Professional Status and Expertise for UML Class Diagram Comprehension: An Empirical Study

Zéphyrin Soh, Zohreh Sharafi, Bertrand Van den Plas, Gerardo Cepeda Porras, Yann-Gaël Guéhéneuc and Giuliano Antoniol

Department of Computer and Software Engineering
École Polytechnique de Montréal, Québec, Canada

June 13, 2012
Outline

Problem and Motivations
  Problem
  Motivations

Related Work
  Expertise Studies
  UML Class Diagram Comprehension

Empirical Study
  Study Design

Results
  RQ1: Status
  RQ2: Expertise
  RQ3: Status vs. Expertise
  RQ4: Question Precision

Conclusion and Future Work
  Conclusion
  Threats to Validity and Future Work
Problem and Motivations

What is experience?

To manage subject/programmer experience:

- **Years** and **education** as main criteria [1]
- Authors sometime combine many criteria

![Diagram showing years and education as main criteria.]

Motivations

Consider two following cases:

- A student who used UML for 4 years during her study
- A professional with 3 years of experience with UML

Who is the best at understanding of UML class diagrams?
Motivations

- Consider two following cases:
  - A student who used UML for 4 years during her study
  - A professional with 3 years of experience with UML

Who is the best at understanding of UML class diagrams?

- Project managers when recruiting a new software designer by prioritized the important “factor”

- Future designers to know “where” to acquire the competitive skills by considering the important “factor”
Related Work

Expertise Studies (1/3)

Previous work on expertise

- Novices spend less time than experts analysts [2]

Related Work

Expertise Studies (1/3)

Previous work on expertise

- Novices spend less time than experts analysts [2]
- Graduate students are faster than junior ones and intermediate professionals [3]

### Related Work

**Expertise Studies (1/3)**

<table>
<thead>
<tr>
<th>Previous work on expertise</th>
</tr>
</thead>
<tbody>
<tr>
<td>▶ Novices spend less time than experts analysts [2]</td>
</tr>
<tr>
<td>▶ Graduate students are faster than junior ones and intermediate professionals [3]</td>
</tr>
<tr>
<td>▶ Experts are better for abstract questions and novices are better for concrete questions [4]</td>
</tr>
</tbody>
</table>

Related Work

Expertise Studies (1/3)

### Previous work on expertise

- Novices spend less time than experts analysts [2]
- Graduate students are faster than junior ones and intermediate professionals [3]
- Experts are better for abstract questions and novices are better for concrete questions [4]
- Experts and novices have different program model for documentation task, no difference for reuse task [5]

---

Related Work

Expertise Studies (2/3)

Comparison

To compare our work with previous work, we consider:

- Object
- Kind of task/question
- Subject categorisation criterion
## Related Work

### Expertise Studies (2/3)

#### Comparison

<table>
<thead>
<tr>
<th>Ref.</th>
<th>Textual Description</th>
<th>Requirements Analysis</th>
<th>Years of Experience Rating Scale of Supervisors</th>
</tr>
</thead>
<tbody>
<tr>
<td>[2]</td>
<td>textual description</td>
<td>requirements analysis</td>
<td>years of experience rating scale of supervisors</td>
</tr>
<tr>
<td>[3]</td>
<td>Java program</td>
<td>change task</td>
<td>students and professionals</td>
</tr>
<tr>
<td>[4]</td>
<td>program + flowcharts</td>
<td>abstract + concrete question</td>
<td>undergraduate students fellow teachers</td>
</tr>
<tr>
<td>[5]</td>
<td>database program</td>
<td>documentation and reuse</td>
<td>students and experts (nomination by colleagues, ...)</td>
</tr>
</tbody>
</table>

Legend:  Same to our work | Different to our work

Related Work
Expertise Studies (3/3)

Limitations

Previous work:

- Did not precisely distinguish years of experience and professionalism:
  - Inexperienced students as novices [5]
  - Senior professionals with less years of programming experience than graduate students [3]

Related Work

Expertise Studies (3/3)

Limitations

Previous work:

- Did not precisely distinguish years of experience and professionalism:
  - Inexperienced students as novices [5]
  - Senior professionals with less years of programming experience than graduate students [3]

We distinguish the years of experience from professionalism

Related Work

Expertise Studies (3/3)

Limitations

Previous work:

- Did not precisely distinguish years of experience and professionalism:
  - Inexperienced students as novices [5]
  - Senior professionals with less years of programming experience than graduate students [3]
  
  We distinguish the years of experience from professionalism

- Studied the source code or textual descriptions of requirements

Related Work

Expertise Studies (3/3)

Limitations

Previous work:

