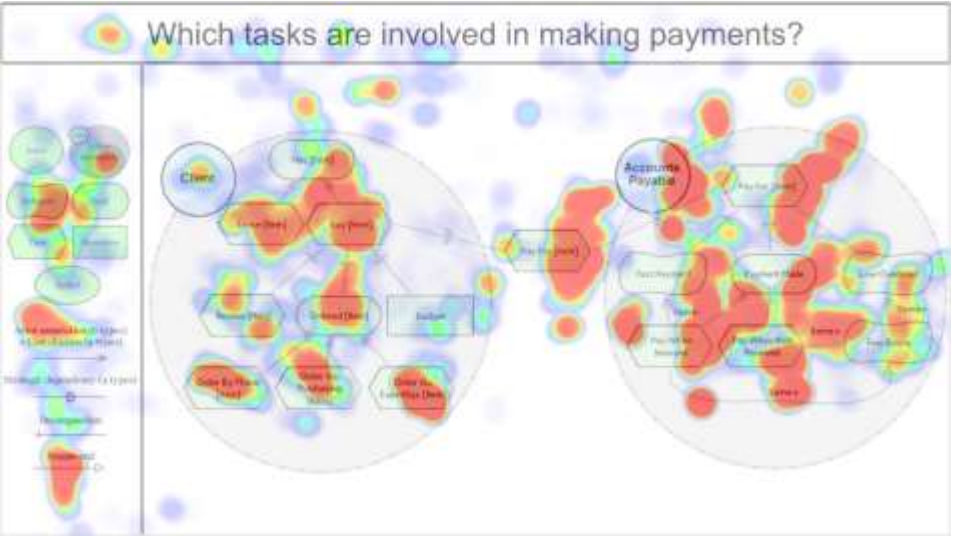


There is no difference in the perception of complexity of the tasks, for both concrete syntaxes