- Did not precisely distinguish years of experience and professionalism:
  - Inexperienced students as novices [5]
  - Senior professionals with less years of programming experience than graduate students [3]

We distinguish the years of experience from professionalism

- Studied the source code or textual descriptions of requirements

We use the UML class diagram

---

Related Work

UML Class Diagram Comprehension (1/2)

UML and eye-tracking

- **Stereotype, color, and layout** facilitate class diagram exploration and comprehension [6]
- Multi-cluster (by requirement) and three-cluster (by stereotype) **layout** positively affect the comprehension of class diagrams [7]

Related Work

UML Class Diagram Comprehension (1/2)

UML and eye-tracking

- **Stereotype, color, and layout** facilitate class diagram exploration and comprehension [6]
- **Multi-cluster (by requirement) and three-cluster (by stereotype) layout** positively affect the comprehension of class diagrams [7]
- **Canonical representation** of the Visitor pattern in class diagram reduce the effort of maintenance task [8]
- **The representations of design patterns** affect the identification of their participants and their roles [9]

---


Related Work

UML Class Diagram Comprehension (2/2)

Subjects’ categories

Previous work used subject’s proficiency as categorisation criterion:

- Subjects’ performance in task realization
- Subjects’ grade in the course they were enrolled
Related Work

UML Class Diagram Comprehension (2/2)

Subjects’ categories

Previous work used subject’s proficiency as categorisation criterion:
- Subjects’ performance in task realization
- Subjects’ grade in the course they were enrolled

Motivations

- No previous work that uses the maintenance task on UML class diagrams and eye-tracking system to study separately the professional status and the expertise
  - Combine expertise studies and UML eye-tracking studies
Empirical Study
Study Design (1/8)

Research Questions

▶ **RQ1:** What is the relation between a designer’s **professional status** and her class diagram comprehension?
Empirical Study

Study Design (1/8)

Research Questions

- **RQ1**: What is the relation between a designer's **professional status** and her class diagram comprehension?
- **RQ2**: What is the relation between a designer's **expertise** and her class diagram comprehension?
Empirical Study

Study Design (1/8)

Research Questions

- **RQ1**: What is the relation between a designer’s **professional status** and her class diagram comprehension?
- **RQ2**: What is the relation between a designer’s **expertise** and her class diagram comprehension?
- **RQ3**: What is the most important factor between **expertise** and **professional status**?
Empirical Study

Study Design (1/8)

Research Questions

- **RQ1**: What is the relation between a designer’s **professional status** and her class diagram comprehension?
- **RQ2**: What is the relation between a designer’s **expertise** and her class diagram comprehension?
- **RQ3**: What is the most important factor between **expertise** and **professional status**?
- **RQ4**: What is the effect of the **question precision** on the comprehension of a UML class diagram?
## Empirical Study

### Study Design (2/8)

### Objects and Tasks

<table>
<thead>
<tr>
<th></th>
<th>ArgoUML</th>
<th>JUnit</th>
<th>QuickUML</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of classes/Interfaces</td>
<td>10</td>
<td>14</td>
<td>16</td>
</tr>
<tr>
<td>Average number of attributes per Class/Interface</td>
<td>0.4</td>
<td>0.57</td>
<td>1.75</td>
</tr>
<tr>
<td>Average number of methods per Class/Interface</td>
<td>8.6</td>
<td>6.14</td>
<td>3.87</td>
</tr>
</tbody>
</table>

= : ArgoUML, JUnit, and QuickUML

: one maintenance task per object
Empirical Study

Study Design (3/8)

Independent variables

= Professional status + Expertise
Empirical Study

Study Design (3/8)

Independent variables

\[ \text{Independent variables} = \text{Professional status} + \text{Expertise} \]

- **Professional status**
  - practitioners (9) (in industry)
  - students (12)

- **Expertise**
  - We used the number of years of experience to categorise experts and novices.

- Pair-wise Wilcoxon comparison (+ Bonferroni correction)
- Categorization with the highest Cliff’s \( \delta \) value
  - experts (12): \{3, 4, 5\} years of experience
  - novices (9): \{1, 2\} years of experience

\[ 12 / 24 \]
Empirical Study

Study Design (3/8)

Independent variables

\[ \text{Independent variables} = \text{Professional status} + \text{Expertise} \]

- **Professional status**
  - = practitioners (9)
  - (in industry)
  - = students (12)

- **Expertise**: We used the number of years of experience to categorise experts and novices.
  - Pair-wise Wilcoxon comparison (+ Bonferroni correction)
  - Categorization with the highest Cliff’s \( \delta \) value
    - = experts (12): \( \{3, 4, 5\} \) years of experience
    - = novices (9): \( \{1, 2\} \) years of experience
Mitigating variable

**Question precision**: The level of details in the formulation of the question:

- **Precise**: state the kind of operation to perform (add/remove) and the kind of target element (class/method/attribute)
- **Not precise**: no operation or target element
Empirical Study

Study Design (5/8)

Dependent variables

- **Accuracy, Time spent**
- **Search effort** = convex hull & spatial density [10]
- **Overall effort** = AFD [9] and NRRF [8]
- **Question comprehension effort** = NDQA and NFQA

AFD: Average Fixation Duration
NRRF: Normalized Rate of Relevant Fixations
NDQA: Normalized Duration in Question Area
NFQA: Normalized Fixations in Question Area

Convex hull area

- Smaller convex set of fixations containing all subject’s fixations
- Smaller convex hull ⇒ close fixations ⇒ less search effort
Empirical Study

Study Design (7/8)

Spatial density

- Number of visited cells / total number of cells
- less visits ⇒ less search effort

In TAUPE [11], cell’s size = 64x64px

Empirical Study

Study Design (8/8)

Overall effort: Fixations’ duration and relevance

We want to add a class named "Constraint" capable of adding some items to the todo list of a designer. How would you do that?

Be specific about the classes / methods / attributes.

Professional status vs. Expertise

Zéphyrin Soh et al.

Problem and Motivations

Problem

Motivations

Related Work

Expertise Studies

UML Class Diagram

Comprehension

Empirical Study

Study Design

Results

RQ1: Status

RQ2: Expertise

RQ3: Status vs. Expertise

RQ4: Question Precision

Conclusion and Future Work

Conclusion

Threats to Validity and Future Work

17 / 24
Empirical Study
Study Design (8/8)
Question Comprehension Effort: Fixations’ count and duration
What is the relation between a designer’s *professional status* and her class diagram comprehension?

Practitioners are more accurate than students
Results

What is the relation between a designer’s *professional status* and her class diagram comprehension?

---

Practitioners are more accurate than students

Students spent around 35% less time than practitioners
Results

**RQ1: Status (1/1)**

What is the relation between a designer’s **professional status** and her class diagram comprehension?

- Practitioners are more accurate than students
- Students spent around 35% less time than practitioners
- No significant difference for other dependent variables
Results

RQ1: Status (1/1)

What is the relation between a designer’s professional status and her class diagram comprehension?

Practitioners are more accurate than students

Students spent around 35% less time than practitioners

- No significant difference for other dependent variables
- Students could be more accurate if spending more time
Professional status vs. Expertise

Zéphyrin Soh et al.

Problem and Motivations
Problem
Motivations

Related Work
Expertise Studies
UML Class Diagram Comprehension

Empirical Study
Study Design

Results
RQ1: Status
RQ2: Expertise (1/1)
RQ3: Status vs. Expertise
RQ4: Question Precision

Conclusion and Future Work
Conclusion
Threats to Validity and Future Work

Results

RQ2: Expertise (1/1)

What is the relation between a designer’s expertise and her class diagram comprehension?

Experts are more accurate than novices
Results

RQ2: Expertise (1/1)

Experts are more accurate than novices
Novices spent around 33% less time than experts
## Results

**RQ2: Expertise (1/1)**

What is the relation between a designer’s *expertise* and her class diagram comprehension?

<table>
<thead>
<tr>
<th>Experts</th>
<th>Novices</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Medal" /></td>
<td><img src="image" alt="Medal" /></td>
</tr>
<tr>
<td><img src="image" alt="Eye" /></td>
<td><img src="image" alt="Clock" /></td>
</tr>
</tbody>
</table>

- Experts are more accurate than novices
- Novices spent around 33% less time than experts
- Experts have a more efficient ability to search relevant elements than novices

---

**Experts are more accurate than novices**

**Novices spent around 33% less time than experts**

**Experts have a more efficient ability to search relevant elements than novices**
Results

RQ2: Expertise (1/1)

Experts are more accurate than novices. Novices spent around 33% less time than experts. Experts have a more efficient ability to search relevant elements than novices. No significant difference for other dependent variables.

What is the relation between a designer’s expertise and her class diagram comprehension?
**Results**

**RQ2: Expertise (1/1)**

What is the relation between a designer’s *expertise* and her class diagram comprehension?

- **Experts are more accurate than novices**
- **Novices spent around 33% less time than experts**
- **Experts have a more efficient ability to search relevant elements than novices**
- **No significant difference for other dependent variables**
- **Novices could be more accurate if spending more time**
Results

RQ3: Status vs. Expertise (1/1)

What is the most important factor between expertise and professional status?