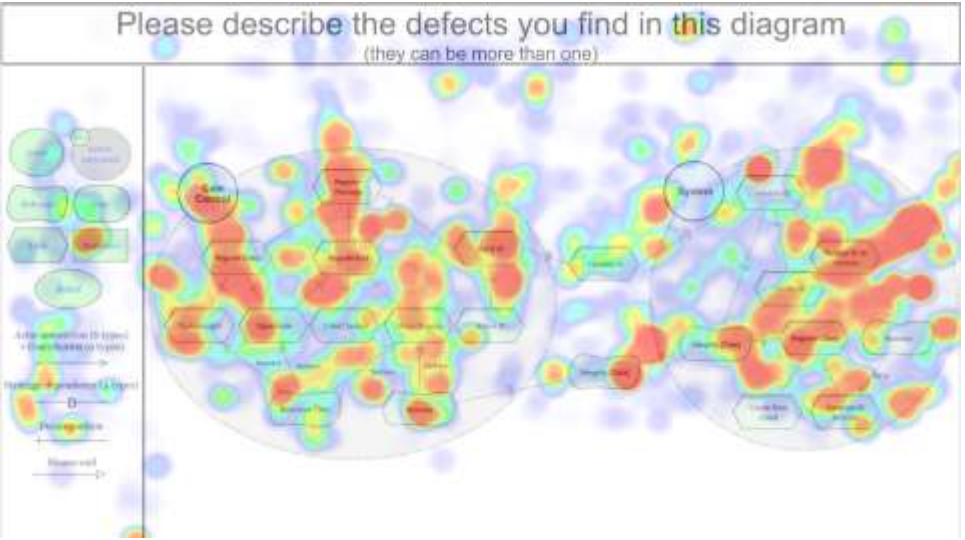


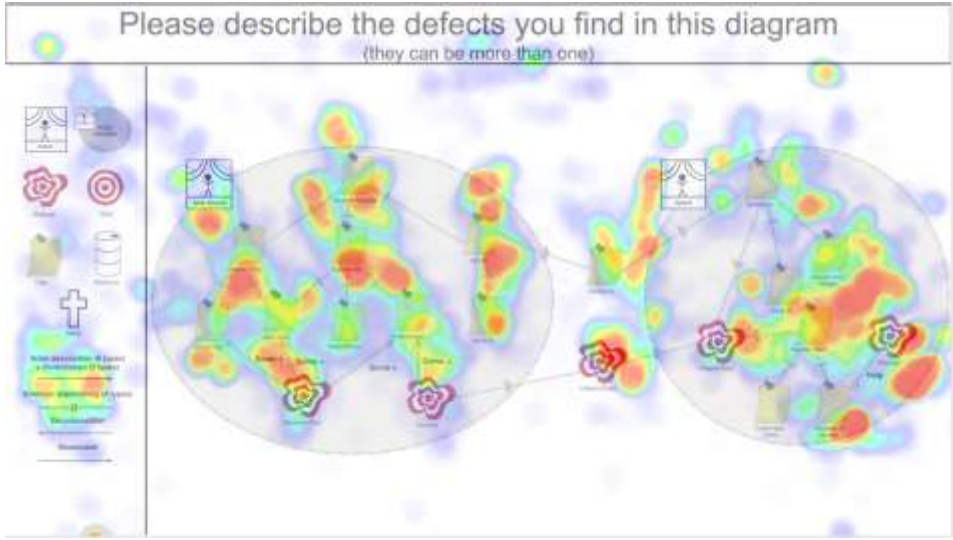
ARE THERE NO STATISTICALLY SIGNIFICANT DIFFERENCES?

Areas that are more frequently gazed during the understand tasks

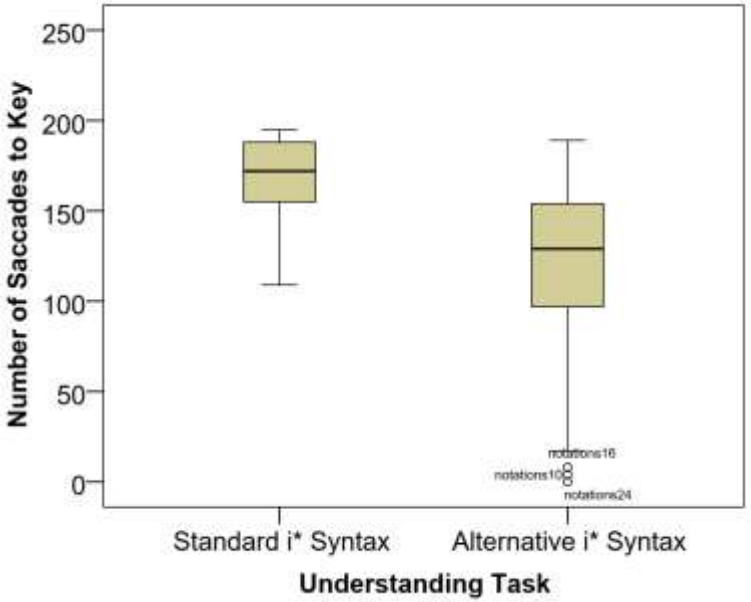


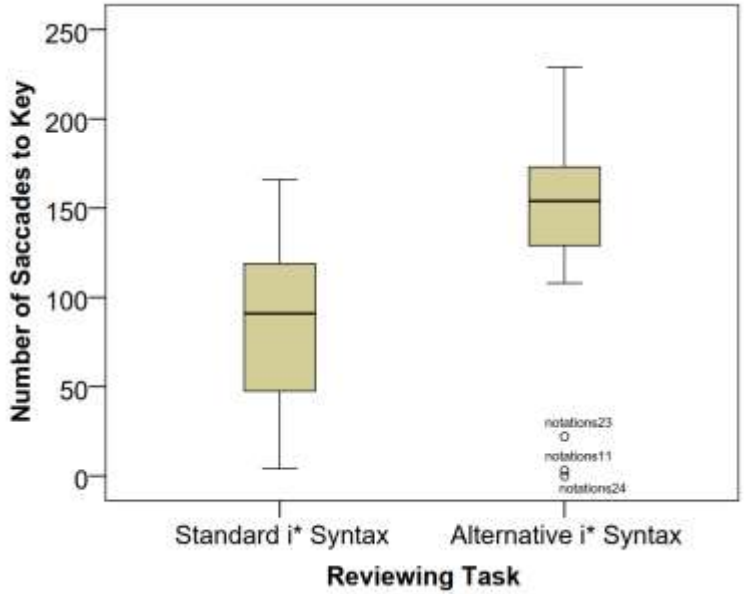
Areas that are more frequently gazed during the review tasks



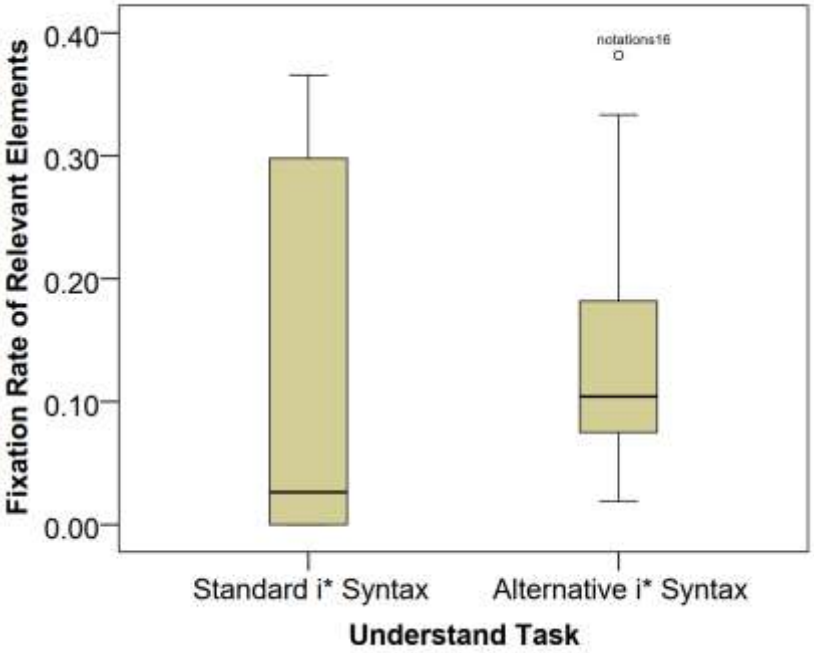


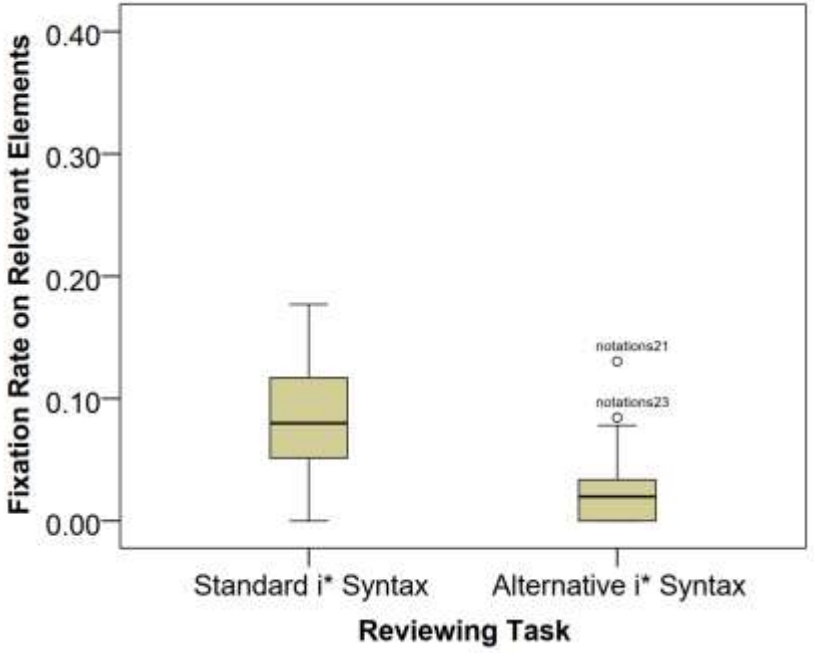
Total number of saccades and saccades to key are higher on understanding tasks for standard *i**, with a statistical significance