Experts are more accurate than practitioners
Results

RQ3: Status vs. Expertise (1/1)

Experts are more accurate than practitioners
Experts spent around 7% less time than practitioners

What is the most important factor between expertise and professional status?
Results

RQ3: Status vs. Expertise (1/1)

What is the most important factor between expertise and professional status?

Experts are more accurate than practitioners

Experts spent around 7% less time than practitioners

When considering expert subjects
Results

RQ3: Status vs. Expertise (1/1)

Experts are more accurate than practitioners

Experts spent around 7% less time than practitioners

- When considering expert subjects
  - Experienced students are more accurate than experienced practitioners
Results
RQ3: Status vs. Expertise (1/1)

Experts are more accurate than practitioners
Experts spent around 7% less time than practitioners
▶ When considering expert subjects
  ▶ Experienced students are more accurate than experienced practitioners
  ▶ Experienced students spent around 37% less time than experienced practitioners
Results

RQ3: Status vs. Expertise (1/1)

Experts are more accurate than practitioners

Experts spent around 7% less time than practitioners

When considering expert subjects

- Experienced students are more accurate than experienced practitioners
- Experienced students spent around 37% less time than experienced practitioners
- The effects of expertise on accuracy and time depend on the status

What is the most important factor between expertise and professional status?
Results

RQ4: Question Precision (1/1)

Question Precision

What is the effect of the question precision on the comprehension of a UML class diagram?

- The accuracy of students benefits from precise question description
- The accuracy of novices benefits from precise question description
Conclusion and Future Work

Status

---

Conclusion (1/1)

Threats to Validity and Future Work
Conclusion and Future Work

Conclusion (1/1)

Status

Expertise

Empirical Study

Study Design

Related Work

Expertise Studies
UML Class Diagram
Comprehension

Conclusion and Future Work

Conclusion

Threats to Validity and Future Work

Results

RQ1: Status
RQ2: Expertise
RQ3: Status vs. Expertise
RQ4: Question Precision

Professional status vs. Expertise

Zéphyrin Soh et al.

Problem and Motivations

Problem
Motivations
Conclusion and Future Work

Status vs. Expertise

Problem and Motivations
- Problem
- Motivations

Related Work
- Expertise Studies
- UML Class Diagram Comprehension

Empirical Study
- Study Design

Results
- RQ1: Status
- RQ2: Expertise
- RQ3: Status vs. Expertise
- RQ4: Question Precision

Conclusion and Future Work

Conclusion
- Threats to Validity and Future Work
Conclusion and Future Work

Threats to Validity and Future Work (1/1)

Threats to Validity and Future Work

- **Construct validity**: We did not use all combination of treatments for each system.
Conclusion and Future Work

Threats to Validity and Future Work (1/1)

Threats to Validity and Future Work

- **Construct validity**: We did not use all combination of treatments for each system

- **Conclusion validity**: Practitioners from the same company + difficulty to find inexperienced practitioners (only one)
  $\Rightarrow$ Practitioners from other company

Conclusion and Future Work

Conclusion

Threats to Validity and Future Work
## Conclusion and Future Work

### Threats to Validity and Future Work

- **Construct validity**: We did not use all combination of treatments for each system.

- **Conclusion validity**: Practitioners from the same company + difficulty to find inexperienced practitioners (only one) ⇒ Practitioners from other company.

- **Internal validity**: We did not limit the time (fatigue bias) ⇒ Limit the experiment time to investigate how much time affect the subject’s accuracy.
Conclusion and Future Work

Threats to Validity and Future Work (1/1)

Threats to Validity and Future Work

- **Construct validity**: We did not use all combination of treatments for each system

- **Conclusion validity**: Practitioners from the same company + difficulty to find inexperienced practitioners (only one)
  ⇒ Practitioners from other company

- **Internal validity**: We did not limit the time (fatigue bias)
  ⇒ Limit the experiment time to investigate how much time affect the subject’s accuracy

- **External validity**: Only three systems and small range of years of experience
  ⇒ Use other systems
Thanks for your attention!

The accuracy of students and novices benefits from precise question descriptions.

Experts vs. Practitioners

Status

Expertise

<table>
<thead>
<tr>
<th>Time spent (s)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>40</td>
</tr>
<tr>
<td>200</td>
<td>60</td>
</tr>
<tr>
<td>300</td>
<td>80</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Time spent (s)</th>
<th>Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>40</td>
</tr>
<tr>
<td>200</td>
<td>60</td>
</tr>
<tr>
<td>300</td>
<td>80</td>
</tr>
</tbody>
</table>

The accuracy of students and novices benefits from precise question descriptions.

Experts vs. Practitioners