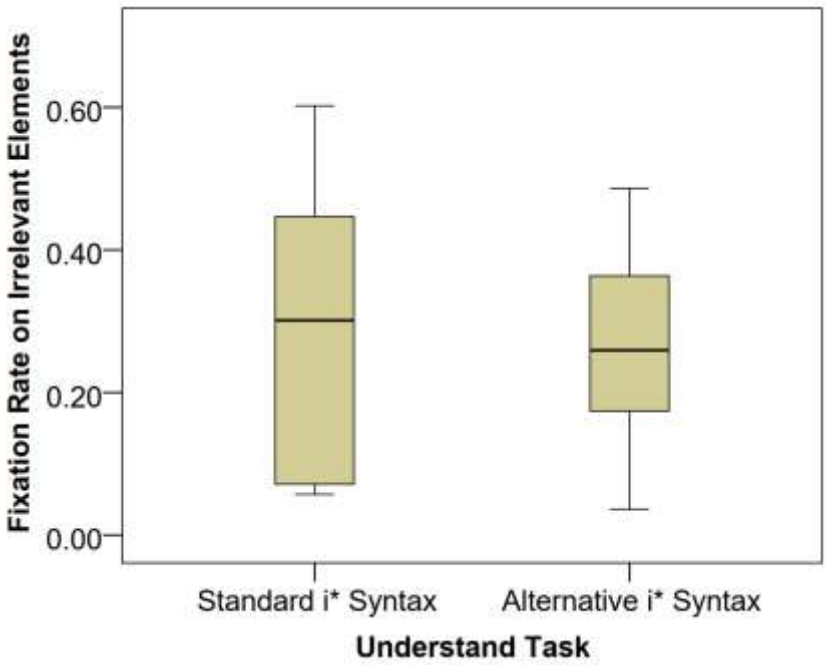


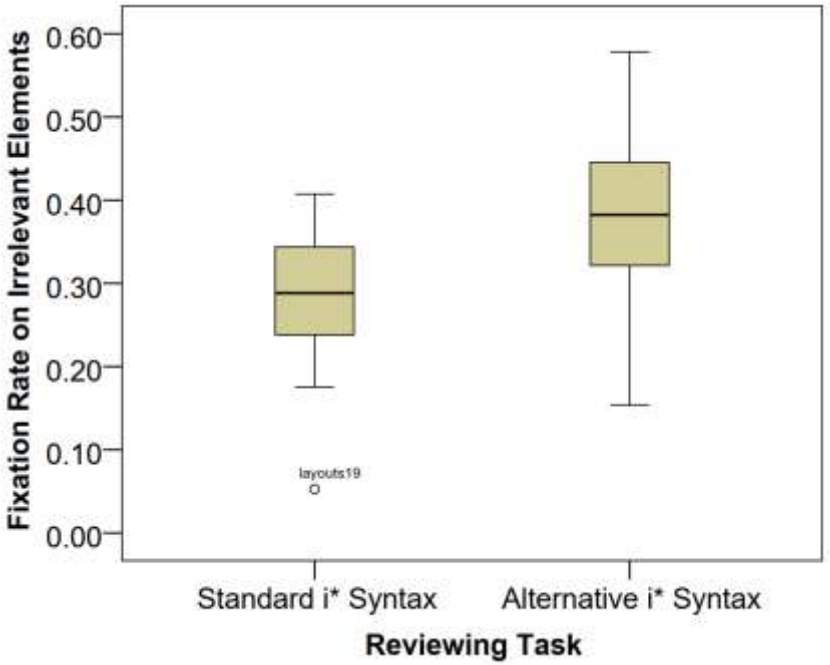
The effort spent looking at the relevant parts of the model decreased with the new *i**...



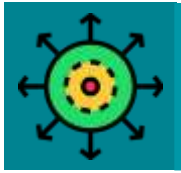
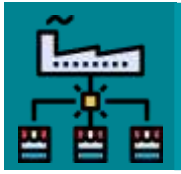


… but the effort on looking at irrelevant parts of the model increased, with the new *i**





THREATS TO VALIDITY

 conclusion	reasonable number of participants; facilitated independent replicas
 internal	mixed order of the tasks; participants used only one of the concrete syntaxes. Size of the models
 external	Participants knowledge. Models not representative
 construct	participants were not informed about what was being tested

INFERENCES

similar speed
and accuracy

no deep
overall impact
of visual effort

better symbol semantic transparency did not imply better model understanding when using the models, due to the context provided by the model, and, when available, the presence of a language key.

Future work

- Further studies should consider the various PoN principles, their interactions, their influence on the actual performance of practitioners in understanding and reviewing social goal models.
- It would be interesting to understand if the new concrete syntax has any drawback (e.g., in model construction) that hinders performance
- Also, why the NASA-TLX questionnaire results do not support the visual effort clear in the heat map, or still, understand the fixation time on relevant/ irrelevant AOIs and how they differ between the two groups of participants.
- It is necessary to assess how consistently our results occur with other users, models and concrete syntaxes
- We plan to replicate the experiment in other contexts, and apply it to bigger and more complex models

THANK YOU

QUESTIONS?